CONTEMPORARY MATHEMATICS

520

Algorithmic Probability and Combinatorics

AMS Special Sessions on
Algorithmic Probability and Combinatorics
October 5–6, 2007
DePaul University
Chicago, Illinois

October 4–5, 2008
University of British Columbia
Vancouver, BC, Canada

Manuel E. Lladser
Robert S. Maier
Marni Mishna
Andrew Rechnitzer
Editors

American Mathematical Society

Algorithmic Probability and Combinatorics

This page intentionally left blank

CONTEMPORARY MATHEMATICS

520

Algorithmic Probability and Combinatorics

AMS Special Sessions on
Algorithmic Probability and Combinatorics
October 5–6, 2007
DePaul University
Chicago, Illinois

October 4–5, 2008
University of British Columbia
Vancouver, BC, Canada

Manuel E. Lladser
Robert S. Maier
Marni Mishna
Andrew Rechnitzer
Editors

American Mathematical Society
Providence, Rhode Island

2000 *Mathematics Subject Classification.* Primary 05–06, 60–06, 41–06, 82–06; Secondary 05A15, 05A16, 60C05, 41A60.

Library of Congress Cataloging-in-Publication Data

AMS Special Session on Algorithmic Probability and Combinatorics (2007 : DePaul University)

Algorithmic probability and combinatorics : AMS Special Session, October 5-6, 2007, DePaul University, Chicago, Illinois : AMS Special Session, October 4-5, 2008, University of British Columbia, Vancouver, BC, Canada / Manuel E. Lladser ... [et al.], editors.

p. cm. – (Contemporary mathematics ; v. 520)

Includes bibliographical references.

ISBN 978-0-8218-4783-1 (alk. paper)

1. Combinatorial analysis—Congresses. 2. Approximation theory—Congresses. 3. Mathematical statistics—Congresses. I. Lladser, Manuel, 1970- II. AMS Special Session on Algorithmic Probability and Combinatorics (2008 : University of British Columbia) III. Title.

QA164.A474 2007
511′.6—dc22

2010011434

Contents

This page intentionally left blank

Preface

This volume contains refereed articles by speakers in the AMS Special Sessions on Algorithmic Probability and Combinatorics, held on October 5–6, 2007 at DePaul University in Chicago, IL, and on October 4–5, 2008 at the University of British Columbia in Vancouver, BC. The articles cover a wide range of topics in analytic combinatorics and in the study, both analytic and computational, of combinatorial probabilistic models. The authors include pure mathematicians, applied mathematicians, and computational physicists. A few of the articles have an expository flavor, with extensive bibliographies, but original research predominates.

This is the first volume that the AMS has published in this interdisciplinary area. Our hope is that these articles give an accurate picture of its variety and vitality, and its ties to other areas of mathematics. These areas include asymptotic analysis, algebraic geometry, special functions, the analysis of algorithms, statistical mechanics, stochastic simulation, and importance sampling.

As co-organizers and co-editors, we thank all participants, contributors, and referees. We are grateful to the American Mathematical Society for assistance in organizing the special sessions, and in the publication of this volume. We especially thank Christine Thivierge of the AMS staff, for her efficient support in the latter.

Manuel E. Lladser
Robert S. Maier
Marni Mishna
Andrew Rechnitzer

This page intentionally left blank

Contemporary Mathematics
Volume **520**, 2010

Walks with small steps in the quarter plane

Mireille Bousquet-Mélou and Marni Mishna

Abstract. Let $\mathcal{S} \subset \{-1,0,1\}^2 \setminus \{(0,0)\}$. We address the enumeration of plane lattice walks with steps in $\mathcal{S}$, that start from $(0,0)$ and remain in the first quadrant $\{(i,j) : i \geqslant 0, j \geqslant 0\}$. *A priori*, there are 2^8 models of this type, but some are trivial. Some others are equivalent to models of walks confined to a half-plane, and can therefore be treated systematically using the kernel method, which leads to a generating function that is algebraic.

We focus on the remaining models, and show that there are 79 inherently different ones. To each of the 79, we associate a group G of birational transformations. We show that this group is finite (in fact dihedral, and of order at most 8) in 23 cases, and is infinite in the other 56 cases. We present a unified way of dealing with 22 of the 23 models associated with a finite group. For each, we find the generating function to be D-finite; and in some cases, algebraic. The 23rd model, known as Gessel's walks, has recently been proved by Bostan *et al.* to have an algebraic (and hence D-finite) generating function. We conjecture that the remaining 56 models, each associated with an infinite group, have generating functions that are non-D-finite.

Our approach allows us to recover and refine some known results, and also to obtain new ones. For instance, we prove that walks with N, E, W, S, SW and NE steps yield an algebraic generating function.

1. Introduction

The enumeration of lattice walks is a classic topic in combinatorics. Many combinatorial objects (trees, maps, permutations, lattice polygons, Young tableaux, queues...) can be encoded as lattice walks, so that lattice path enumeration has many applications. Given a lattice, for instance the hypercubic lattice $\mathbb{Z}^d$, and a finite set of steps $\mathcal{S} \subset \mathbb{Z}^d$, a typical problem is to determine how many n-step walks with steps taken from $\mathcal{S}$, starting from the origin, are confined to a certain region $\mathcal{A}$ of the space. If $\mathcal{A}$ is the whole space, then the length generating function of these walks is a simple rational series. If $\mathcal{A}$ is a half-space, bounded by a rational hyperplane, then the associated generating function is an algebraic series. Instances of the latter problem have been studied in many articles since at least the end of the 19^{th} century [**1, 6**]. It is now understood that the *kernel method* provides a

2000 *Mathematics Subject Classification.* Primary 05A15.

MBM was supported by the French "Agence Nationale de la Recherche," project SADA ANR-05-BLAN-0372.

MM was supported by a Canadian NSERC Discovery grant.

systematic solution to all such problems, which are, in essence, one-dimensional [**2, 14**]. Other generic approaches to half-space problems are provided in [**20, 25**].

A natural next class of problems is the enumeration of walks constrained to lie in the intersection of two rational half-spaces: typically, in the quarter plane. Thus far, a number of instances have been solved, but no unified approach has yet emerged, and the problem is far from being completely understood. The generating functions that have been found exhibit a more complicated structure than those of half-space problems. Some are algebraic, but for reasons that are poorly understood combinatorially [**9, 13, 26, 40**]. Some are, more generally, D-finite, meaning that the generating function satisfies a linear differential equation with polynomial coefficients [**11, 15, 31, 47**]. Some are not D-finite, having infinitely many singularities in the complex plane [**15, 41**].

We focus in this paper on walks in the plane confined to the first quadrant. The four examples of Figure 1 illustrate the complexity of this problem:

- **Kreweras' walks (steps W, S, and NE)**: These were first counted in 1965 by Kreweras [**37**]. He obtained a complicated expression for the number of n-step walks ending at (i, j), which simplifies drastically when $j = 0$. It was then proved by Gessel that the associated 3-variable generating function (which counts all quarter plane walks by the length and the coordinates of the endpoint) is algebraic [**26**]. Since then, simpler derivations of this series have been obtained [**13, 42**], including an automated proof [**36**], and a purely bijective one for walks ending at the origin [**4**]. See also [**13, 21, 23**], where the stationary distribution of a related Markov chain in the quarter plane is obtained and found to be algebraic.
- **Gessel's walks (steps E, W, NE and SW)**: Around 2001, Gessel conjectured a simple hypergeometric formula for the number of n-step walks ending at the origin. This conjecture was proved recently by Kauers, Koutschan and Zeilberger [**35**]. Even more recently, Bostan and Kauers proved that the associated 3-variable generating function is in fact algebraic [**9**]. Strangely enough, the simple numbers conjectured by Gessel had not been recognized as the coefficients of an algebraic series before. Both approaches involve, among other tools, heavy computer algebra calculations.
- **Gouyou-Beauchamps's walks (steps E, W, NW and SE)**: Gouyou-Beauchamps discovered in 1986 a simple hypergeometric formula for walks ending on the x-axis [**29**]. We derive in this paper similar expressions for the total number of walks, and for those ending at a prescribed position. The associated series are D-finite, but transcendental. These walks are related to Young tableaux of height at most 4 [**30**]. An affine deformation transforms them into square lattice walks (with N, S, E and W steps) confined to the wedge $0 \leqslant j \leqslant i$. An enumeration of these walks involving the number of visits to the diagonal appears in [**32, 43, 44**].
- **A non-D-finite case (steps NE, NW and SE)**: Mishna and Rechnitzer derived a complicated expression for the generating function of these walks, from which they were able to prove that this series has infinitely many singularities, and thus cannot be D-finite [**41**].

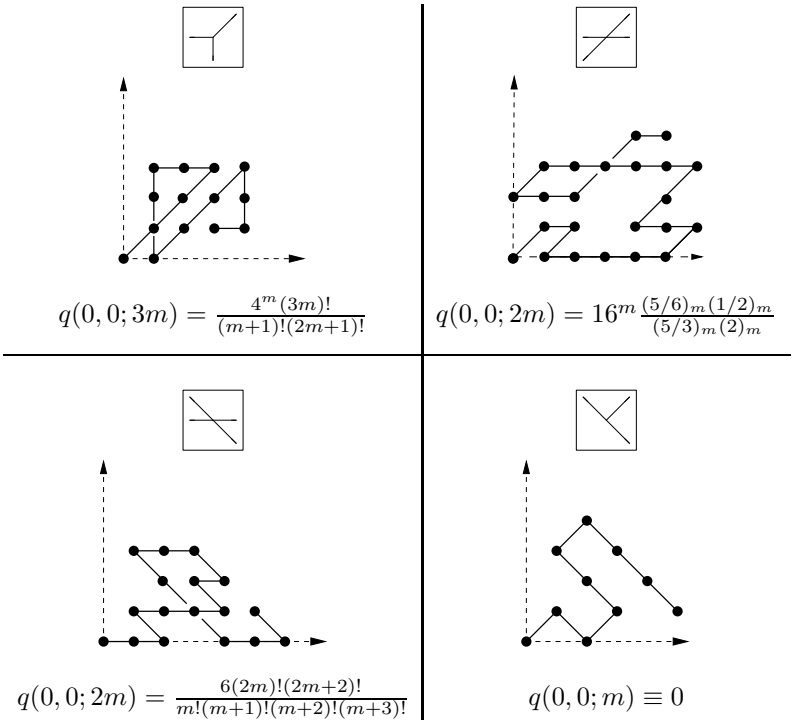


FIGURE 1. Four models of walks in the quarter plane: Kreweras' walks, Gessel's walks, Gouyou-Beauchamps's walks, and the non-D-finite example of Mishna and Rechnitzer. The numbers $q(0,0;n)$ count walks of length n confined to the quarter plane that start and end at the origin. We denote $(a)_n := a(a+1)\cdots(a+n-1)$.

Observe that we have defined $\mathcal{S}$, the set of steps, as a subset of $\mathbb{Z}^2$, but that we often use a more intuitive terminology, referring to $(1,1)$ as a NE step, for instance. We occasionally abuse the coordinate notation and directly write our steps using x's and y's, writing for example $x\bar{y}$ for a SE step.

Ideally, we seek generic results or combinatorial conditions which ensure D-finite (or even algebraic) generating functions. One criterion of this type states that if the set $\mathcal{S}$ is invariant under reflection around a vertical axis (we say briefly that it has a vertical symmetry) and consists of steps (i,j) such that $|i| \leqslant 1$, then the generating function of quarter plane walks with steps in $\mathcal{S}$ is D-finite [**11, 15**]. This is also true when the set of steps is left invariant by a Weyl group and the walks are confined to a corresponding Weyl chamber [**28**]. We are not aware of any other such criteria.

1.1. Results. We restrict our attention to the quarter plane and to *small* steps (*i.e.*, the step set $\mathcal{S}$ is a subset of $\{-1,0,1\}^2 \setminus \{(0,0)\}$). This includes the four examples in Figure 1. We first narrow down the 2^8 possible cases to 79 distinct non-trivial problems (Section 2). Their step sets $\mathcal{S}$ are listed in Tables 1 to 4 in Section 8. These problems fall into two categories, depending on whether a certain

transformation group associated with $\mathcal{S}$ is finite or infinite (Section 3). The 23 models associated with a finite group turn out to be those that satisfy at least one of the following conditions:

- the step set possesses a vertical symmetry,
- the vector sum of the vectors in the step set is 0.

In Section 4 we develop certain general tools that apply to all models; in particular, we explain how to write for each of them a functional equation that defines the generating function of the walks. Then, we describe a *uniform way* to solve this equation for all of the models associated with a finite group, except one (Gessel's walks). The solutions are all D-finite, and even algebraic in three cases (Sections 5 and 6). Note that Gessel's walks are also known to have an algebraic generating function [**9**]. We conjecture that the solutions to all models with an infinite group are non-D-finite. We conclude in Section 7 with some comments and questions. The tables of Section 8 list the 79 models, classified according to the order of the corresponding group, and provide references to both the existing literature and the relevant result of this paper.

1.2. Comments and detailed outline of the paper. The following technical and/or bibliographical comments may be of interest to readers who have worked on similar problems. One will also find here a more detailed description of the contents of the paper.

The starting point of our approach is a functional equation defining the generating function $Q(x, y; t)$ that counts quarter plane walks by the length (variable t) and the coordinates of the endpoint (variables x and y). This equation merely reflects a step by step construction of quarter plane walks. For instance, the equation obtained for Kreweras' walks (the first example in Figure 1) reads:

$$(1 - t(1/x + 1/y + xy))\, Q(x, y; t) = 1 - (t/x)Q(0, y; t) - (t/y)Q(x, 0; t).$$

Note that there is no obvious way to derive from the above identity an equation for, say, $Q(0, 0; t)$ or $Q(1, 1; t)$. Following Zeilberger's terminology [**56**], we say that the variables x and y are *catalytic*. One of our objectives is to provide some general principles that may be applied to any such linear equation with two catalytic variables. The case of linear equations with *one* catalytic variable is well-understood, and the solutions are always algebraic [**14**].

One key tool in our approach is a certain group $G(\mathcal{S})$ of birational transformations that leaves the *kernel* of the functional equation (that is, the coefficient of $Q(x, y; t)$) unchanged (Section 3). We have borrowed this group from *the little yellow book* by Fayolle, Iasnogorodski and Malyshev [**21**], in which the authors study the stationary distributions of Markov chains with small steps in the quarter plane. Ever since it was imported from probability theory to combinatorics, this group has proved useful (in several disguises, such as the *obstinate*, *algebraic*, or *iterated* kernel method) to solve various enumeration problems, dealing with walks [**11, 13, 15, 33, 40, 41**], and also with other objects, such as permutations [**12**] or set partitions [**17, 55**]—the common feature of all these problems being that they boil down to solving a linear equation with two catalytic variables. A striking observation, which applies to all solutions obtained so far, is that the solution is D-finite if and only if the group is finite. We find that exactly 23 out of our 79 quarter plane models give rise to a finite group.

We then focus on these 23 models. For each of them, we derive in Section 4 an identity between various specializations of $Q(x, y; t)$ which we call the *orbit sum* (or *half-orbit sum*, when there is an x/y symmetry in $G(\mathcal{S})$).

In Section 5, we show how to derive $Q(x, y; t)$ from the orbit sum in 19 out of the 20 models that have a finite group and no x/y symmetry. The number of n-step walks ending at (i, j) is obtained by extracting the coefficient of $x^i y^j t^n$ in a rational series which is easily obtained from the step set and the group. This implies that the generating function $Q(x, y; t)$ is D-finite. The form of the solution is reminiscent of a formula obtained by Gessel and Zeilberger for the enumeration of walks confined to a Weyl chamber, when the set of steps is invariant under the associated Weyl group [**28**]. Indeed, when the quarter plane problem happens to be a Weyl chamber problem, our method can be seen as an algebraic version of the reflection principle (which is the basis of [**28**]). However, its range of applications seems to be more general. We work out in detail three cases: walks with N, W and SE steps (equivalent to Young tableaux with at most three rows), walks with N, S, E, W, SE and NW steps (which do not seem to have been solved before, but behave very much like the former case), and finally walks with E, W, NW and SE steps (studied in [**29**]), for which we obtain new explicit results.

The results of Section 6 may be considered more surprising: For the three models that have a finite group and an x/y symmetry, we derive the series $Q(x, y; t)$ from the half-orbit sum and find, remarkably, that $Q(x, y; t)$ is always algebraic. We work out in detail all cases: walks with S, W and NE steps (Kreweras' walks), walks with E, N and SW steps (the reverse steps of Kreweras' steps) and walks with N, S, E, W, NE and SW steps, which, to our knowledge, have never been studied before. In particular, we find that the series $Q(1, 1; t)$ that counts walks of the last type, regardless of their endpoint, satisfies a simple quartic equation.

1.3. Preliminaries and notation. Let A be a commutative ring and x an indeterminate. We denote by $A[x]$ (resp. $A[[x]]$) the ring of polynomials (resp. formal power series) in x with coefficients in A. If A is a field, then $A(x)$ denotes the field of rational functions in x, and $A((x))$ the field of Laurent series in x. These notations are generalized to polynomials, fractions and series in several indeterminates. We let $\bar{x} = 1/x$, so that $A[x, \bar{x}]$ is the ring of Laurent polynomials in x with coefficients in A. The coefficient of x^n in a Laurent series $F(x)$ is denoted $[x^n]F(x)$. The *valuation* of a Laurent series $F(x)$ is the smallest d such that x^d occurs in $F(x)$ with a non-zero coefficient.

The main family of series that we use is that of power series in t with coefficients in $A[x, \bar{x}]$, that is, series of the form

$$F(x; t) = \sum_{n \geqslant 0, i \in \mathbb{Z}} f(i; n) x^i t^n,$$

where for all n, almost all coefficients $f(i; n)$ are zero. The *positive part* of $F(x; t)$ in x is the following series, which has coefficients in $x\mathbb{Q}[x]$:

$$[x^{>}]F(x; t) := \sum_{n \geqslant 0, i > 0} f(i; n) x^i t^n.$$

We define similarly the negative, non-negative and non-positive parts of $F(x; t)$ in x, which we denote respectively by $[x^{<}]F(x; t), [x^{\geqslant}]F(x; t)$ and $[x^{\leqslant}]F(x; t)$.

In our generating functions, the indeterminate t keeps track of the length of the walks. We record the coordinates of the endpoints with the variables x and y. In order to simplify the notation, we often omit the dependence of our series in t, writing for instance $Q(x,y)$ instead of $Q(x,y;t)$ for the generating function of quarter plane walks.

Recall that a power series $F(x_1,\dots,x_k) \in \mathbb{K}[[x_1,\dots,x_k]]$, where $\mathbb{K}$ is a field, is *algebraic* (over $\mathbb{K}(x_1,\dots,x_k)$) if it satisfies a non-trivial polynomial equation $P(x_1,\dots,x_k,F(x_1,\dots,x_k)) = 0$. It is *transcendental* if it is not algebraic. It is *D-finite* (or *holonomic*) if the vector space over $\mathbb{K}(x_1,\dots,x_k)$ spanned by all partial derivatives of $F(x_1,\dots,x_k)$ has finite dimension. This means that for all $i \leqslant k$, the series F satisfies a (non-trivial) linear differential equation in x_i with coefficients in $\mathbb{K}[x_1,\dots,x_k]$. We refer to [**38, 39**] for a study of these series. All algebraic series are D-finite. In Section 5 we use the following result.

PROPOSITION 1. *If $F(x,y;t)$ is a rational power series in t, with coefficients in $\mathbb{C}(x)[y,\bar{y}]$, then $[y^>]F(x,y;t)$ is algebraic over $\mathbb{C}(x,y,t)$. If the latter series has coefficients in $\mathbb{C}[x,\bar{x},y]$, its positive part in x, that is, the series $[x^>][y^>]F(x,y;t)$, is a 3-variable D-finite series (in x, y and t).*

The first statement is a simple adaptation of [**25**, Thm. 6.1]. The key tool is to expand $F(x,y;t)$ in partial fractions of y. The second statement relies on the fact that the diagonal of a D-finite series is D-finite [**38**]. One first observes that there is some k such that $[y^>]F(x,y;tx^k)$ has polynomial coefficients in x and y, and then applies [**38**, p. 377, Remark (4)].

Below we also use the fact that a series $F(t)$ with real coefficients such that $[t^n]F(t) \sim \kappa\mu^n n^{-k}$ with $k \in \{1,2,3,\dots\}$ cannot be algebraic [**22**].

2. The number of non-equivalent non-simple models

Since we restrict ourselves to walks with "small" steps (sometimes called walks with small variations), there are only a finite number of cases to study, namely 2^8, the number of sets $\mathcal{S}$ formed of small steps. However, some of these models are trivial (for instance $\mathcal{S} = \varnothing$, or $\mathcal{S} = \{\bar{x}\}$). More generally, it sometimes happens that one of the two constraints imposed by the quarter plane holds automatically, at least when the other constraint is satisfied. Such models are equivalent to problems of walks confined to a half-space: their generating functions are always algebraic and can be derived automatically using the *kernel method* [**14, 2**]. We show in Section 2.1 that of the $2^8 = 256$ models, only 138 are truly two-constraint problems and are thus worth considering in greater detail. Then, some of the remaining problems coincide up to an x/y symmetry and are thus equivalent. As shown in Section 2.2, one finally obtains 79 inherently different, truly two-constraint problems.

2.1. Easy algebraic cases. Let us say that a step (i,j) is x-positive if $i > 0$. We define similarly x-negative, y-positive and y-negative steps. There are a number of reasons that may make the enumeration of quarter plane walks with steps in $\mathcal{S}$ a simple problem:

(1) If $\mathcal{S}$ contains no x-positive step, we can ignore its x-negative steps, which will never be used in a quarter plane walk: we are thus back to counting walks with vertical steps on a (vertical) half-line. The solution of this problem is always algebraic, and even rational if $\mathcal{S} = \varnothing$ or $\mathcal{S} = \{y\}$ or $\mathcal{S} = \{\bar{y}\}$;

(2) Symmetrically, if $\mathcal{S}$ contains no y-positive step, the problem is simple with an algebraic solution;
(3) If $\mathcal{S}$ contains no x-negative step, all walks with steps in $\mathcal{S}$ that start from $(0,0)$ lie in the half-plane $i \geqslant 0$. Thus any walk lying weakly above the x-axis is automatically a quarter plane walk, and the problem boils down to counting walks confined to the upper half-plane: the corresponding generating function is always algebraic;
(4) Symmetrically, if $\mathcal{S}$ contains no y-negative step, the problem is simple with an algebraic solution.

We can thus restrict our attention to sets $\mathcal{S}$ containing x-positive, x-negative, y-positive and y-negative steps. An inclusion-exclusion argument shows that the number of such sets is 161. More precisely, the polynomial that counts them by cardinality is

$$P_1(z) = (1+z)^8 - 4(1+z)^5 + 2(1+z)^2 + 4(1+z)^3 - 4(1+z) + 1 \\ = 2\,z^2 + 20\,z^3 + 50\,z^4 + 52\,z^5 + 28\,z^6 + 8\,z^7 + z^8.$$

In the expression for $P_1(z)$, one of the 4 terms $(1+z)^5$ counts sets with no x-positive step, one term $(1+z)^2$ those with no x-positive nor x-negative step, one term $(1+z)^3$ those with no x-positive nor y-positive step, and so on. All the sets $\mathcal{S}$ we have discarded correspond to problems that either are trivial or can be solved automatically using the kernel method.

Among the remaining 161 sets $\mathcal{S}$, some do not contain any step with both coordinates non-negative: in this case the only quarter plane walk is the empty walk. These sets are subsets of $\{\bar{x}, \bar{y}, x\bar{y}, \bar{x}\bar{y}, \bar{x}y\}$. But, as we have assumed at this stage that $\mathcal{S}$ contains x-positive and y-positive steps, both $x\bar{y}$ and $\bar{x}y$ must belong to $\mathcal{S}$. Hence we exclude 2^3 of our 161 step sets, which leaves us with 153 sets, the generating polynomial of which is

$$P_2(z) = P_1(z) - z^2(1+z)^3 = z^2 + 17\,z^3 + 47\,z^4 + 51\,z^5 + 28\,z^6 + 8\,z^7 + z^8.$$

Another, slightly less obvious, source of simplicity of the model is when one of the quarter plane constraints implies the other. Assume that all walks with steps in $\mathcal{S}$ that end at a non-negative abscissa automatically end at a non-negative ordinate (we say, for short, that the x-condition forces the y-condition). This implies in particular that the steps $\bar{y}$ and $x\bar{y}$ do not belong to $\mathcal{S}$. As we have assumed that $\mathcal{S}$ contains a y-negative step, $\bar{x}\bar{y}$ must be in $\mathcal{S}$. But then x cannot belong to $\mathcal{S}$, otherwise some walks with a non-negative final abscissa would have a negative final ordinate, as with x followed by $\bar{x}\bar{y}$. We are left with sets $\mathcal{S} \subset \{\bar{x}, y, xy, \bar{x}y, \bar{x}\bar{y}\}$ containing $\bar{x}\bar{y}$, and also xy (because we need at least one x-positive step). Observe that these five steps are those lying above the first diagonal. Conversely, it is easy to realize that for any such set, the x-condition forces the y-condition. The generating polynomial of such *super-diagonal* sets is $z^2(1+z)^3$. Symmetrically, we need not consider *sub-diagonal* sets. An inclusion-exclusion argument reduces the generating polynomial of non-simple cases to

$$P_3(z) = P_2(z) - 2z^2(1+z)^3 + z^2 = 11\,z^3 + 41\,z^4 + 49\,z^5 + 28\,z^6 + 8\,z^7 + z^8,$$

that is to say, to 138 sets $\mathcal{S}$.

2.2. Symmetries. The eight symmetries of the square act on the step sets. However, only the x/y symmetry (reflection across the first diagonal) leaves the quarter plane fixed. Thus two step sets obtained from one another by applying this symmetry lead to equivalent counting problems. As we want to count non-equivalent problems, we need to determine how many among the 138 sets $\mathcal{S}$ that are left have the x/y symmetry. We repeat the arguments of the previous subsection, counting only symmetric models. We successively obtain

$$\begin{aligned} P_1^{\mathrm{sym}} &= (1+z)^2(1+z^2)^3 - 2(1+z)(1+z^2) + 1, \\ P_2^{\mathrm{sym}} &= P_1^{\mathrm{sym}} - z^2(1+z)(1+z^2), \\ P_3^{\mathrm{sym}} &= P_2^{\mathrm{sym}} - z^2 = 3\,z^3 + 5\,z^4 + 5\,z^5 + 4\,z^6 + 2\,z^7 + z^8. \end{aligned}$$

For instance, the term we subtract from P_1^{sym} to obtain P_2^{sym} counts symmetric subsets of $\{\bar{x}, \bar{y}, x\bar{y}, \bar{x}\bar{y}, \bar{x}y\}$ containing $x\bar{y}$ and $\bar{x}y$. The generating polynomial of (inherently different) models that are neither trivial, nor equivalent to a one-constraint problem, is thus

$$\frac{1}{2}\,(P_3 + P_3^{\mathrm{sym}}) = 7\,z^3 + 23\,z^4 + 27\,z^5 + 16\,z^6 + 5\,z^7 + z^8.$$

This gives a total of 79 models, shown in Tables 1 to 4.

3. The group of the walk

Let $\mathcal{S}$ be a set of small steps containing x-positive, x-negative, y-positive and y-negative steps. This includes the 79 sets we wish to study. Let $S(x,y)$ denote the generating polynomial of the steps of $\mathcal{S}$:

$$S(x,y) = \sum_{(i,j)\in\mathcal{S}} x^i y^j. \tag{1}$$

It is a Laurent polynomial in x and y. Recall that $\bar{x}$ stands for $1/x$, and $\bar{y}$ for $1/y$. Let us write

$$S(x,y) = A_{-1}(x)\bar{y} + A_0(x) + A_1(x)y = B_{-1}(y)\bar{x} + B_0(y) + B_1(y)x. \tag{2}$$

By assumption, A_1, B_1, A_{-1} and B_{-1} are non-zero. Clearly, $S(x,y)$ is left unchanged by the following rational transformations:

$$\Phi\colon (x,y) \mapsto \left(\bar{x}\frac{B_{-1}(y)}{B_1(y)}, y\right) \quad \text{and} \quad \Psi\colon (x,y) \mapsto \left(x, \bar{y}\frac{A_{-1}(x)}{A_1(x)}\right).$$

Note that both Φ and Ψ are involutions, and thus *bi* rational transformations. By composition, they generate a group that we denote $G(\mathcal{S})$, or G if there is no risk of confusion. This group is isomorphic to a dihedral group D_n of order $2n$, with $n \in \mathbb{N} \cup \{\infty\}$. For each $g \in G$, one has $S(g(x,y)) = S(x,y)$. The *sign* of g is 1 (resp. -1) if g is the product of an even (resp. odd) number of generators Φ and Ψ.

Examples

1. Assume $\mathcal{S}$ is left unchanged by a reflection across a vertical line. This is equivalent to saying that $S(x,y) = S(\bar{x},y)$, or that $B_1(y) = B_{-1}(y)$, or that $A_i(x) = A_i(\bar{x})$ for $i = -1, 0, 1$. Then the orbit of (x,y) under the action of G reads

$$(x,y) \overset{\Phi}{\longleftrightarrow} (\bar{x},y) \overset{\Psi}{\longleftrightarrow} (\bar{x}, C(x)\bar{y}) \overset{\Phi}{\longleftrightarrow} (x, C(x)\bar{y}) \overset{\Psi}{\longleftrightarrow} (x,y),$$

with $C(x) = \frac{A_{-1}(x)}{A_1(x)}$, so that G is finite of order 4.

Note that there may exist rational transformations on (x, y) that leave $S(x, y)$ unchanged but are *not* in G. For instance, if $\mathcal{S} = \{\mathsf{N}, \mathsf{S}, \mathsf{E}, \mathsf{W}\}$, the map $(x, y) \mapsto (y, x)$ leaves $S(x, y)$ unchanged, but the orbit of (x, y) under G is $\{(x, y), (\bar{x}, y), (\bar{x}, \bar{y}), (x, \bar{y})\}$.

2. Consider the case $\mathcal{S} = \{\bar{x}, y, x\bar{y}\}$. We have $A_{-1}(x) = x$, $A_1(x) = 1$, $B_{-1}(y) = 1$, $B_1(y) = \bar{y}$. The transformations are

$$\Phi\colon (x, y) \mapsto (\bar{x}y, y) \quad \text{and} \quad \Psi\colon (x, y) \mapsto (x, x\bar{y}),$$

and they generate a group of order 6:

$$(x, y) \xleftrightarrow{\Phi} (\bar{x}y, y) \xleftrightarrow{\Psi} (\bar{x}y, \bar{x}) \xleftrightarrow{\Phi} (\bar{y}, \bar{x}) \xleftrightarrow{\Psi} (\bar{y}, x\bar{y}) \xleftrightarrow{\Phi} (x, x\bar{y}) \xleftrightarrow{\Psi} (x, y). \tag{3}$$

3. Consider now the case $\mathcal{S} = \{\bar{x}, \bar{y}, xy\}$, which differs from the previous one by a rotation of 90 degrees. We have $A_{-1}(x) = 1$, $A_1(x) = x$, $B_{-1}(y) = 1$, $B_1(y) = y$. The two transformations are

$$\Phi\colon (x, y) \mapsto (\bar{x}\bar{y}, y) \quad \text{and} \quad \Psi\colon (x, y) \mapsto (x, \bar{x}\bar{y}),$$

and they also generate a group of order 6:

$$(x, y) \xleftrightarrow{\Phi} (\bar{x}\bar{y}, y) \xleftrightarrow{\Psi} (\bar{x}\bar{y}, x) \xleftrightarrow{\Phi} (y, x) \xleftrightarrow{\Psi} (y, \bar{x}\bar{y}) \xleftrightarrow{\Phi} (x, \bar{x}\bar{y}) \xleftrightarrow{\Psi} (x, y).$$

As shown by the following lemma, this is not a coincidence.

LEMMA 2. *Let $\mathcal{S}$ and $\tilde{\mathcal{S}}$ be two sets of steps differing by one of the 8 symmetries of the square. Then the groups $G(\mathcal{S})$ and $G(\tilde{\mathcal{S}})$ are isomorphic.*

PROOF. The group of symmetries of the square is generated by the two reflections Δ (across the first diagonal) and V (across a vertical line). Hence it suffices to prove the lemma when $\tilde{S} = \Delta(S)$ and when $\tilde{S} = V(S)$. We denote by Φ and Ψ the transformations associated with $\mathcal{S}$, and by $\tilde{\Phi}$ and $\tilde{\Psi}$ those associated with $\tilde{\mathcal{S}}$.

Assume $\tilde{S} = \Delta(S)$. We have $\tilde{A}_i(x) = B_i(x)$ and $\tilde{B}_i(y) = A_i(y)$. Denote by δ the involution that swaps the coordinates of a pair: $\delta(x, y) = (y, x)$. An elementary calculation gives

$$\tilde{\Phi} = \delta \circ \Psi \circ \delta \quad \text{and} \quad \tilde{\Psi} = \delta \circ \Phi \circ \delta$$

so that the groups $G(\mathcal{S})$ and $G(\tilde{\mathcal{S}})$ are conjugate by δ.

Assume now $\tilde{S} = V(S)$. We have $\tilde{A}_i(x) = A_i(\bar{x})$ and $\tilde{B}_i(y) = B_{-i}(y)$. Denote by v the involution that replaces the first coordinate of a pair by its reciprocal: $v(x, y) = (\bar{x}, y)$. An elementary calculation gives

$$\tilde{\Phi} = v \circ \Phi \circ v \quad \text{and} \quad \tilde{\Psi} = v \circ \Psi \circ v$$

so that the groups $G(\mathcal{S})$ and $G(\tilde{\mathcal{S}})$ are conjugate by v. □

THEOREM 3. *Out of the 79 models under consideration, exactly 23 are associated with a finite group:*

- *16 have a vertical symmetry and thus a group of order 4,*
- *5 have a group of order 6,*
- *2 have a group of order 8.*

PROOF. Given that Φ and Ψ are involutions, the group they generate is finite of order $2n$ if and only if $\Theta := \Psi \circ \Phi$ has finite order n. It is thus easy to prove that one of the groups $G(\mathcal{S})$ has order $2n$: one computes the m^{th} iterate Θ^m for $1 \leqslant m \leqslant n$, and checks that only the last of these transformations is the identity.

We have already seen in the examples above that models with a vertical symmetry have a group of order 4. We leave it to the reader to check that the models of Tables 2 and 3 have groups of order 6 and 8, respectively. These tables give the orbit of (x, y) under the action of G, the elements being listed in the following order: (x, y), $\Phi(x, y)$, $\Psi \circ \Phi(x, y)$, and so on.

Proving that one of the groups $G(\mathcal{S})$ is infinite is a more difficult task. We apply two different strategies, depending on $\mathcal{S}$. The first one uses valuations and works for the five step sets of Figure 2. These are the sets of our collection for which all elements (i, j) satisfy $i + j \geqslant 0$. We are very grateful to Jason Bell, who suggested to us a second strategy which turned out to apply to the remaining cases.

FIGURE 2. Five step sets with an infinite group.

1. The valuation argument

Let z be an indeterminate, and let x and y be Laurent series in z with coefficients in $\mathbb{Q}$, of respective valuations a and b. We assume that the *trailing* coefficients of these series, namely $[z^a]x$ and $[z^b]y$, are positive. Let us define x' by $\Phi(x, y) = (x', y)$. Then the trailing coefficient of x' (and y) is positive, and the valuation of x' (and y) only depends on a and b:

$$\phi(a, b) := (\mathrm{val}(x'), \mathrm{val}(y)) = \begin{cases} \left(-a + b(v_{-1}^{(y)} - v_1^{(y)}), b\right) & \text{if } b \geqslant 0, \\ \left(-a + b(d_{-1}^{(y)} - d_1^{(y)}), b\right) & \text{if } b \leqslant 0, \end{cases}$$

where $v_i^{(y)}$ (resp. $d_i^{(y)}$) denotes the valuation (resp. degree) in y of $B_i(y)$, for $i = \pm 1$. Similarly, $\Psi(x, y) := (x, y')$ is well-defined, and the valuations of x and y' only depend on a and b:

$$\psi(a, b) := (\mathrm{val}(x), \mathrm{val}(y')) = \begin{cases} \left(a, -b + a(v_{-1}^{(x)} - v_1^{(x)})\right) & \text{if } a \geqslant 0, \\ \left(a, -b + a(d_{-1}^{(x)} - d_1^{(x)})\right) & \text{if } a \leqslant 0, \end{cases}$$

where $v_i^{(x)}$ (resp. $d_i^{(x)}$) denotes the valuation (resp. degree) in x of $A_i(x)$, for $i = \pm 1$.

In order to prove that G is infinite, it suffices to prove that the group G' generated by ϕ and ψ is infinite. To prove the latter statement, it suffices to exhibit $(a, b) \in \mathbb{Z}^2$ such that the orbit of (a, b) under the action of G' is infinite.

Let $\mathcal{S}$ be one of the five sets of Figure 2. Then $A_{-1}(x) = x$ and $B_{-1}(y) = y$, so that $v_{-1}^{(x)} = d_{-1}^{(x)} = v_{-1}^{(y)} = d_{-1}^{(y)} = 1$. Also, $v_1^{(x)} = v_1^{(y)} = -1$ as $\mathcal{S}$ contains the steps $\bar{x}y$ and $x\bar{y}$. Hence the transformations ϕ and ψ read:

$$\phi(a, b) = \begin{cases} (-a + 2b, b) & \text{if } b \geqslant 0, \\ \left(-a + b(1 - d_1^{(y)}), b\right) & \text{if } b \leqslant 0, \end{cases}$$

$$\psi(a, b) = \begin{cases} (a, 2a - b) & \text{if } a \geqslant 0, \\ \left(a, -b + a(1 - d_1^{(x)})\right) & \text{if } a \leqslant 0. \end{cases}$$

It is easy to check, by induction on $n \geqslant 0$, that

$$(\psi \circ \phi)^n(1, 2) = (2n + 1, 2n + 2) \quad \text{and} \quad \phi(\psi \circ \phi)^n(1, 2) = (2n + 3, 2n + 2).$$

(All these pairs have positive entries, so that we never need to know $d_1^{(y)}$ or $d_1^{(x)}$.) Hence the orbit of $(1,2)$ under the action of ϕ and ψ is infinite, and so are the groups G' and G.

We believe, from our computer experiments, that the groups G' associated with the remaining 51 models of Table 4 are finite, and hence, cannot be used to prove that G is infinite. Instead, we use for these models a different argument based on the fixed points of $\Theta = \Psi \circ \Phi$.

2. The fixed point argument

We are left with 51 models. Thanks to Lemma 2, we only need to prove that (roughly) a quarter of them are associated with a finite group: if $G(\mathcal{S})$ is infinite, then $G(\tilde{\mathcal{S}})$ is infinite for all sets $\tilde{\mathcal{S}}$ that differ from $\mathcal{S}$ by a symmetry of the square[1]. This leaves 14 models to study, listed in Table 5.

Assume $\Theta = \Psi \circ \Phi$ is well-defined in the neighborhood of $(a,b) \in \mathbb{C}^2$, and that this point is fixed by Θ. Note that a and b are algebraic over $\mathbb{Q}$. Let us write $\Theta = (\Theta_1, \Theta_2)$, where Θ_1 and Θ_2 are the two coordinates of Θ. Each Θ_i sends the pair (x,y) to a rational function of x and y. The local expansion of Θ around (a,b) reads

$$\Theta(a+u, b+v) = (a,b) + (u,v)J_{a,b} + O(u^2) + O(v^2) + O(uv),$$

where $J_{a,b}$ is the Jacobian matrix of Θ at (a,b):

$$J_{a,b} = \begin{pmatrix} \dfrac{\partial \Theta_1}{\partial x}(a,b) & \dfrac{\partial \Theta_2}{\partial x}(a,b) \\ \dfrac{\partial \Theta_1}{\partial y}(a,b) & \dfrac{\partial \Theta_2}{\partial y}(a,b) \end{pmatrix}.$$

Iterating the above expansion gives, for $m \geqslant 1$,

$$\Theta^m(a+u, b+v) = (a,b) + (u,v)J_{a,b}^m + O(u^2) + O(v^2) + O(uv).$$

Assume $G(\mathcal{S})$ is finite of order $2n$, so that Θ is of order n. Then $\Theta^n(a+u, b+v) = (a,b) + (u,v)$, and the above equation shows that $J_{a,b}^n$ is the identity matrix. In particular, all eigenvalues of $J_{a,b}$ are roots of unity.

This gives us a strategy for proving that a group $G(\mathcal{S})$ is infinite: find a fixed point (a,b) for Θ, and compute the characteristic polynomial $\chi(X)$ of the Jacobian matrix $J_{a,b}$. This is a polynomial in X with coefficients in $\mathbb{Q}(a,b)$. In order to decide whether the roots of χ are roots of unity, we eliminate a and b (which are algebraic numbers) from the equation $\chi(X) = 0$ to obtain a polynomial $\bar{\chi}(X) \in \mathbb{Q}[X]$ that vanishes for all eigenvalues of $J_{a,b}$: if none of its factors is cyclotomic, we can conclude that $G(\mathcal{S})$ is infinite. As all cyclotomic polynomials of given degree are known, this procedure is effective.

Let us treat one case in detail, say $\mathcal{S} = \{x, y, \bar{y}, \bar{x}\bar{y}\}$ (the first case in Table 5). We have

$$\Theta(x,y) = \Psi \circ \Phi(x,y) = (\bar{x}\bar{y}, x + \bar{y}).$$

Every pair (a,b) such that $a^4 + a^3 = 1$ and $b = 1/a^2$ is fixed by Θ. Let us choose one such pair. The Jacobian matrix reads

$$J_{a,b} = \begin{pmatrix} -1 & 1 \\ -a^3 & -a^4 \end{pmatrix},$$

[1]Why a quarter, rather than an eighth? Recall that, if two (distinct) models differ by an x/y symmetry, only one of them appears in Table 4.

and its characteristic polynomial is

$$\chi(X) := \det(X\,\mathrm{Id} - J_{a,b}) = X^2 + X(1+a^4) + a^3 + a^4.$$

Let X be a root of this polynomial. By eliminating a (which satisfies $a^4 + a^3 = 1$), we obtain

$$\bar{\chi}(X) := X^8 + 9\,X^7 + 31\,X^6 + 62\,X^5 + 77\,X^4 + 62\,X^3 + 31\,X^2 + 9\,X + 1 = 0.$$

This polynomial is irreducible, and distinct from all cyclotomic polynomials of degree 8. Hence none of its roots are roots of unity, no power of $J_{a,b}$ is equal to the identity matrix, and the group $G(\mathcal{S})$ is infinite.

This strategy turns out to work for all models of Table 5. This table gives, for each model, the algebraic equations defining the fixed point (a, b) that we choose (for instance, the "condition" $a^4 + a^3 - 1$ occurring on the first line means that $a^4 + a^3 - 1 = 0$), and a polynomial $\bar{\chi}(X) \in \mathbb{Q}[X]$ that vanishes at all eigenvalues of the Jacobian matrix, in factored form. One then checks that no factor of this polynomial is cyclotomic.

It may be worth noting that this second strategy does *not* work for the five models of Figure 2: in the first three cases, Θ has no fixed point; in the last two cases, it has a fixed point (a, b), but the sixth power of the Jacobian matrix $J_{a,b}$ is the identity (of course, Θ^6 is *not* the identity; more precisely, the expansion of $\Theta^6(a + u, b + v)$ involves cubic terms in u and v). □

Remark. To put this discussion in a larger framework, let us mention that the group of birational transformations of $\mathbb{C}^2$ (or of the projective plane $\mathbb{P}^2(\mathbb{C})$), called the *Cremona group*, has been the object of many studies in algebraic geometry since the end of the 19th century [**34, 54**]. It seems possible that, given the attention already paid to the classification of finite subgroups of this group (see, e.g., [**3, 7**]), there is a generic or automatic test for finiteness that can be applied to all our examples.

4. General tools

Let $\mathcal{S}$ be one of the 79 step sets of Tables 1 to 4. Let $\mathcal{Q}$ be the set of walks that start from $(0, 0)$, take their steps from $\mathcal{S}$ and always remain in the first quadrant. Let $q(i, j; n)$ be the number of such walks that have length n and end at position (i, j). Denote by $Q(x, y; t) \equiv Q(x, y)$ the associated generating function:

$$Q(x, y; t) = \sum_{i,j,n \geqslant 0} q(i, j; n) x^i y^j t^n.$$

It is a formal power series in t with coefficients in $\mathbb{Q}[x, y]$.

4.1. A functional equation.

Lemma 4. *As a power series in t, the generating function $Q(x, y) \equiv Q(x, y; t)$ of walks with steps taken from $\mathcal{S}$ starting from $(0, 0)$ and staying in the first quadrant is characterized by the functional equation*

$$K(x, y)xyQ(x, y) = xy - txA_{-1}(x)Q(x, 0) - tyB_{-1}(y)Q(0, y) + t\epsilon Q(0, 0),$$

where

$$K(x, y) = 1 - tS(x, y) = 1 - t \sum_{(i,j)\in\mathcal{S}} x^i y^j$$

is called the kernel *of the equation, the polynomials* $A_{-1}(x)$ *and* $B_{-1}(y)$ *are the coefficients of* $\bar{y}$ *and* $\bar{x}$ *in* $S(x,y)$*, as described by* (2)*, and* ϵ *is* 1 *if* $(-1,-1)$ *is one of the allowed steps, and* 0 *otherwise.*

Proof. We construct walks step by step, starting from the empty walk and concatenating a new step at the end of the walk at each stage. The empty walk has weight 1. The generating function of walks obtained by adding a step of $\mathcal{S}$ at the end of a walk of $\mathcal{Q}$ is $tS(x,y)Q(x,y)$. However, some of these walks exit the quadrant: those obtained by concatenating a y-negative step to a walk ending at ordinate 0, and those obtained by concatenating an x-negative step to a walk ending at abscissa 0. Walks ending at ordinate (resp. abscissa) 0 are counted by the series $Q(x,0)$ (resp. $Q(0,y)$). Hence we must subtract the series $t\bar{y}A_{-1}(x)Q(x,0)$ and $t\bar{x}B_{-1}(y)Q(0,y)$. However, if $(-1,-1)\in\mathcal{S}$, we have subtracted twice the series counting walks obtained by concatenating this step to a walk ending at $(0,0)$: we must thus add the series $\epsilon t\bar{x}\bar{y}Q(0,0)$. This inclusion-exclusion argument gives

$$Q(x,y) = 1 + tS(x,y)Q(x,y) - t\bar{y}A_{-1}(x)Q(x,0) - t\bar{x}B_{-1}(y)Q(0,y) + \epsilon t\bar{x}\bar{y}Q(0,0),$$

which, multiplied by xy, gives the equation of the lemma.

The fact that it characterizes $Q(x,y;t)$ completely (as a power series in t) comes from the fact that the coefficient of t^n in $Q(x,y;t)$ can be computed inductively using this equation. This is of course closely related to the fact that we have used a recursive description of walks in $\mathcal{Q}$ to obtain the equation. □

4.2. Orbit sums. We have seen in Section 3 that all transformations g of the group G associated with the step set $\mathcal{S}$ leave the polynomial $S(x,y)$ unchanged. Hence they also leave the kernel $K(x,y) = 1 - tS(x,y)$ unchanged. Write the equation of Lemma 4 as

$$K(x,y)xyQ(x,y) = xy - F(x) - G(y) + t\epsilon Q(0,0),$$

with $F(x) = txA_{-1}(x)Q(x,0)$ and $G(y) = tyB_{-1}(y)Q(0,y)$. Replacing (x,y) by $\Phi(x,y) = (x',y)$ gives

$$K(x,y)x'yQ(x',y) = x'y - F(x') - G(y) + t\epsilon Q(0,0).$$

The difference between the former and latter identities reads:

$$K(x,y)\big(xyQ(x,y) - x'yQ(x',y)\big) = xy - x'y - F(x) + F(x').$$

The term $G(y)$ has disappeared. We can repeat this process, and add to this identity the functional equation of Lemma 4, evaluated at $(x',y') = \Psi(x',y)$. This gives:

$$K(x,y)\big(xyQ(x,y) - x'yQ(x',y) + x'y'Q(x',y')\big) \\ = xy - x'y + x'y' - F(x) - G(y') + t\epsilon Q(0,0).$$

Now the term $F(x')$ has disappeared. If G is finite of order $2n$, we can repeat the procedure until we come back to $(\Psi\circ\Phi)^n(x,y) = (x,y)$. That is to say, we form the alternating sum of the equations over the orbit of (x,y). All unknown functions on the right-hand side finally vanish, giving the following proposition, where we use the notation

$$\text{for } g\in G, \qquad g(A(x,y)) := A(g(x,y)).$$

PROPOSITION 5 (Orbit sums). *Assume the group $G(\mathcal{S})$ is finite. Then*

$$\sum_{g\in G} \mathrm{sgn}(g) g(xyQ(x,y;t)) = \frac{1}{K(x,y;t)} \sum_{g\in G} \mathrm{sgn}(g) g(xy). \tag{4}$$

Observe that the right-hand side is a rational function in x, y and t. We show in Section 5 that this identity implies immediately that 19 of the 23 models having a finite group have a D-finite solution.

The 4 remaining models are Gessel's model $\{x, \bar{x}, xy, \bar{x}\bar{y}\}$ (which we do not solve in this paper) and the three models with steps $\mathcal{S}_1 = \{\bar{x}, \bar{y}, xy\}$, $\mathcal{S}_2 = \{x, y, \bar{x}\bar{y}\}$, and $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$. For each of these three models, the orbit of (x, y) is

$$(x,y) \overset{\Phi}{\longleftrightarrow} (\bar{x}\bar{y},y) \overset{\Psi}{\longleftrightarrow} (\bar{x}\bar{y},x) \overset{\Phi}{\longleftrightarrow} (y,x) \overset{\Psi}{\longleftrightarrow} (y,\bar{x}\bar{y}) \overset{\Phi}{\longleftrightarrow} (x,\bar{x}\bar{y}) \overset{\Psi}{\longleftrightarrow} (x,y),$$

and exhibits an x/y symmetry. That is, (y, x) belongs to the orbit of (x, y). Moreover, if $g((x, y)) = (y, x)$, then $\mathrm{sgn}(g) = -1$. Thus the right-hand side of (4) vanishes, leaving

$$xyQ(x,y) - \bar{x}Q(\bar{x}\bar{y},y) + \bar{y}Q(\bar{x}\bar{y},x) = xyQ(y,x) - \bar{x}Q(y,\bar{x}\bar{y}) + \bar{y}Q(x,\bar{x}\bar{y}).$$

But this identity directly follows from the obvious relation $Q(x, y) = Q(y, x)$ and does not bring much information. In Section 6, we solve these three obstinate models by summing the functional equation over one half of the orbit only. Given that $A_{-1}(x) = B_{-1}(x)$, the identity resulting from this half-orbit summation reads as follows.

PROPOSITION 6 (Half-orbit sums). *Let $\mathcal{S}_1 = \{\bar{x}, \bar{y}, xy\}$ and $\mathcal{S}_2 = \{x, y, \bar{x}\bar{y}\}$. If the set of steps is $\mathcal{S}_1$, $\mathcal{S}_2$ or $\mathcal{S}_1 \cup \mathcal{S}_2$, then*

$$xyQ(x,y) - \bar{x}Q(\bar{x}\bar{y},y) + \bar{y}Q(\bar{x}\bar{y},x) = \frac{xy - \bar{x} + \bar{y} - 2txA_{-1}(x)Q(x,0) + t\epsilon Q(0,0)}{K(x,y)}.$$

REMARK. For Gessel's walks, the orbit of (x, y) is shown in Table 3. Proposition 5 reads

$$\sum_{g\in G} \mathrm{sgn}(g) g(xyQ(x,y)) = 0,$$

although no obvious symmetry explains this identity.

4.3. The roots of the kernel. Recall that the kernel of the main functional equation (Lemma 4) is

$$K(x,y) = 1 - t \sum_{(i,j)\in\mathcal{S}} x^i y^j.$$

LEMMA 7. *Let*

$$\Delta(x) = (1 - tA_0(x))^2 - 4t^2 A_{-1}(x) A_1(x).$$

Let $Y_0(x)$ and $Y_1(x)$ denote the two roots of the kernel $K(x, y)$, where $K(x, y)$ is viewed as a polynomial in y. These roots are Laurent series in t with coefficients in $\mathbb{Q}(x)$:

$$Y_0(x) = \frac{1 - tA_0(x) - \sqrt{\Delta(x)}}{2tA_1(x)}, \qquad Y_1(x) = \frac{1 - tA_0(x) + \sqrt{\Delta(x)}}{2tA_1(x)}. \tag{5}$$

Their valuations in t are respectively 1 and -1. Moreover, $1/K(x,y)$ is a power series in t with coefficients in $\mathbb{Q}[x,\bar{x},y,\bar{y}]$, and the coefficient of y^j in this series can be easily extracted using

$$\frac{1}{K(x,y)} = \frac{1}{\sqrt{\Delta(x)}}\left(\frac{1}{1-\bar{y}Y_0(x)} + \frac{1}{1-y/Y_1(x)} - 1\right). \tag{6}$$

PROOF. The equation $K(x,Y)=0$ also reads

$$Y = t\left(A_{-1}(x) + YA_0(x) + Y^2A_1(x)\right). \tag{7}$$

Solving this quadratic provides the expressions for $Y_0(x)$ and $Y_1(x)$ given above. As $\Delta(x) = 1 + O(t)$, the series Y_1 has valuation -1 in t, and first term $1/(tA_1(x))$. The equation

$$Y_0(x)Y_1(x) = \frac{A_{-1}(x)}{A_1(x)}$$

then shows that $Y_0(x)$ has valuation 1. This is also easily seen from (7), which in turn implies that Y_0 has coefficients in $\mathbb{Q}[x,\bar{x}]$. The equation

$$Y_0(x) + Y_1(x) = \frac{1}{tA_1(x)} - \frac{A_0(x)}{A_1(x)}$$

shows that for $n \geqslant 1$, the coefficient of t^n in $Y_1(x)$ is also a Laurent polynomial in x. This is not true, in general, of the coefficients of t^{-1} and t^0.

The identity (6) results from a partial fraction expansion in y. Note that both Y_0 and $1/Y_1$ have valuation 1 in t. Hence the expansion in y of $1/K(x,y)$ gives

$$[y^j]\frac{1}{K(x,y)} = \begin{cases} \dfrac{Y_0(x)^{-j}}{\sqrt{\Delta(x)}}, & \text{if } j \leqslant 0 \\[2ex] \dfrac{Y_1(x)^{-j}}{\sqrt{\Delta(x)}}, & \text{if } j \geqslant 0. \end{cases} \tag{8}$$

□

4.4. Canonical factorization of the discriminant $\Delta(x)$. The kernel can be seen as a polynomial in y. Its discriminant is then a (Laurent) polynomial in x:

$$\Delta(x) = (1 - tA_0(x))^2 - 4t^2A_{-1}(x)A_1(x).$$

Say Δ has valuation $-\delta$, and degree d in x. Then it admits $\delta + d$ roots $X_i \equiv X_i(t)$, for $1 \leqslant i \leqslant \delta + d$, which are Puiseux series in t with complex coefficients. Exactly δ of them, say $X_1, \dots, X_\delta$, are finite (and actually vanish) at $t = 0$. The remaining d roots, $X_{\delta+1}, \dots, X_{\delta+d}$, have a negative valuation in t and thus diverge at $t = 0$. (We refer to [**52**, Chapter 6] for generalities on solutions of algebraic equations with coefficients in $\mathbb{C}(t)$.) We write

$$\Delta(x) = \Delta_0\Delta_-(\bar{x})\Delta_+(x),$$

with

$$\Delta_-(\bar{x}) \equiv \Delta_-(\bar{x};t) = \prod_{i=1}^{\delta}(1 - \bar{x}X_i),$$

$$\Delta_+(x) \equiv \Delta_+(x;t) = \prod_{i=\delta+1}^{\delta+d}(1 - x/X_i),$$

and

$$\Delta_0 \equiv \Delta_0(t) = (-1)^\delta \frac{[\bar{x}^\delta]\Delta(x)}{\prod_{i=1}^{\delta} X_i} = (-1)^d [x^d]\Delta(x) \prod_{i=\delta+1}^{\delta+d} X_i.$$

It can be seen that Δ_0 (resp. $\Delta_-(\bar{x})$, $\Delta_+(x)$) is a formal power series in t with constant term 1 and coefficients in $\mathbb{C}$ (resp. $\mathbb{C}[\bar{x}]$, $\mathbb{C}[x]$). The above factorization is an instance of the canonical factorization of series in $\mathbb{Q}[x,\bar{x}][[t]]$, which was introduced by Gessel [**25**], and has proved useful in several walk problems since then [**10, 13, 16**]. It will play a crucial role in Section 6.

5. D-finite solutions via orbit sums

In this section, we first show that 19 of the 23 models having a finite group can be solved from the corresponding orbit sum. This includes the 16 models having a vertical symmetry, plus 3 others. Then, we work out the last 3 models in detail, obtaining closed form expressions for the number of walks ending at prescribed positions.

5.1. A general result.

PROPOSITION 8. *For the 23 models associated with a finite group, except for the four cases $\mathcal{S} = \{\bar{x}, \bar{y}, xy\}$, $\mathcal{S} = \{x, y, \bar{x}\bar{y}\}$, $\mathcal{S} = \{x, y, \bar{x}, \bar{y}, xy, \bar{x}\bar{y}\}$ and $\mathcal{S} = \{x, \bar{x}, xy, \bar{x}\bar{y}\}$, the following holds. The rational function*

$$R(x,y;t) = \frac{1}{K(x,y;t)} \sum_{g\in G} \mathrm{sgn}(g) g(xy)$$

is a power series in t with coefficients in $\mathbb{Q}(x)[y,\bar{y}]$ (Laurent polynomials in y, having coefficients in $\mathbb{Q}(x)$). Moreover, the positive part in y of $R(x,y;t)$, denoted $R^+(x,y;t)$, is a power series in t with coefficients in $\mathbb{Q}[x,\bar{x},y]$. Extracting the positive part in x of $R^+(x,y;t)$ gives $xyQ(x,y;t)$. In brief,

$$xyQ(x,y;t) = [x^>][y^>]R(x,y;t). \tag{9}$$

In particular, $Q(x,y;t)$ is D-finite. The number of n-step walks ending at (i,j) is

$$q(i,j;n) = [x^{i+1}y^{j+1}] \left(\sum_{g\in G} \mathrm{sgn}(g) g(xy) \right) S(x,y)^n, \tag{10}$$

where

$$S(x,y) = \sum_{(p,q)\in\mathcal{S}} x^p y^q.$$

PROOF. We begin with the 16 models associated with a group of order 4 (Table 1). We shall then address the three remaining cases, $\mathcal{S} = \{\bar{x}, y, x\bar{y}\}$, $\mathcal{S} = \{x, \bar{x}, x\bar{y}, \bar{x}y\}$ and $\mathcal{S} = \{x, \bar{x}, y, \bar{y}, x\bar{y}, \bar{x}y\}$.

All models with a group of order 4 exhibit a vertical symmetry. That is, $K(x,y) = K(\bar{x},y)$. As discussed in the examples of Section 3, the orbit of (x,y) reads

$$(x,y) \overset{\Phi}{\longleftrightarrow} (\bar{x},y) \overset{\Psi}{\longleftrightarrow} (\bar{x}, C(x)\bar{y}) \overset{\Phi}{\longleftrightarrow} (x, C(x)\bar{y}) \overset{\Psi}{\longleftrightarrow} (x,y),$$

with $C(x) = \frac{A_{-1}(x)}{A_1(x)}$. The orbit sum of Proposition 5 reads

$$xyQ(x,y) - \bar{x}yQ(\bar{x},y) + \bar{x}\bar{y}C(x)Q(\bar{x},C(x)\bar{y}) + x\bar{y}C(x)Q(x,C(x)\bar{y}) = R(x,y).$$

Clearly, both sides of this identity are series in t with coefficients in $\mathbb{Q}(x)[y, \bar{y}]$. Let us extract the positive part in y: only the first two terms of the left-hand side contribute, and we obtain

$$xyQ(x,y) - \bar{x}yQ(\bar{x},y) = R^+(x,y).$$

It is now clear from the left-hand side that $R^+(x,y)$ has coefficients in $\mathbb{Q}[x,\bar{x},y]$. Extracting the positive part in x gives the expression (9) for $xyQ(x,y)$, since the second term of the left-hand side does not contribute.

Let us now examine the cases $\mathcal{S} = \{\bar{x}, y, x\bar{y}\}$, $\mathcal{S} = \{x, \bar{x}, x\bar{y}, \bar{x}y\}$ and $\mathcal{S} = \{x, \bar{x}, y, \bar{y}, x\bar{y}, \bar{x}y\}$. For each of them, the orbit of (x,y) consists of pairs of the form $(x^a y^b, x^c y^d)$ for integers a, b, c and d (see Tables 2 and 3). This implies that $R(x,y)$ is a series in t with coefficients in $\mathbb{Q}[x,\bar{x},y,\bar{y}]$. When extracting the positive part in x and y from the orbit sum of Proposition 5, it is easily checked, in each of the three cases, that only the term $xyQ(x,y)$ remains in the left-hand side. The expression for $xyQ(x,y)$ follows.

Proposition 1 then implies that $Q(x,y;t)$ is D-finite. The expression for $q(i,j;n)$ follows from a mere coefficient extraction. □

5.2. Two models with algebraic specializations: $\{\bar{x}, y, x\bar{y}\}$ and $\{x, \bar{x}, y, \bar{y}, x\bar{y}, \bar{x}y\}$. Consider the case $\mathcal{S} = \{\bar{x}, y, x\bar{y}\} \equiv \{\mathsf{W}, \mathsf{N}, \mathsf{SE}\}$. A walk w with steps taken from $\mathcal{S}$ remains in the first quadrant if each of its prefixes contains more N steps that SE steps, and more SE steps than W steps. These walks are thus in bijection with Young tableaux of height at most 3 (Figure 3), or, via the Schensted correspondence [**49**], with involutions having no decreasing subsequence of length 4. The enumeration of Young tableaux is a well-understood topic. In particular, the number of tableaux of a given shape—and hence the number of n-step walks ending at a prescribed position—can be written in closed form using the hook-length formula [**49**]. It is also known that the total number of tableaux of size n and height at most 3 is the nth Motzkin number [**48**]. Here, we recover these two results (and refine the latter), using orbit and half-orbit sums. Then, we show that the case $\mathcal{S} = \{x, \bar{x}, y, \bar{y}, x\bar{y}, \bar{x}y\}$, which, to our knowledge, has never been solved, behaves very similarly. In particular, the total number of n-step walks confined to the quadrant is also related to Motzkin numbers (Proposition 10).

W	6	7	10	
SE	3	5	9	
N	1	2	4	8

N-N-SE-N-SE-W-W-N-SE-W

FIGURE 3. A Young tableau of height 3 and the corresponding quarter plane walk with steps in $\{\mathsf{W}, \mathsf{N}, \mathsf{SE}\}$.

PROPOSITION 9. *The generating function of walks with steps* W, N, SE *confined to the quarter plane is the non-negative part (in x and y) of a rational series in t having coefficients in $\mathbb{Q}[x,\bar{x},y,\bar{y}]$:*

$$Q(x,y;t) = [x^{\geqslant} y^{\geqslant}]\tilde{R}(x,y;t), \quad \textit{with} \quad \tilde{R}(x,y;t) = \frac{(1-\bar{x}\bar{y})\left(1-\bar{x}^2 y\right)\left(1-x\bar{y}^2\right)}{1-t(\bar{x}+y+x\bar{y})}.$$

In particular, $Q(x,y)$ is D-finite. The number of walks of length $n=3m+2i+j$ ending at (i,j) is

$$q(i,j;n)=\frac{(i+1)(j+1)(i+j+2)(3m+2i+j)!}{m!(m+i+1)!(m+i+j+2)!}.$$

In particular,

$$q(0,0;3m)=\frac{2(3m)!}{m!(m+1)!(m+2)!}\sim\sqrt{3}\,\frac{3^{3m}}{\pi m^4},$$

so that $Q(0,0;t)$—and hence $Q(x,y;t)$—is transcendental.

However, the specialization $Q(x,1/x;t)$ is algebraic of degree 2:

$$Q(x,1/x;t)=\frac{1-t\bar{x}-\sqrt{1-2\,\bar{x}t+t^2\bar{x}^2-4\,t^2x}}{2xt^2}.$$

In particular, the total number of n-step walks confined to the quadrant is the n^{th} Motzkin number:

$$Q(1,1;t)=\frac{1-t-\sqrt{(1+t)(1-3t)}}{2t^2}=\sum_{n\geqslant 0}t^n\sum_{k=0}^{\lfloor n/2\rfloor}\frac{1}{k+1}\binom{n}{2k}\binom{2k}{k}.$$

Proof. The orbit of (x,y) under the action of G is shown in (3). The first result of the proposition is a direct application of Proposition 8, with $\tilde{R}(x,y)=R(x,y)/(xy)$. It is then an easy task to extract the coefficient of $x^iy^jt^n$ in $\tilde{R}(x,y;t)$, using

$$[x^iy^j](\bar{x}+y+x\bar{y})^n=\frac{(3m+2i+j)!}{m!(m+i)!(m+i+j)!}$$

if $n=3m+2i+j$.

The algebraicity of $Q(x,\bar{x})$ can be proved as follows: let us form the alternating sum of the three equations obtained from Lemma 4 by replacing (x,y) by the first three elements of the orbit. These elements are those in which y occurs with a non-negative exponent. This gives

$$K(x,y)(xyQ(x,y)-\bar{x}y^2Q(\bar{x}y,y)+\bar{x}^2yQ(\bar{x}y,\bar{x}))\\ =xy-\bar{x}y^2+\bar{x}^2y-tx^2Q(x,0)-t\bar{x}Q(0,\bar{x}).$$

We now specialize this equation to two values of y. First, replace y by $\bar{x}$: the second and third occurrence of Q in the left-hand side cancel out, leaving

$$K(x,\bar{x})Q(x,\bar{x})=1-tx^2Q(x,0)-t\bar{x}Q(0,\bar{x}).$$

For the second specialization, replace y by the root $Y_0(x)$ of the kernel (see (5)). This is a well-defined substitution, as $Y_0(x)$ has valuation 1 in t. The left-hand side vanishes, leaving

$$0=xY_0(x)-\bar{x}Y_0(x)^2+\bar{x}^2Y_0(x)-tx^2Q(x,0)-t\bar{x}Q(0,\bar{x}).$$

By combining the last two equations, we obtain

$$K(x,\bar{x})Q(x,\bar{x})=1-xY_0(x)+\bar{x}Y_0(x)^2-\bar{x}^2Y_0(x).$$

The expression for $Q(x,\bar{x})$ follows. □

The case $\mathcal{S}=\{\mathsf{N},\mathsf{S},\mathsf{W},\mathsf{E},\mathsf{SE},\mathsf{NW}\}$ is very similar to the previous one. In particular, the orbit of (x,y) is the same in both cases. The proof of the previous proposition translates almost verbatim. Remarkably, Motzkin numbers occur again.

PROPOSITION 10. *The generating function of walks with steps* N, S, W, E, SE, NW *confined to the quarter plane is the non-negative part (in x and y) of a rational function:*

$$Q(x,y;t) = [x^{\geqslant} y^{\geqslant}]\tilde{R}(x,y;t), \quad \text{with} \quad \tilde{R}(x,y;t) = \frac{(1-\bar{x}\bar{y})\,(1-\bar{x}^2 y)\,(1-x\bar{y}^2)}{1-t(x+y+\bar{x}+\bar{y}+x\bar{y}+\bar{x}y)}.$$

In particular, $Q(x,y;t)$ is D-finite. The specialization $Q(x,1/x;t)$ is algebraic of degree 2:

$$Q(x,1/x;t) = \frac{1-tx-t\bar{x}+\sqrt{(1-t(x+\bar{x}))^2-4t^2(1+x)(1+\bar{x})}}{2t^2\,(1+x)\,(1+\bar{x})}.$$

In particular, the total number of n-step walks confined to the quadrant is 2^n times the n^{th} Motzkin number:

$$Q(1,1;t) = \frac{1-2t-\sqrt{(1+2t)(1-6t)}}{8t^2} = \sum_{n\geqslant 0} t^n \sum_{k=0}^{\lfloor n/2\rfloor} \frac{2^n}{k+1}\binom{n}{2k}\binom{2k}{k}.$$

5.3. The case $\mathcal{S} = \{x, \bar{x}, x\bar{y}, \bar{x}y\}$. Consider quadrant walks made of E, W, NW and SE steps. These walks are easily seen to be in bijection with pairs of non-intersecting prefixes of Dyck paths: To pass from such a pair to a quadrant walk, parse the pair of paths from left to right and assign a direction (E, W, NW or SE) to each pair of steps, as described in Figure 4. Hence the number of walks ending at a prescribed position is given by a 2-by-2 Gessel–Viennot determinant [**27**], and we can expect a closed form expression for this number.

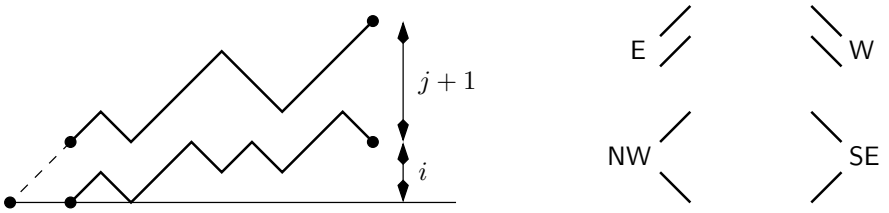


FIGURE 4. A pair of non-intersecting prefixes of Dyck paths corresponding to the quarter plane walk E-W-E-E-NW-SE-W-E-E-NW.

Also, note that the transformation $(i,j) \mapsto (i+j, j)$ maps these quarter plane walks bijectively to walks with E, W, N, S steps confined to $\{(i,j) : 0 \leqslant j \leqslant i\}$. In this form, they were studied by Gouyou-Beauchamps, who proved that the number of n-step walks ending on the x-axis is a product of Catalan numbers [**29**]. His interest in these walks came from a bijection he had established between walks ending on the x-axis and Young tableaux of height at most 4 [**30**]. This bijection is far from being as obvious as the one that relates tableaux of height at most 3 to quarter plane walks with W, N and SW steps (Section 5.2).

Here, we first specialize Proposition 8 to obtain the number of n-step walks ending at (i,j) (Proposition 11). Then, we perform an indefinite summation on i, or j, or both i and j to obtain additional closed form expressions, including Gouyou-Beauchamps's (Corollary 12).

PROPOSITION 11. *The generating function of walks with steps* E, W, NW, SE *confined to the quarter plane is the non-negative part (in x and y) of a rational function:*

$$Q(x,y;t) = [x^{\geqslant} y^{\geqslant}]\tilde{R}(x,y;t),$$

with

$$\tilde{R}(x,y;t) = \frac{(1-\bar{x})\,(1+\bar{x})\,(1-\bar{y})\,\left(1-\bar{x}^2 y\right)\,(1-x\bar{y})\,(1+x\bar{y})}{1-t(x+\bar{x}+x\bar{y}+\bar{x}y)}.$$

In particular, $Q(x,y;t)$ is D-finite. The number of walks of length $n = 2m+i$ ending at (i,j) is

$$q(i,j;2m+i) = \frac{(i+1)(j+1)(i+j+2)(i+2j+3)}{(2m+i+1)(2m+i+2)(2m+i+3)^2}\binom{2m+i+3}{m-j}\binom{2m+i+3}{m+1}.$$

In particular,

$$q(0,0;2m) = \frac{6(2m)!(2m+2)!}{m!(m+1)!(m+2)!(m+3)!} \sim \frac{24\cdot 4^{2m}}{\pi m^5},$$

so that $Q(0,0;t)$—and hence $Q(x,y;t)$—is transcendental.

Before we prove this proposition, let us perform summations on i and j. Recall that a hypergeometric sequence $(f(k))_k$ is *Gosper summable* (in k) if there exists another hypergeometric sequence $(g(k))_k$ such that $f(k) = g(k+1) - g(k)$. In this case, indefinite summation can be performed in closed form [**46**, Chapter 5]:

$$\sum_{k=k_0}^{k_1} f(k) = g(k_1+1) - g(k_0).$$

The numbers $q(i,j;n)$ of Proposition 11 have remarkable Gosper properties, from which we now derive Gouyou-Beauchamps's result for walks ending on the x-axis, and more.

COROLLARY 12. *The numbers $q(i,j;n)$ are Gosper-summable in i and in j. Hence sums of the form $\sum_{i=i_0}^{i_1} q(i,j;n)$ and $\sum_{j=j_0}^{j_1} q(i,j;n)$ have closed form expressions. In particular, the number of walks of length n ending at ordinate j is*

$$q(-,j;n) := \sum_{i\geqslant 0} q(i,j;n) = \begin{cases} \dfrac{(j+1)(2m)!(2m+2)!}{(m-j)!(m+1)!^2(m+j+2)!} & \text{if } n=2m, \\[2ex] \dfrac{2(j+1)(2m+1)!(2m+2)!}{(m-j)!(m+1)!(m+2)!(m+j+2)!} & \text{if } n=2m+1. \end{cases}$$

Similarly, the number of walks of length $n = 2m+i$ ending at abscissa i is

$$q(i,-;2m+i) := \sum_{j\geqslant 0} q(i,j;2m+i) = \frac{(i+1)(2m+i)!(2m+i+2)!}{m!(m+1)!(m+i+1)!(m+i+2)!}.$$

In particular, as many $2m$-step walks end on the x- and y-axes:

$$q(-,0;2m) = q(0,-;2m) = \frac{(2m)!(2m+2)!}{m!(m+1)!^2(m+2)!}.$$

The numbers $q(i,-;n)$ and $q(-,j;n)$ defined above are again Gosper-summable in i and j respectively. Hence sums of the form

$$\sum_{i=i_0}^{i_1} q(i,-;n) = \sum_{i=i_0}^{i_1}\sum_{j} q(i,j;n) \quad \text{and} \quad \sum_{j=j_0}^{j_1} q(-,j;n) = \sum_{j=j_0}^{j_1}\sum_{i} q(i,j;n),$$

which count walks ending between certain vertical or horizontal lines, have simple closed-form expressions. In particular, the total number of quarter plane walks of length n is

$$q(-,-;n) := \sum_{i,j\geqslant 0} q(i,j;n) = \begin{cases} \dfrac{(2m)!(2m+1)!}{m!^2(m+1)!^2} & \text{if } n=2m, \\ \dfrac{(2m+1)!(2m+2)!}{m!(m+1)!^2(m+2)!} & \text{if } n=2m+1. \end{cases}$$

The asymptotic behaviors of these numbers are found to be

$$q(-,-;n) \sim c_1 \cdot 4^n/n^2, \quad q(0,-;n) \sim c_2 \cdot 4^n/n^3, \quad q(-,0;n) \sim c_3 \cdot 4^n/n^3,$$

which shows that the series $Q(1,1;t)$, $Q(0,1;t)$ and $Q(1,0;t)$ are transcendental.

Proof of Proposition 11 and Corollary 12. The orbit of (x,y) under the action of G, of cardinality 8, is shown in Table 3. The expression for $Q(x,y;t)$ given in Proposition 11 is a direct application of Proposition 8, with $\tilde{R}(x,y) = R(x,y)/(xy)$. It is then an easy task to extract the coefficient of $x^i y^j t^n$ in $\tilde{R}(x,y;t)$, using

$$x+\bar{x}+\bar{x}y+x\bar{y} = (1+\bar{y})(x+\bar{x}y) \quad \text{and} \quad [x^i y^j](x+\bar{x}+\bar{x}y+x\bar{y})^n = \binom{n}{m+i}\binom{n}{m-j}$$

for $n = 2m+i$. This proves Proposition 11.

For the first part of the corollary, that is, the expressions for $q(-,j;n)$ and $q(i,-;n)$, it suffices to check the following identities, which we have obtained using the implementation of Gosper's algorithm found in the `sumtools` package of Maple:

$$\begin{aligned} q(2i,j;2m) &= g_1(i,j;m) - g_1(i+1,j;m), \\ q(2i+1,j;2m+1) &= g_2(i,j;m) - g_2(i+1,j;m), \\ q(i,j;2m+i) &= g_3(i,j;m) - g_3(i,j+1;m), \end{aligned}$$

with

$$\begin{aligned} g_1(i,j;m) &= \frac{2\,(1+j)\,(m+1+2\,i\,(i+1+j))\,(2\,m)!\;(2\,m+1)!}{(m-i-j)!\;(m-i+1)!\;(m+i+1)!\;(m+i+j+2)!}, \\ g_2(i,j;m) &= \frac{2\,(1+j)\,(m+1+j\,(1+2\,i)+2\;(1+i)^2)\,(2\,m+1)!\;(2\,m+2)!}{(m-i-j)!\;(m-i+1)!\,(m+i+2)!\;(m+i+j+3)!}, \\ g_3(i,j;m) &= \frac{(1+i)\,(m+1+(i+j+1)\,(1+j))\,(2\,m+i)!\;(2\,m+i+2)!}{(m-j)!\;(m+1)!\;(m+i+2)!\;(m+i+j+2)!}. \end{aligned}$$

These three identities respectively lead to

$$\begin{aligned} q(-,j;2m) &= g_1(0,j;m), \\ q(-,j;2m+1) &= g_2(0,j;m), \\ q(i,-;2m+i) &= g_3(i,0;m), \end{aligned}$$

as stated in the corollary.

For the second part, we have used the following identities, also discovered (and proved) using MAPLE:

$$\begin{aligned} q(2i,-;2m) &= g_4(i;m) - g_4(i+1;m),\\ q(2i+1,-;2m+1) &= g_5(i;m) - g_5(i+1;m),\\ q(-,j;2m) &= g_6(j;m) - g_6(j+1;m),\\ q(-,j;2m+1) &= g_7(j;m) - g_7(j+1;m), \end{aligned}$$

with

$$\begin{aligned} g_4(i;m) &= \frac{(2\,m)!\;(2\,m+1)!}{(m-i)!\;(m-i+1)!\;(m+i)!\;(m+i+1)!},\\ g_5(i;m) &= \frac{(2\,m+1)!\;(2\,m+2)!}{(m-i)!\;(m-i+1)!\;(m+i+1)!\;(m+i+2)!},\\ g_6(j;m) &= \frac{(2\,m)!\;(2\,m+1)!}{(m-j)!\,m!\;(m+1)!\;(m+j+1)!},\\ g_7(j;m) &= \frac{(2\,m+1)!\;(2\,m+2)!}{(m-j)!\;(m+1)!\;(m+2)!\;(m+j+1)!}. \end{aligned}$$

Note that these identities give two ways to determine the total number of n-step walks in the quadrant, as

$$q(-,-;2m) = g_4(0;m) = g_6(0;m) \quad \text{and} \quad q(-,-;2m+1) = g_5(0;m) = g_7(0;m).$$

□

6. Algebraic solutions via half-orbit sums

In this section we solve in a unified manner the three models whose orbit has an x/y symmetry: $\mathcal{S}_1 = \{\bar{x}, \bar{y}, xy\}$, $\mathcal{S}_2 = \{x, y, \bar{x}\bar{y}\}$ and $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$. Remarkably, in all three cases the generating function $Q(x, y; t)$ is found to be algebraic. Our approach uses the *algebraic kernel method* introduced by the first author to solve the case $\mathcal{S} = \mathcal{S}_1$, that is, Kreweras' model [**13**, Section 2.3]. We refer to the introduction for more references on this model. The case $\mathcal{S} = \mathcal{S}_2$ was solved by the second author in [**40**], and the case $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$ is, to our knowledge, new.

Recall from Proposition 6 that for each of these three models,

$$xyQ(x,y) - \bar{x}Q(\bar{x}\bar{y},y) + \bar{y}Q(\bar{x}\bar{y},x) = \frac{xy - \bar{x} + \bar{y} - 2txA_{-1}(x)Q(x,0) + t\epsilon Q(0,0)}{K(x,y)}.$$

Extract from this equation the coefficient of y^0: in the left-hand side, only the second term contributes, and its contribution is $\bar{x}Q_d(\bar{x})$, where $Q_d(x) \equiv Q_d(x;t)$ is the generating function of walks ending on the diagonal:

$$Q_d(x;t) = \sum_{n,i \geqslant 0} t^n x^i q(i,i;n).$$

The coefficient of y^0 in the right-hand side can be easily extracted using (8). This gives

$$-\bar{x}Q_d(\bar{x}) = \frac{1}{\sqrt{\Delta(x)}}\left(xY_0(x) - \bar{x} + \frac{1}{Y_1(x)} - 2txA_{-1}(x)Q(x,0) + t\epsilon Q(0,0)\right),$$

or, given the expression (5) of Y_0 and the fact that $Y_0Y_1 = \bar{x}$,

$$\frac{x}{tA_1(x)} - \bar{x}Q_d(\bar{x}) = \frac{1}{\sqrt{\Delta(x)}}\left(\frac{x(1-tA_0(x))}{tA_1(x)} - \bar{x} - 2txA_{-1}(x)Q(x,0) + t\epsilon Q(0,0)\right).$$

Let us write the canonical factorization of $\Delta(x) = \Delta_0\Delta_+(x)\Delta_-(\bar{x})$ (see Section 4.4). Multiplying the equation by $A_1(x)\sqrt{\Delta_-(\bar{x})}$ gives

$$\sqrt{\Delta_-(\bar{x})}\left(\frac{x}{t} - \bar{x}A_1(x)Q_d(\bar{x})\right) = \frac{1}{\sqrt{\Delta_0\Delta_+(x)}}\left(\frac{x(1-tA_0(x))}{t} - \bar{x}A_1(x) - 2txA_{-1}(x)A_1(x)Q(x,0) + t\epsilon A_1(x)Q(0,0)\right). \tag{11}$$

Each term in this equation is a Laurent series in t with coefficients in $\mathbb{Q}[x,\bar{x}]$. Moreover, very few positive powers of x occur in the left-hand side, while very few negative powers in x occur in the right-hand side. We shall extract from this equation the positive and negative parts in x, and this will give algebraic expressions for the unknown series $Q_d(x)$ and $Q(x,0)$. From now on, we consider each model separately.

6.1. The case $\mathcal{S} = \{\bar{x}, \bar{y}, xy\}$. We have $A_{-1}(x) = 1$, $A_0(x) = \bar{x}$, $A_1(x) = x$ and $\epsilon = 0$. The discriminant $\Delta(x)$ reads $(1-t\bar{x})^2 - 4t^2x$. The curve $\Delta(x;t) = 0$ has a rational parametrization in terms of the series $W \equiv W(t)$, defined as the only power series in t satisfying

$$W = t(2 + W^3). \tag{12}$$

Replacing t by $W/(2+W^3)$ in $\Delta(x)$ gives the canonical factorization as $\Delta(x) = \Delta_0\Delta_+(x)\Delta_-(\bar{x})$ with

$$\Delta_0 = \frac{4t^2}{W^2}, \qquad \Delta_+(x) = 1 - xW^2, \qquad \Delta_-(\bar{x}) = 1 - \frac{W(W^3+4)}{4x} + \frac{W^2}{4x^2}.$$

Extracting the positive part in x from (11) gives

$$\frac{x}{t} = -\frac{\left(2\,t^2x^2Q\left(x,0\right) - x + 2\,t\right)W}{2t^2\sqrt{1-xW^2}} + \frac{W}{t},$$

from which we obtain an expression for $Q(x,0)$ in terms of W. Extracting the non-positive part in x from (11) gives

$$\sqrt{1 - \frac{W\left(W^3+4\right)}{4x} + \frac{W^2}{4x^2}}\ \left(\frac{x}{t} - Q_d\left(\bar{x}\right)\right) - \frac{x}{t} = -\frac{W}{t},$$

from which we obtain an expression for $Q_d(\bar{x})$. We recover the results of [**13**].

Proposition 13. *Let $W \equiv W(t)$ be the power series in t defined by* (12). *Then the generating function of quarter plane walks formed of* W, S *and* NE *steps, and ending on the x-axis, is*

$$Q(x,0;t) = \frac{1}{tx}\left(\frac{1}{2t} - \frac{1}{x} - \left(\frac{1}{W} - \frac{1}{x}\right)\sqrt{1-xW^2}\right).$$

Consequently, the length generating function of walks ending at $(i,0)$ is

$$[x^i]Q(x,0;t) = \frac{W^{2i+1}}{2\cdot 4^i\ t}\left(C_i - \frac{C_{i+1}W^3}{4}\right),$$

where $C_i = \binom{2i}{i}/(i+1)$ is the i-th Catalan number. The Lagrange inversion formula gives the number of such walks of length $m = 3m + 2i$ as

$$q(i,0;3m+2i) = \frac{4^m(2i+1)}{(m+i+1)(2m+2i+1)}\binom{2i}{i}\binom{3m+2i}{m}.$$

The generating function of walks ending on the diagonal is

$$Q_d(x;t) = \frac{W-\bar{x}}{t\sqrt{1-xW(1+W^3/4)+x^2W^2/4}} + \bar{x}/t.$$

Note that $Q(0,0)$ is algebraic of degree 3, $Q(x,0)$ is algebraic of degree 6, and $Q(x,y)$ (which can be expressed in terms of $Q(x,0)$, and $Q(0,y) = Q(y,0)$ using the functional equation we started from) is algebraic of degree 12.

6.2. The case $\mathcal{S} = \{x, y, \bar{x}\bar{y}\}$. The steps of this model are obtained by reversing the steps of the former model. In particular, the series $Q(0,0)$ counting walks that start and end at the origin is the same in both models. This observation was used in [**40**] to solve the latter case. We present here a self-contained solution.

We have $A_{-1}(x) = \bar{x}$, $A_0(x) = x$, $A_1(x) = 1$ and $\epsilon = 1$. The discriminant $\Delta(x)$ is now $(1-tx)^2 - 4t^2\bar{x}$, and is obtained by replacing x by $\bar{x}$ in the discriminant of the previous model. In particular, the canonical factors of $\Delta(x)$ are

$$\Delta_0 = \frac{4t^2}{W^2}, \qquad \Delta_+(x) = 1 - \frac{W(W^3+4)}{4}\,x + \frac{W^2}{4}\,x^2, \qquad \Delta_-(\bar{x}) = 1 - \bar{x}W^2,$$

where $W \equiv W(t)$ is the power series in t defined by (12). Extracting the coefficient of x^0 in (11) gives

$$-\frac{W^2}{2t} = -\frac{W\left(W^4 + 4\,W + 8t\,Q(0,0)\right)}{16t},$$

from which we obtain an expression for $Q(0,0)$. Extracting the non-negative part in x from (11) gives

$$\frac{x}{t} - \frac{W^2}{2t} = -\frac{\left(2xt^2Q(x,0) - xt^2Q(0,0) + t - x^2 + x^3t\right)W}{2xt^2\sqrt{1-xW(W^3+4)/4+x^2W^2/4}} + \frac{W}{2xt},$$

from which we obtain an expression for $Q(x,0)$. Finally, extracting the negative part in x from (11) gives

$$\left(\frac{x}{t} - \frac{Q_d\left(\bar{x}\right)}{x}\right)\sqrt{1-\frac{W^2}{x}} - \frac{x}{t} + \frac{W^2}{2t} = -\frac{W}{2xt},$$

from which we obtain an expression for $Q_d(\bar{x})$. We have thus recovered, and completed, the results of [**40**]. Note in particular how simple the number of walks of length n ending at a diagonal point (i,i) is.

Proposition 14. *Let $W \equiv W(t)$ be the power series in t defined by* (12). *Then the generating function of quarter plane walks formed of* N, E *and* SW *steps and ending on the x-axis is*

$$Q(x,0;t) = \frac{W\left(4-W^3\right)}{16t} - \frac{t-x^2+tx^3}{2xt^2} - \frac{\left(2\,x^2 - xW^2 - W\right)\sqrt{1-xW(W^3+4)/4\,+x^2W^2/4}}{2txW}.$$

The generating function of walks ending on the diagonal is

$$Q_d(x;t) = \frac{xW(x+W)-2}{2tx^2\sqrt{1-xW^2}} + \frac{1}{tx^2}.$$

Consequently, the length generating function of walks ending at (i,i) is

$$[x^i]Q_d(x;t) = \frac{W^{2i+1}}{4^{i+1}t\,(i+2)}\binom{2i}{i}\left(2i+4-(2i+1)W^3\right).$$

The Lagrange inversion formula gives the number of such walks of length $n = 3m+2i$ as

$$q(i,i;3m+2i) = \frac{4^m(i+1)^2}{(m+i+1)(2m+2i+1)}\binom{2i+1}{i}\binom{3m+2i}{m}.$$

Note that $Q(0,0)$ is algebraic of degree 3, $Q(x,0)$ is algebraic of degree 6, and $Q(x,y)$ (which can be expressed in terms of $Q(x,0)$, $Q(0,y)=Q(y,0)$ and $Q(0,0)$ using the functional equation we started from) is algebraic of degree 12.

6.3. The case $\mathcal{S} = \{x, y, \bar{x}, \bar{y}, xy, \bar{x}\bar{y}\}$. We have $A_{-1}(x) = 1+\bar{x}$, $A_0(x) = x+\bar{x}$, $A_1(x) = 1+x$ and $\epsilon = 1$. The discriminant $\Delta(x)$ is now

$$(1-t(x+\bar{x}))^2 - 4t^2(1+x)(1+\bar{x}),$$

and is symmetric in x and $\bar{x}$. Two of its roots, say X_1 and X_2, have valuation 1 in t, and the other two roots are $1/X_1$ and $1/X_2$. By studying the two elementary symmetric functions of X_1 and X_2 (which are the coefficients of $\Delta_-(\bar{x})$), we are led to introduce the power series $Z \equiv Z(t)$, satisfying

$$Z = t\,\frac{1-2\,Z+6\,Z^2-2\,Z^3+Z^4}{(1-Z)^2} \tag{13}$$

and having no constant term. Replacing t by its expression in terms of Z in $\Delta(x)$ provides the canonical factors of $\Delta(x)$ as

$$\Delta_0 = \frac{t^2}{Z^2}, \qquad \Delta_+(x) = 1 - 2\,Z\frac{1+Z^2}{(1-Z)^2}\,x + Z^2x^2, \qquad \Delta_-(\bar{x}) = \Delta_+(\bar{x}).$$

As in the previous case, extracting from (11) the coefficient of x^0 gives an expression for $Q(0,0)$:

$$Q(0,0) = \frac{Z(1-2\,Z-Z^2)}{t(1-Z)^2}.$$

Extracting then the positive and negative parts of (11) in x gives expressions for $Q(x,0)$ and $Q_d(\bar{x})$.

Proposition 15. *Let $Z \equiv Z(t)$ be the power series with no constant term satisfying* (13)*, and let*

$$\Delta_+(x) = 1 - 2\,Z\frac{1+Z^2}{(1-Z)^2}\,x + Z^2x^2.$$

Then the generating function of quarter plane walks formed of N, S, E, W, SE *and* NW *steps, and ending on the x-axis is*

$$Q(x,0;t) = \frac{\left(Z(1-Z)+2xZ-(1-Z)\,x^2\right)\sqrt{\Delta_+(x)}}{2txZ(1-Z)(1+x)^2} \\ - \frac{Z(1-Z)^2+Z\left(Z^3+4Z^2-5Z+2\right)x-\left(1-2Z+7Z^2-4Z^3\right)x^2+x^3Z(1-Z)^2}{2txZ(1-Z)^2(1+x)^2}.$$

The generating function of walks ending on the diagonal is

$$Q_d(x;t) = \frac{1-Z-2xZ+x^2Z(Z-1)}{tx(1+x)(Z-1)\sqrt{\Delta_+(x)}} + \frac{1}{tx(1+x)}.$$

Note that $Q(0,0)$ is algebraic of degree 4, $Q(x,0)$ is algebraic of degree 8, and $Q(x,y)$ (which can be expressed in terms of $Q(x,0)$, $Q(0,y)=Q(y,0)$ and $Q(0,0)$ using the functional equation we started from) is algebraic of degree 16. However, $Q \equiv Q(1,1)$ has degree 4 only, and the algebraic equation it satisfies has a remarkable form:

$$Q\,(1+tQ)\left(1+2t\,Q+2\,t^2Q^2\right) = \frac{1}{1-6t}.$$

Also, Motzkin numbers seem to be lurking around, as in Proposition 10.

Corollary 16. *Let $N \equiv N(t)$ be the only power series in t satisfying*

$$N = t(1+2N+4N^2).$$

Up to a factor t, this series is the generating function of the numbers $2^n M_n$, where M_n is the n^{th} Motzkin number:

$$N = \sum_{n\geqslant 0} t^{n+1} \sum_{k=0}^{\lfloor n/2\rfloor} \frac{2^n}{k+1}\binom{n}{2k}\binom{2k}{k}.$$

Then the generating function of all walks in the quadrant with steps N, S, E, W, SE *and* NW *is*

$$Q(1,1;t) = \frac{1}{2t}\left(\sqrt{\frac{1+2N}{1-2N}}-1\right),$$

and the generating function of walks in the quadrant ending at the origin is

$$Q(0,0;t) = \frac{(1+4N)^{3/2}}{2Nt} - \frac{1}{2t^2} - \frac{2}{t}.$$

The proof is elementary once the algebraic equations satisfied by $Q(1,1)$ and $Q(0,0)$ are obtained.

7. Final comments and questions

The above results raise, in our opinion, numerous natural questions. Here are some of them. The first two families of questions are of a purely combinatorial, or even bijective, nature ("explain why some results are so simple"). Others are more closely related to the method used in this paper. We also raise a question of an algorithmic nature.

7.1. Explain closed form expressions. We have obtained remarkable hypergeometric expressions for the number of walks in many cases (Propositions 9 to 14). Are there direct combinatorial explanations? Let us emphasize a few examples that we consider worth investigating.

- **Kreweras' walks and their reverse:** The number of Kreweras' walks ($\mathcal{S} = \{\bar{x}, \bar{y}, xy\}$) ending at $(i, 0)$ is remarkably simple (Proposition 13). A combinatorial explanation has been found when $i = 0$, in connection with the enumeration of planar triangulations [**4**]. To our knowledge, the generic case remains open. If we consider instead the reverse collection of steps ($\mathcal{S} = \{x, y, \bar{x}\bar{y}\}$), then it is the number of walks ending at (i, i) that is remarkably simple (Proposition 14). This is a new result, which we should like to see explained in a more combinatorial manner.
- **Motzkin numbers:** this famous sequence of numbers arises in the solution of the cases $\mathcal{S} = \{\bar{x}, y, x\bar{y}\}$ and $\mathcal{S} = \{x, \bar{x}, y, \bar{y}, x\bar{y}, \bar{x}y\}$ (Propositions 9 and 10). The first problem is equivalent to the enumeration of involutions with no decreasing subsequence of length 4, and the occurrence of Motzkin numbers follows from restricting a bijection of Françon and Viennot [**24**]. The solution to the second problem is, to our knowledge, new, and deserves a more combinatorial solution. Can one find a direct explanation for why the respective counting sequences for the total number of walks of these two models differ by a power of 2? Is there a connection with the 2^n phenomenon of [**19**]?
- **Gessel's walks:** although we have not solved this case ($\mathcal{S} = \{x, \bar{x}, xy, \bar{x}\bar{y}\}$) in this paper, we cannot resist advertising Gessel's former conjecture, which has now become Kauers–Koutschan–Zeilberger's theorem:

$$q(0, 0; 2m) = 16^m \frac{(5/6)_m (1/2)_m}{(5/3)_m (2)_m}.$$

Certain other closed form expressions obtained in this paper are less mysterious. As discussed at the beginning of Section 5.3, walks with E, W, NW and SE steps are in bijection with pairs of non-intersecting walks. The Gessel–Viennot method (which, as our approach, is an inclusion-exclusion argument) expresses the number of walks ending at (i, j) as a 2-by-2 determinant, thus justifying the closed form expressions of Proposition 11. The extension of this theory by Stembridge [**53**] allows one to let i, or j, or both i and j vary, and the number of walks is now expressed as a pfaffian. Hence the closed forms of Corollary 12 are not unexpected. However, one may try to find direct proofs not involving the inclusion-exclusion principle. Moreover, the following question may be interesting *per se*:

- **Walks with E, W, NW and SE steps:** can one explain bijectively why as many $2m$-step walks end on the x- and y-axes? Recall that those ending on the x-axis were counted bijectively in [**29**].

Another well-understood case is that of quarter plane walks with N, E, S and W steps. The number of such walks ending at the origin is a product of Catalan numbers, and this has been explained bijectively, first in a recursive manner [**18**], and more recently directly, using again certain planar maps as intermediate objects [**5**]. Another argument, based on the reflection principle and thus involving minus signs, appears in [**31**] and applies to more general endpoints.

7.2. Explain algebraic series. A related problem is to explain combinatorially, via a direct construction, why the three models of Section 6 have algebraic generating functions. Given the connection between Kreweras' walks and planar triangulations [**4**], this could be of the same complexity as proving directly that families of planar maps have an algebraic generating function. (Much progress has been made recently on this problem by Schaeffer, Di Francesco and their coauthors.) And what about Gessel's walks (with steps E, W, NE and SW), which we have not solved in this paper, but have very recently been proved to have an algebraic generating function as well [**9**]?

7.3. Models with a vertical symmetry. When $\mathcal{S}$ is invariant under a reflection across a vertical axis, the group $G(\mathcal{S})$ is of order 4 and Proposition 8 gives the generating function $Q(x, y; t)$ as the positive part of a rational function. We have not worked out the coefficient extraction in any of these 16 cases. This may be worth doing, with the hope of finding closed form expressions in some cases. However, according to [**8**], there is little hope of finding an algebraic solution for $Q(1, 1; t)$.

7.4. Models with an infinite group. Two of the 56 models that are associated with an infinite group (Table 4) have been proved to have a non-D-finite generating function [**41**]. We conjecture that this holds for all models with an infinite group. This conjecture is based on our experimental attempts to discover a differential equation satisfied by the generating function, and much strengthened by the further attempts of Bostan and Kauers [**8**], which are based on the calculation of 1000 terms of each generating function. It also relies on the fact that all equations with two catalytic variables and an infinite group that have been solved so far have a non-D-finite solution. How could one prove this conjecture, for instance in the case $\mathcal{S} = \{\mathsf{N}, \mathsf{E}, \mathsf{NE}, \mathsf{SW}\}$? And what is the importance, if any, that can be attributed to the fact that the step sets with finite groups either exhibit a symmetry across the y-axis or have a vector sum of zero?

7.5. Automatic derivation of differential equations. Our D-finite but transcendental solutions are expressed as the positive part (in x and y) of a 3-variable rational series in t, x and y (Proposition 8). Can one derive automatically from these expressions differential equations satisfied by $Q(0, 0; t)$ and $Q(1, 1; t)$? This would be a convenient way to fill in the gap between our work and the paper [**8**], where differential equations are conjectured for the series $Q(1, 1; t)$.

For the algebraic solutions of Section 6, it is easy to derive from our results first an algebraic equation satisfied by $Q(0, 0; t)$ (or $Q(1, 1; t)$), and then a differential equation satisfied by this series, using the Maple package Gfun [**50**].

7.6. Variations and extensions. It is natural to ask to which similar problems the approach used in this paper could be adapted. To make this question more precise, let us emphasize that such problems may involve, for instance, putting weights on the walks, allowing more general steps, or considering higher dimensions. However, the very first question is whether Gessel's model can be dealt with using the ideas of this paper!

Weighted paths. One may try to adapt our approach to solve refined enumeration problems. For instance, some authors have studied the enumeration of walks in a wedge, keeping track not only of the number of steps, but also of the number of

contacts, or *visits* to the boundary lines [**32, 43, 44**]. Of course, other statistics could be considered.

Another natural way to add weights, of a more probabilistic nature, consists in studying Markov chains confined to the quarter plane. The weight of a walk is then its probability. An entire book is devoted to the determination of the *stationary distributions* of such chains [**21**]. These distributions are governed by functional equations similar to ours, but *without the length variable* t (since only the stationary regime is considered). This difference makes the problem rather different in nature, and indeed, the tools involved in [**21**] are much more analytic than algebraic. A natural way to set the problem back in the algebraic playground (to which our power series methods belong) is to keep track of the length of the trajectories, which boils down to studying the law of the chain at time n. This was done in [**13**] for a probabilistic version of Kreweras' walks, using a variant of the method presented in this paper. An asymptotic analysis of the solution should then yield the limiting/stationary distribution. This was however *not* done in [**13**]. Instead, we enriched our algebraic approach with a few basic analytic arguments to solve directly the equation that describes the stationary distribution. This solution is in our opinion more elementary than the original ones [**21, 23**]. It is worth noting that the bivariate series that describes the stationary distribution is algebraic, and that this Kreweras chain is actually the main algebraic example of [**21**]. In view of the results presented in this paper, it is natural to ask if one could design probabilistic versions of the other three algebraic models (Propositions 14 and 15, plus Gessel's model) that would also yield algebraic stationary distributions.

Conversely, it is natural to ask whether certain tools from [**21**], other than the group of the walk, could be adapted to our power series context. We are thinking in particular of the material of Chapter 4, which is devoted to the case where the group of the walk is finite, and (under this hypothesis) to the conditions under which the stationary distribution has an algebraic generating function.

More general steps. The fact that we only allow "small" steps plays a crucial role in our approach, and more precisely in the definition of the group of the walk (Section 3). However, this does not mean that models with larger steps are definitely beyond reach. First, it is always possible to write a functional equation for the series $Q(x, y; t)$, based on a step-by-step construction of the walk. If no step has a coordinate smaller than -1, the right-hand side of the equation only involves $Q(0, y; t)$ and $Q(x, 0; t)$, but otherwise more unknown functions, depending exclusively on x or y, appear. Another important difference with the present setting is that the kernel now has degree larger than 2 (in x or y). One can still define a group, but acting on pairs (x, y) that cancel the kernel. We refer to [**15**] for the solution of a simple example, with steps $(2, -1)$ and $(-1, 2)$.

Higher dimensions. Finally, a natural question is to address 3-dimensional problems, such as those studied experimentally in [**8**]. Provided one focuses on walks with small steps, the key ingredients of our approach—the functional equation and the group of the walk—can indeed be adapted in a straightforward manner to this higher-dimensional context.

8. Tables

Tables 1–4 below list the 79 step sets $\mathcal{S}$ we consider, classified according to the order of the group $G(\mathcal{S})$.

Tables 1–3 contain the 23 step sets $\mathcal{S}$ for which $G(\mathcal{S})$ is finite: dihedral of order 4, 6, and 8, respectively. The orbit of (x, y) under the action of this group is listed in the second column. The third column displays the steps of $\mathcal{S}$. The fourth column shows the sequences $q(-, -; n)$ and $q(0, 0; n)$, indexed by n, which respectively count all quarter plane walks and quarter plane walks ending at the origin. We give the '`A` number' of each sequence, according to the current version of the OEIS (On-Line Encyclopedia of Integer Sequences) [**51**]. In the rightmost column we give references, both to this paper and to other papers.

TABLE 1. The 16 step sets $\mathcal{S}$ that are associated with a group isomorphic to D_2. In each case the 3-variable generating function $Q(x, y; t)$ is D-finite.

#	$G(\mathcal{S})$	$\mathcal{S}$	Sequences $q(-, -; n)$, $q(0, 0; n)$	References
1	(x, y), $(\bar{x}, y)$, $(\bar{x}, \bar{y})$, $(x, \bar{y})$		1,2,6,18,60,200,700,2450,8820,31752,... [A005566] 1,0,2,0,10,0,70,0,588,0,5544,0,56628,... [A005568]	Prop. 8; [**5, 11, 18, 28, 31**]
2			1,1,4,9,36,100,400,1225,4900,15876,... [A018224] 1,0,1,0,4,0,25,0,196,0,1764,0,17424,... [A001246]	Prop. 8; [**11, 28, 47**]
3			1,2,10,39,210,960,5340,26250,... [A151312] 1,0,2,0,18,0,255,0,4522,0,91896,0,... [A151362]	Prop. 8; [**11, 28**]
4			1,3,18,105,684,4550,31340,219555,... [A151331] 1,0,3,6,38,160,905,4830,28308,... [A172361]	Prop. 8; [**11, 28**]
5	(x, y), $(\bar{x}, y)$, $(\bar{x}, \bar{y}\frac{1}{x+\bar{x}})$, $(x, \bar{y}\frac{1}{x+\bar{x}})$		1,1,3,7,19,49,139,379,1079,3011,... [A151266] 1,0,0,0,2,0,0,0,28,0,0,0,660,0,0,0,... [A151332]	Prop. 8; [**11, 40**]
6			1,2,9,34,151,659,2999,13714,63799,... [A151307] 1,0,1,3,4,20,65,175,742,2604,9072,... [A151357]	Prop. 8; [**11**]

#	$G(\mathcal{S})$	$\mathcal{S}$	Sequences $q(-,-;n)$, $q(0,0;n)$	References
7	(x,y), $(\bar{x},y)$, $(\bar{x},\bar{y}\,\frac{1}{x+1+\bar{x}})$, $(x,\bar{y}\,\frac{1}{x+1+\bar{x}})$		1,2,7,23,84,301,1127,4186,15891,... [A151291] 1,0,1,0,4,0,20,0,126,0,882,0,6732,... [A151341]	Prop. 8; [**11**]
8			1,3,15,74,392,2116,11652,64967,... [A151326] 1,0,2,3,12,40,145,560,2240,9156,... [A151368]	Prop. 8; [**11**]
9	(x,y), $(\bar{x},y)$, $(\bar{x},\bar{y}\,\frac{x+\bar{x}}{x+1+\bar{x}})$, $(x,\bar{y}\,\frac{x+\bar{x}}{x+1+\bar{x}})$		1,2,8,29,129,535,2467,10844,50982,... [A151302] 1,0,1,0,6,0,55,0,644,0,8694,0,128964,... [A151345]	Prop. 8; [**11**]
10			1,3,16,86,509,3065,19088,120401,... [A151329] 1,0,2,3,20,60,345,1400,7770,36876,... [A151370]	Prop. 8; [**11**]
11	(x,y), $(\bar{x},y)$, $(\bar{x},\bar{y}\,(x+1+\bar{x}))$, $(x,\bar{y}\,(x+1+\bar{x}))$		1,1,3,5,17,34,121,265,969,2246,8351,... [A151261] 1,0,1,0,4,0,20,0,126,0,882,0,6732,... [A151341]	Prop. 8; [**11**]
12			1,2,7,26,105,444,1944,8728,39999,... [A151297] 1,0,2,3,12,40,145,560,2240,9156,... [A151368]	Prop. 8; [**11**]
13	(x,y), $(\bar{x},y)$, $(\bar{x},\bar{y}\,\frac{x+1+\bar{x}}{x+\bar{x}})$, $(x,\bar{y}\,\frac{x+1+\bar{x}}{x+\bar{x}})$		1,1,5,13,61,199,939,3389,16129,... [A151275] 1,0,1,0,6,0,55,0,644,0,8694,0,128964,... [A151345]	Prop. 8; [**11**]
14			1,2,11,49,277,1479,8679,49974,... [A151314] 1,0,2,3,20,60,345,1400,7770,36876,... [A151370]	Prop. 8; [**11**]
15	(x,y), $(\bar{x},y)$, $(\bar{x},\bar{y}(x+\bar{x}))$, $(x,\bar{y}(x+\bar{x}))$		1,1,2,3,8,15,39,77,216,459,1265,2739,... [A151255] 1,0,0,0,2,0,0,0,28,0,0,0,660,... [A151332]	Prop. 8; [**11, 40**]
16			1,2,6,21,76,290,1148,4627,19038,... [A151287] 1,0,1,3,4,20,65,175,742,2604,9072,... [A151357]	Prop. 8; [**11**]

TABLE 2. The 5 step sets $\mathcal{S}$ that are associated with a group $G(\mathcal{S})$ that is isomorphic to D_3. In each case the 3-variable generating function $Q(x, y; t)$ is D-finite; and in cases 3, 4, and 5, it is algebraic. In cases 1 and 2 there are algebraic specializations; e.g., $Q(1, 1; t)$, the generating function of the sequence $q(-, -; n)$, is algebraic.

#	$G(\mathcal{S})$	$\mathcal{S}$	Sequences $q(-,-;n)$, $q(0,0;n)$	References
1	(x,y), $(\bar{x}y,y)$, $(\bar{x}y,\bar{x})$, $(\bar{y},\bar{x})$, $(\bar{y},\bar{y}x)$, $(x,\bar{y}x)$		1,1,2,4,9,21,51,127,323,835,2188,... [A001006] 1,0,0,1,0,0,5,0,0,42,0,0,462,0,0,... [A005789]	Prop. 9; [**28**]
2			1,2,8,32,144,672,3264,16256,82688,... [A129400] 1,0,2,2,12,30,130,462,1946,7980,... [A151366]	Prop. 10; [**28**]
3	(x,y), $(\bar{x}\bar{y},y)$, $(\bar{x}\bar{y},x)$, (y,x), $(y,\bar{x}\bar{y})$, $(x,\bar{x}\bar{y})$		1,1,3,7,17,47,125,333,939,2597,7183,... [A151265] 1,0,0,2,0,0,16,0,0,192,0,0,2816,... [A006335]	Prop. 13; [**13, 37**]
4			1,2,4,10,26,66,178,488,1320,3674,... [A151278] 1,0,0,2,0,0,16,0,0,192,0,0,2816,... [A006335]	Prop. 14; [**40**]
5			1,3,14,67,342,1790,9580,52035,... [A151323] 1,0,3,4,26,80,387,1596,7518,34656,... [A151372]	Prop. 15; —

TABLE 3. The 2 step sets $\mathcal{S}$ that are associated with a group $G(\mathcal{S})$ that is isomorphic to D_4. In each case the 3-variable generating function $Q(x, y; t)$ is D-finite; and in case 2, it is algebraic.

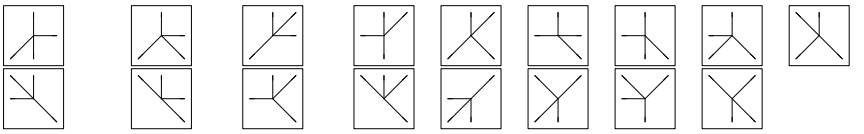
#	$G(\mathcal{S})$	$\mathcal{S}$	Sequences $q(-,-;n)$, $q(0,0;n)$	References
1	(x,y), $(y\bar{x},y)$, $(y\bar{x},y\bar{x}^2)$, $(\bar{x},y\bar{x}^2)$, $(\bar{x},\bar{y})$, $(x\bar{y},\bar{y})$, $(x\bar{y},x^2\bar{y})$, $(x,\bar{y}x^2)$		1,1,3,6,20,50,175,490,1764,5292,... [A005558] 1,0,1,0,3,0,14,0,84,0,594,0,4719,... [A005700]	Prop. 11; [**28, 29, 44**]
2	(x,y), $(\bar{x}\bar{y},y)$, $(\bar{x}\bar{y},x^2y)$, $(\bar{x},x^2y)$, $(\bar{x},\bar{y})$, $(xy,\bar{y})$, $(xy,\bar{x}^2\bar{y})$, $(x,\bar{y}\bar{x}^2)$		1,2,7,21,78,260,988,3458,13300,... [A060900] 1,0,2,0,11,0,85,0,782,0,8004,0,88044,... [A135404, A151360]	—; [**9, 35, 45**]

TABLE 4. The 56 step sets $\mathcal{S}$ that are associated with an infinite group.

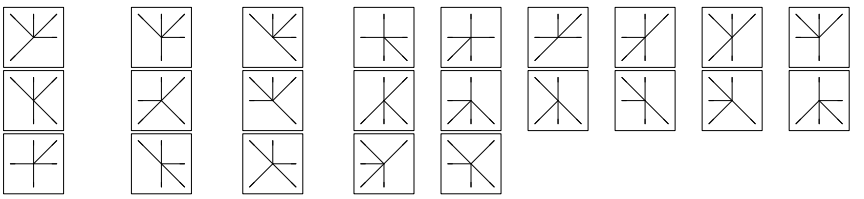
Three steps:

 (see [**41**])

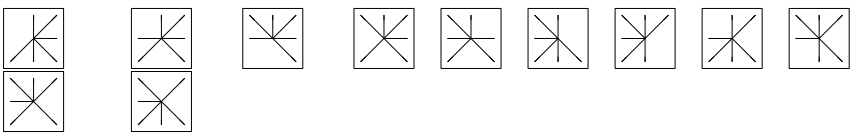
Four steps:



Five steps:



Six steps:



Seven steps:

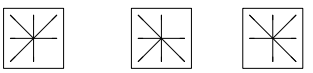


TABLE 5. Computations proving that for many of the step sets $\mathcal{S}$ of Table 4, the associated group $G(\mathcal{S})$ is infinite. (See Section 3.)

	Fixed point	Condition	$\bar{\chi}(X)$
Four steps:			
	$\left(a, a^{-2}\right)$	a^4+a^3-1	$X^8+9X^7+31X^6+62X^5+77X^4+62X^3+31X^2+9X+1$
	$\left(a, \frac{1-a^2}{a^2}\right)$	$a^5-a^4+a^3+2a^2-1$	$X^{10}+18\,X^9+125\,X^8+439\,X^7+897\,X^6+1131\,X^5+897\,X^4+439\,X^3+125\,X^2+18\,X+1$
	(a, a)	a^4+a^3-1	$X^8-19X^7-X^6-124X^5+3X^4-124X^3-X^2-19X+1$
	$\left(b^2/2, b\right)$	b^6+b^4-4	$X^6+9\,X^5+59\,X^4+70\,X^3+59\,X^2+9\,X+1$

	Fixed point	Condition	$\bar{\chi}(X)$
Five steps:			
	$(-1,1)$		$2\,X^2+3X+2$
	$\left(a,\frac{a^2}{1-a^2}\right)$	$a^5+a^3+2a^2-1$	$X^{10}+16\,X^9+106\,X^8+371\,X^7+764\,X^6+967\,X^5+764\,X^4$ $+371\,X^3+106\,X^2+16\,X+1$
	(a,a)	a^3-a-1	$X^6+8\,X^5+28\,X^4+41\,X^3+28\,X^2+8\,X+1$
	$\left(a,a^{-2}\right)$	a^3-a-1	$X^6+2\,X^5+6\,X^4+5\,X^3+6\,X^2+2\,X+1$
	$\left(\frac{2b^2}{1-b^2},b\right)$	$3b^6+b^4+b^2-1$	$X^6+14\,X^5+87\,X^4+100\,X^3+87\,X^2+14\,X+1$
	$(a,1/a)$	a^2+a+1	$X^4+17\,X^3+81\,X^2+17\,X+1$

	Fixed point	Condition	$\bar\chi(X)$
Six steps:			
	$\big(a, 2+a-3a^2-3a^3\big)$	$3a^6+6a^5+2a^4$ $-5a^3-4a^2+1$	$X^{12}+23\,X^{11}+283\,X^{10}+1861\,X^9+7461\,X^8$ $+14225\,X^7+18249\,X^6+14225\,X^5+7461\,X^4$ $+1861\,X^3+283\,X^2+23\,X+1$
	$\Big(\frac{2+2b-b^3}{b^2(1+b+b^2)}, b\Big)$	$b^7+b^6+2b^5+5b^4$ $+4b^3-4b^2-8b-4$	$X^{14}+37\,X^{13}+567\,X^{12}+4853\,X^{11}+26197\,X^{10}+89695\,X^9$ $+194611\,X^8+250446\,X^7+194611\,X^6+89695\,X^5$ $+26197\,X^4+4853\,X^3+567\,X^2+37\,X+1$
	(a,a)	a^4+a^3-1	$X^8+117\,X^7+3671\,X^6+13396\,X^5+19683\,X^4$ $+13396\,X^3+3671\,X^2+117\,X+1$
Seven steps:			
	$(a,1/a)$	a^3-a-1	$X^6+28\,X^5+224\,X^4+345\,X^3+224\,X^2+28\,X+1$

Acknowledgements. We are grateful to Michael Albert, Jason Bell, Pierrette Cassou-Nogues, Julie Déserti, Arnaud Jehanne and Andrew Rechnitzer for various discussions and advice related to this paper.

References

[1] D. André. *Solution directe du problème résolu par M. Bertrand*, C. R. Acad. Sci. Paris **105** (1887), 436–437.

[2] C. Banderier and Ph. Flajolet, *Basic analytic combinatorics of directed lattice paths*, Theoret. Comput. Sci. **281** (2002), no. 1–2, 37–80.

[3] L. Bayle and A. Beauville, *Birational involutions of* $\mathbf{P}^2$, Asian J. Math. **4** (2000), no. 1, 11–17.

[4] O. Bernardi, *Bijective counting of Kreweras walks and loopless triangulations*, J. Combin. Theory Ser. A **114** (2007), no. 5, 931–956.

[5] O. Bernardi, *Bijective counting of tree-rooted maps and shuffles of parenthesis systems*, Electron. J. Combin. **14** (2007), no. 1, research paper 9, 36 pp.

[6] J. Bertrand, *Solution d'un problème*, C. R. Acad. Sci. Paris **105** (1887), 369.

[7] J. Blanc, *Finite abelian subgroups of the Cremona group of the plane*, C. R. Math. Acad. Sci. Paris (2007), no. 1, 21–26.

[8] A. Bostan and M. Kauers. *Automatic classification of restricted lattice walks*, pp. 203–217 in *Formal Power Series and Algebraic Combinatorics (Proceedings of FPSAC '09)*, 2009.

[9] A. Bostan and M. Kauers, *The complete generating function for Gessel walks is algebraic*, preprint, 2009. Available as arxiv:0909.1965.

[10] M. Bousquet-Mélou, *Walks on the slit plane: Other approaches*, Adv. in Appl. Math. **27** (2001), no. 2–3, 243–288.

[11] M. Bousquet-Mélou, *Counting walks in the quarter plane*, pp. 49–67 in *Mathematics and Computer Science, II (Versailles,* 2002*)* (B. Chauvin, Ph. Flajolet, D. Gardy and A. Mokkadem, eds.), Birkhäuser, Boston/Basel, 2002.

[12] M. Bousquet-Mélou, *Four classes of pattern-avoiding permutations under one roof: Generating trees with two labels*, Electron. J. Combin. **9** (2002/3), no. 2, research paper 19, 31 pp.

[13] M. Bousquet-Mélou, *Walks in the quarter plane: Kreweras' algebraic model*, Ann. Appl. Probab. **15** (2005), no. 2, 1451–1491.

[14] M. Bousquet-Mélou and M. Petkovšek, *Linear recurrences with constant coefficients: The multivariate case*, Discrete Math. **225** (2000), no. 1–3, 51–75.

[15] M. Bousquet-Mélou and M. Petkovšek, *Walks confined in a quadrant are not always D-finite*, Theoret. Comput. Sci. **307** (2003), no. 2, 257–276.

[16] M. Bousquet-Mélou and G. Schaeffer, *Walks on the slit plane*, Probab. Theory Related Fields **124** (2002), no. 3, 305–344.

[17] M. Bousquet-Mélou and G. Xin, *On partitions avoiding 3-crossings*, Sém. Lothar. Combin. **54** (2005/7), article B54e, 21 pp.

[18] R. Cori, S. Dulucq, and G. Viennot, *Shuffle of parenthesis systems and Baxter permutations*, J. Combin. Theory Ser. A **43** (1986), no. 1, 1–22.

[19] E. Duchi and R. A. Sulanke, *The* 2^{n-1} *factor for multi-dimensional lattice paths with diagonal steps*, Sém. Lothar. Combin. **51** (2004), article B51c, 16 pp.

[20] P. Duchon, *On the enumeration and generation of generalized Dyck words*, Discrete Math. **225** (2000), no. 1–3, 121–135.

[21] G. Fayolle, R. Iasnogorodski, and V. Malyshev, *Random Walks in the Quarter-Plane: Algebraic Methods, Boundary Value Problems and Applications*, no. 40 in Applications of Mathematics (New York), Springer-Verlag, New York/Berlin, 1999.

[22] Ph. Flajolet, Analytic models and ambiguity of context-free languages. *Theoret. Comput. Sci.*, **49** (1987), no. 2–3, 283–309.

[23] L. Flatto and S. Hahn, *Two parallel queues created by arrivals with two demands, I*, SIAM J. Appl. Math. **44** (1984), no. 5, 1041–1053.

[24] J. Françon and G. Viennot, *Permutations selon leurs pics, creux, doubles montées et double descentes, nombres d'Euler et nombres de Genocchi*, Discrete Math. **28** (1979), no. 1, 21–35.

[25] I. M. Gessel, *A factorization for formal Laurent series and lattice path enumeration*, J. Combin. Theory Ser. A, **28** (1980), no. 3, 321–337.

[26] I. M. Gessel, *A probabilistic method for lattice path enumeration*, J. Statist. Plann. Inference **14** (1986), no. 1, 49–58.
[27] I. M. Gessel and X. Viennot, *Binomial determinants, paths, and hook length formulae*, Adv. in Math. **58** (1985), no. 3, 300–321.
[28] I. M. Gessel and D. Zeilberger, *Random walk in a Weyl chamber*, Proc. Amer. Math. Soc. **115** (1992), no. 1, 27–31.
[29] D. Gouyou-Beauchamps, *Chemins sous-diagonaux et tableaux de Young*, pp. 112–125 in *Combinatoire Énumérative (Montréal, Québec, 1985)*, no. 1234 in Lecture Notes in Mathematics, Springer-Verlag, New York/Berlin, 1986.
[30] D. Gouyou-Beauchamps, *Standard Young tableaux of height* 4 *and* 5, European J. Combin. **10** (1989), no. 1, 69–82.
[31] R. K. Guy, C. Krattenthaler, and B. E. Sagan, *Lattice paths, reflections, & dimension-changing bijections*, Ars Combin. **34** (1992), 3–15.
[32] E. J. Janse van Rensburg, *Adsorbing staircase walks and staircase polygons*, Ann. Comb. **3** (1999), no. 2–4, 451–473.
[33] E. J. Janse van Rensburg, T. Prellberg, and A. Rechnitzer, *Partially directed paths in a wedge*, J. Combin. Theory Ser. A **115** (2008), no. 4, 623–650.
[34] S. Kantor, *Neue Theorie der eindeutigen periodischen Transformationen in der Ebene*, Acta Math. **19** (1895), no. 1, 115–193.
[35] M. Kauers, C. Koutschan, and D. Zeilberger, *Proof of Ira Gessel's lattice path conjecture*, Proc. Nat. Acad. Sci. USA **106** (2009), no. 28, 11502–11505.
[36] M. Kauers and D. Zeilberger, *The quasi-holonomic Ansatz and restricted lattice walks*, J. Difference Equ. Appl. **14** (2008), no. 10–11, 1119–1126.
[37] G. Kreweras, *Sur une classe de problèmes liés au treillis des partitions d'entiers*, Cahiers du B.U.R.O. **6** (1965), 5–105.
[38] L. Lipshitz, *The diagonal of a D-finite power series is D-finite*, J. Algebra **113** (1988), no. 2, 373–378.
[39] L. Lipshitz, *D-finite power series*, J. Algebra **122** (1989), no. 2, 353–373.
[40] M. Mishna, *Classifying lattice walks restricted to the quarter plane*, J. Combin. Theory Ser. A **116** (2009), no. 2, 460–477.
[41] M. Mishna and A. Rechnitzer, *Two non-holonomic lattice walks in the quarter plane*, Theoret. Comput. Sci. **410** (2009), no. 38–40, 3616–3630.
[42] H. Niederhausen, *The ballot problem with three candidates*, European J. Combin. **4** (1983), no. 2, 161–167.
[43] H. Niederhausen, *A note on the enumeration of diffusion walks in the first octant by their number of contacts with the diagonal*, J. Integer Seq. **8** (2005), no. 4, article 05.4.3, 7 pp.
[44] H. Niederhausen, *Random walks in octants, and related structures*, J. Statist. Plann. Inference **135** (2005), no. 1, 165–196.
[45] M. Petkovšek and H. S. Wilf, *On a conjecture of Ira Gessel*, preprint, 2008. Available as arxiv:0807.3202.
[46] M. Petkovšek, H. S. Wilf, and D. Zeilberger, $A = B$, A. K. Peters Ltd., Wellesley, MA, 1996.
[47] D. Poulalhon and G. Schaeffer, *Counting, Coding, and Sampling with Words*, Chapter 9 of *Applied Combinatorics on Words* (M. Lothaire, ed.), vol. 104 of The Encyclopedia of Mathematics and its Applications, Cambridge Univ. Press, Cambridge, UK, 2005.
[48] A. Regev, *Asymptotic values for degrees associated with strips of Young diagrams*, Adv. in Math. **41** (1981), no. 2, 115–136.
[49] B. E. Sagan, *The Symmetric Group. Representations, Combinatorial Algorithms, and Symmetric Functions*, no. 203 of *Graduate Texts in Mathematics*, Springer-Verlag, New York/Berlin, 2nd edition, 2001.
[50] B. Salvy and P. Zimmermann, GFUN: *A* MAPLE *package for the manipulation of generating and holonomic functions in one variable*, ACM Trans. Math. Software **20** (1994), no. 2, 163–177.
[51] N. J. A. Sloane and S. Plouffe, *The Encyclopedia of Integer Sequences*, Academic Press, San Diego, CA, 1995. Published electronically in a continuously updated form as *The On-Line Encyclopedia of Integer Sequences (OEIS)*, available at `http://www.research.att.com/~njas/sequences`, 2010.
[52] R. P. Stanley, *Enumerative Combinatorics, Vol. 2*, no. 62 in Cambridge Studies in Advanced Mathematics, Cambridge Univ. Press, Cambridge, UK, 1999.

[53] J. R. Stembridge, *Nonintersecting paths, Pfaffians, and plane partitions*, Adv. in Math. **83** (1990), no. 1, 96–131.
[54] A. Wiman, *Zur Theorie der endlichen Gruppen von birationalen Transformationen in der Ebene*, Math. Ann. **48** (1896), no. 1–2, 195–240.
[55] G. Xin and T. Y. J. Zhang, *Enumeration of bilaterally symmetric 3-noncrossing partitions*, Discrete Math. **309** (2009), no. 8, 2497–2509.
[56] D. Zeilberger, *The umbral transfer-matrix method: I. Foundations*, J. Comb. Theory Ser. A **91** (2000), no. 1–2, 451–463.

CNRS, LaBRI, Université Bordeaux 1, 351 cours de la Libération, 33405 Talence Cedex, France
E-mail address: mireille.bousquet@labri.fr
URL: http://www.labri.fr/perso/bousquet

Dept. of Mathematics, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia V5A 1S6, Canada
E-mail address: mmishna@sfu.ca
URL: http:www.math.sfu.ca/~mmishna

This page intentionally left blank

Contemporary Mathematics
Volume **520**, 2010

Quantum random walk on the integer lattice: Examples and phenomena

Andrew Bressler, Torin Greenwood, Robin Pemantle, and Marko Petkovšek

Abstract. We apply results of Baryshnikov, Bressler, Pemantle, and collaborators, to compute limiting probability profiles for various quantum walks in one and two dimensions. Using analytical machinery, we obtain some features of the limit distribution that are not evident in an empirical intensity plot of the time 10,000 distribution. Some conjectures are stated, and computational techniques are discussed as well.

1. Introduction

The quantum walk on the integer lattice is a quantum analogue of the discrete-time finite-range random walk (hence the abbreviation QRW). The process was first constructed in the 1990's by [**ADZ93**], with the idea of using such a process for quantum computing. A mathematical analysis of one particular one-dimensional QRW, called the Hadamard QRW, was put forward in 2001 by [**ABN$^+$01**]. Those interested in a survey of the present state of knowledge may wish to consult [**Kem03**], as well as the more recent expository works [**Ken07, VA08, Kon08**]. Among other properties, they showed that the motion of the quantum walker is ballistic: at time n, the location of the particle is typically found at distance $\Theta(n)$ from the origin. This contrasts with the diffusive behavior of the classical random walk, which is found at distance $\Theta(\sqrt{n})$ from the origin. A rigorous and more comprehensive analysis via several methodologies was given by [**CIR03**], and a thorough study of the general one-dimensional QRW with two chiralities appeared in [**BP07**]. A number of papers on the subject of the quantum walk appear in the physics literature in the early 2000's.

Studies of lattice quantum walks in more than one dimension are less numerous. The first mathematical such study, of which we are aware, is [**IKK04**], though some numerical results are found in [**MBSS02**]. Ballistic behavior is established in [**IKK04**], along with the possibility of bound states. Further aspects of the limiting distribution are discussed in [**WKKK08**]. A rigorous treatment of the general lattice QRW may be found in the preprint [**BBBP08**]. In particular,

2000 *Mathematics Subject Classification.* Primary 05A16; Secondary 41A63, 41A60, 82C10.

Key words and phrases. Rational function, generating function, shape, quantum random walk, ballistic rescaling, feasible region.

The third author was supported in part by NSF Grant no. DMS-063821.

asymptotic formulae are given for the n-step transition amplitudes. Drawing on this work, the present paper examines a number of examples of QRWs in one and two dimensions. We prove the existence of phenomena new to the QRW literature, as well as resolving some computational issues arising in the application of results from [**BBBP08**] to specific quantum walks.

An outline of the remainder of the paper is as follows. In Section 2 we define the QRW and summarize some known results. Section 3 is concerned with one-dimensional QRWs. We develop some theoretical results specific to one dimension, which hold for an arbitrary number of chiralities. We work an example to illustrate the new phenomena as well as some techniques of computation. Section 4 is concerned with examples in two dimensions. In particular, we compute the bounding curves for some examples previously examined in [**BBBP08**].

2. Background

2.1. Construction. To specify a lattice quantum walk one needs the dimension $d \geqslant 1$, the number of chiralities $k \geqslant d+1$, a sequence of k vectors $\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(k)} \in \mathbb{Z}^d$, and a unitary $k \times k$ matrix U. The state space for the QRW is

$$\Omega := L^2\left(\mathbb{Z}^d \times \{1, \ldots, k\}\right) .$$

A Hilbert space basis for Ω is the set of elementary states $\delta_{\mathbf{r},j}$, as $\mathbf{r}$ ranges over $\mathbb{Z}^d$ and $1 \leqslant j \leqslant k$; we shall also denote $\delta_{\mathbf{r},j}$ simply by $(\mathbf{r}, j)$. Let $I \otimes U$ denote the unitary operator on Ω whose value on the elementary state $(\mathbf{r}, j)$ is equal to $\sum_{i=1}^{k} U_{ij}(\mathbf{r}, i)$. Let T denote the operator whose action on the elementary states is given by $T(\mathbf{r}, j) = (\mathbf{r} + \mathbf{v}^{(j)}, j)$. The QRW operator $\mathcal{S} = \mathcal{S}_{d,k,U,\{\mathbf{v}^{(j)}\}}$ is defined by

$$\mathcal{S} := T \cdot (I \otimes U) . \tag{2.1}$$

2.2. Interpretation. The elementary state $(\mathbf{r}, j)$ is interpreted as a particle known to be in location $\mathbf{r}$ and having chirality j. The chirality is an observable that can take k values; chirality and location are simultaneously observable. Introduction of chirality to the model is necessary for the existence of nontrivial translation-invariant unitary operators, as was observed by [**Mey96**]. A single step of the QRW consists of two parts: first, leave the location alone but modify the state by applying U; then leave the state alone and make a deterministic move by an increment $\mathbf{v}^{(j)}$ corresponding to the new chirality, j. The QRW is translation invariant, meaning that if σ is any translation operator $(\mathbf{r}, j) \mapsto (\mathbf{r} + \mathbf{u}, j)$ then $\mathcal{S} \circ \sigma = \sigma \circ \mathcal{S}$. The n-step operator is $\mathcal{S}^n$. Using bracket notation, we denote the amplitude for finding the particle in chirality j and location $\mathbf{x} + \mathbf{r}$ after n steps, starting in chirality i and location $\mathbf{x}$, by

$$a(i, j, \mathbf{r}, n) := \langle (\mathbf{x}, i) \mid \mathcal{S}^n \mid (\mathbf{x} + \mathbf{r}, j) \rangle . \tag{2.2}$$

By translation invariance, this quantity is independent of $\mathbf{x}$. The squared modulus $|a(i, j, \mathbf{r}, n)|^2$ is interpreted as the probability of finding the particle in chirality j and location $\mathbf{x} + \mathbf{r}$ after n steps, starting in chirality i and location $\mathbf{x}$, if a measurement is made. Unlike the classical random walk, the quantum random walk can be measured only at one time without disturbing the process. We may therefore study limit laws for the quantities $a(i, j, \mathbf{r}, n)$, but not joint distributions of these.

2.3. Generating functions. In what follows, we let $\mathbf{x}$ denote the vector $(x_1, \ldots, x_d)$. Given a lattice QRW, for $1 \leqslant i, j \leqslant k$ we may define a power series in $d+1$ variables via

$$F_{ij}(\mathbf{x}, y) := \sum_{n \geqslant 0} \sum_{\mathbf{r} \in \mathbb{Z}^d} a(i, j, \mathbf{r}, n) \mathbf{x}^{\mathbf{r}} y^n \,. \tag{2.3}$$

Here and throughout, $\mathbf{x}^{\mathbf{r}}$ denotes the monomial power $x_1^{r_1} \cdots x_d^{r_d}$. We let $\mathbf{F}$ denote the generating matrix $(F_{ij})_{1 \leqslant i,j \leqslant k}$, which is a $k \times k$ matrix with entries in the ring of formal power series in $d+1$ variables.

LEMMA 2.1 ([**BP07**, Proposition 3.1]). *Let $M(\mathbf{x})$ denote the $k \times k$ diagonal matrix whose diagonal entries are $\mathbf{x}^{\mathbf{v}^{(1)}}, \ldots, \mathbf{x}^{\mathbf{v}^{(k)}}$. Then*

$$\mathbf{F}(\mathbf{x}, y) = (I - y\, M(\mathbf{x}) U)^{-1} \,. \tag{2.4}$$

Consequently, there are polynomials $P_{ij}(\mathbf{x}, y)$ such that

$$F_{ij} = \frac{P_{ij}}{Q} \,, \tag{2.5}$$

where $Q(\mathbf{x}, y) := \det(I - y\, M(\mathbf{x}) U)$.

Let $\mathbf{z}$ denote the vector $(\mathbf{x}, y) \in \mathbb{C}^{d+1}$.

LEMMA 2.2 (torality [**BBBP08**, Proposition 2.1]). *If $Q(\mathbf{z}) = Q(\mathbf{x}, y) = 0$ and $\mathbf{x}$ lies on the unit torus $T^d = \{|x_1| = \cdots = |x_d| = 1\}$ in $\mathbb{C}^d$, then $|y| = 1$, so that $\mathbf{z}$ lies on the unit torus $T^{d+1} := \{|x_1| = \ldots |x_d| = |y| = 1\}$ in $\mathbb{C}^{d+1}$.*

PROOF. If $\mathbf{x} \in T^d$ then $M(\mathbf{x})$ is unitary, hence $M(\mathbf{x})U$ is unitary. And the zeroes of $Q(\mathbf{x}, y)$, in y, are the reciprocals of eigenvalues of $M(\mathbf{x})U$. □

Accordingly, let

$$\mathcal{V} := \left\{ \mathbf{z} \in \mathbb{C}^{d+1} : Q(\mathbf{z}) = 0 \right\}$$

denote the algebraic variety which is the common pole of the generating functions F_{ij}. Let $\mathcal{V}_1 := \mathcal{V} \cap T^{d+1}$ denote the intersection of the singular variety $\mathcal{V}$ with the unit torus $T^{d+1} \subset \mathbb{C}^{d+1}$. An important map on $\mathcal{V}$ is the logarithmic Gauss map $\mu \colon \mathcal{V} \to \mathbb{CP}^d$, introduced as follows. Let $\nabla_{\log} Q \colon \mathbb{C}^{d+1} \to \mathbb{C}^{d+1}$ (and in particular, $\nabla_{\log} Q \colon \mathcal{V} \to \mathbb{C}^{d+1}$) be defined by

$$\nabla_{\log} Q(\mathbf{z}) := \left(z_1 \frac{\partial Q}{\partial z_1}, \ldots, z_{d+1} \frac{\partial Q}{\partial z_{d+1}} \right) = \left(x_1 \frac{\partial Q}{\partial x_1}, \ldots, x_d \frac{\partial Q}{\partial x_d}, y \frac{\partial Q}{\partial y} \right) . \tag{2.6}$$

Its projective counterpart μ is defined by

$$\mu(\mathbf{z}) := \left(z_1 \frac{\partial Q}{\partial z_1} \,:\, \ldots \,:\, z_{d+1} \frac{\partial Q}{\partial z_{d+1}} \right) = \left(x_1 \frac{\partial Q}{\partial x_1} \,:\, \ldots \,:\, x_d \frac{\partial Q}{\partial x_d} \,:\, y \frac{\partial Q}{\partial y} \right) . \tag{2.7}$$

The map μ is defined only at points of $\mathcal{V}$ where the gradient ∇Q does not vanish. In this paper we shall be concerned only with instances of QRW satisfying

$$\nabla Q \text{ vanishes nowhere on } \mathcal{V}_1 \,. \tag{2.8}$$

This condition holds generically.

2.4. Known results. It follows from Lemma 2.2 that the image $\mu[\mathcal{V}_1]$ is contained in the real subspace $\mathbb{RP}^d \subset \mathbb{CP}^d$. Also, under the hypothesis (2.8), $\partial Q/\partial y$ cannot vanish on $\mathcal{V}_1$, hence we may interpret the range of μ as $\mathbb{R}^d \subset \mathbb{RP}^d$ via the identification $(x_1 : \ldots : x_d : y) \leftrightarrow ((x_1/y), \ldots, (x_d/y))$. In what follows, we draw heavily on two results from [**BBBP08**].

THEOREM 2.3 (shape theorem [**BBBP08**, Theorem 4.2]). *Assume* (2.8) *and let $G \subset \mathbb{R}^d$ be the closure of the image of $\mathcal{V}_1$ under μ. If K is any compact subset of G^c, then as $n \to \infty$,*

$$a(i, j, \mathbf{r} = n\hat{\mathbf{r}}, n) = O(e^{-cn})$$

for some $c = c(K) > 0$, uniformly as $\hat{\mathbf{r}}$ varies over K. (One needs $\mathbf{r} = n\hat{\mathbf{r}} \in \mathbb{Z}^d$.)

In other words, under ballistic rescaling, the *feasible region* of non-exponential decay is contained in G. The converse, and much more, is provided by the second result, also from the same theorem. For $\mathbf{z} \in \mathcal{V}_1$, let $\kappa(\mathbf{z})$ denote the curvature of the real hypersurface $\arg \mathcal{V}_1 = (1/\mathrm{i}) \log \mathcal{V}_1 \subset \mathbb{R}^{d+1}$ at the point $\arg \mathbf{z} = (1/\mathrm{i}) \log \mathbf{z}$, where log is applied to vectors coordinatewise and manifolds pointwise.

THEOREM 2.4 (asymptotics in the feasible region). *Suppose Q satisfies* (2.8). *For $\hat{\mathbf{r}} \in G \subset \mathbb{R}^d$, let $Z(\hat{\mathbf{r}}) \subset \mathcal{V}_1 \subset T^{d+1}$ denote the set $\mu^{-1}(\hat{\mathbf{r}})$ of pre-images in $\mathcal{V}_1$ of $\hat{\mathbf{r}}$ under μ. If $\kappa(\mathbf{z}) \neq 0$ for all $\mathbf{z} \in Z(\hat{\mathbf{r}})$, then as $n \to \infty$,*

$$a(i, j, \mathbf{r} = n\hat{\mathbf{r}}, n) = \pm (2\pi |\mathbf{r}^*|)^{-d/2} \left[\sum_{\mathbf{z} \in Z(\hat{\mathbf{r}})} \frac{P_{ij}(\mathbf{z})}{|\nabla_{\log} Q(\mathbf{z})|} |\kappa(\mathbf{z})|^{-1/2} e^{\mathrm{i}\omega_{\mathbf{z}}(\mathbf{r}, n)} \right] + O\left(n^{-(d+1)/2}\right),$$

where $\mathbf{r}^* := (\mathbf{r}, n) = n(\hat{\mathbf{r}}, 1) \in \mathbb{Z}^{d+1}$ and the phase $\omega_{\mathbf{z}}(\mathbf{r}, n)$ of the summand labeled by $\mathbf{z}$ is given by $-\mathbf{r}^* \cdot \arg(\mathbf{z}) - \pi\tau(\mathbf{z})/4$, with $\tau(\mathbf{z}) = (\tau_+ - \tau_-)(\mathbf{z})$ the index of the quadratic form defining the curvature at the point $\arg \mathbf{z} \in \arg \mathcal{V}_1$. The overall $\pm$ is $+$, resp. $-$, if $\nabla_{\log} Q$ is a positive, resp. negative, multiple of $\mathbf{r}^*$, i.e., if $y\,\partial Q/\partial y > 0$, resp. < 0, at all $\mathbf{z} \in \mathcal{V}_1$.

3. One-dimensional QRW with three or more chiralities

3.1. Hadamard QRW. The Hadamard QRW is the one-dimensional QRW with two chiralities that is defined in [**ADZ93**] and analyzed in [**ABN$^+$01**] and [**CIR03**]. It has unitary matrix $U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, which is a constant multiple of a Hadamard matrix, such matrices being ones whose entries are all ± 1. Applying an affine map to the state space, we may assume without loss of generality that the steps $v^{(1)}, v^{(2)}$ are $0, 1$. Up to a rapidly oscillating factor due to a phase difference in two summands in the amplitude, it is shown in these early works that the rescaled amplitudes $n^{1/2} a(i, j, n\hat{r}, n)$ converge as $n \to \infty$ to a profile $f(\hat{r})$ supported on the feasible interval $J := \left[\frac{1}{2} - \frac{\sqrt{2}}{4}, \frac{1}{2} + \frac{\sqrt{2}}{4} \right] \approx [0.15, 0.85]$. The function f is continuous on the interior of J and diverges like $|\hat{r} - \hat{r}_0|^{-1/2}$ near either endpoint $\hat{r}_0$ of J. These results are extended in [**BP07**] to arbitrary unitary matrices. The limiting profiles are all qualitatively similar; a plot for the Hadamard QRW is shown in Figure 1, with the upper envelope showing what happens when the phases of the summands of Theorem 2.4, of which there are only two, line up.

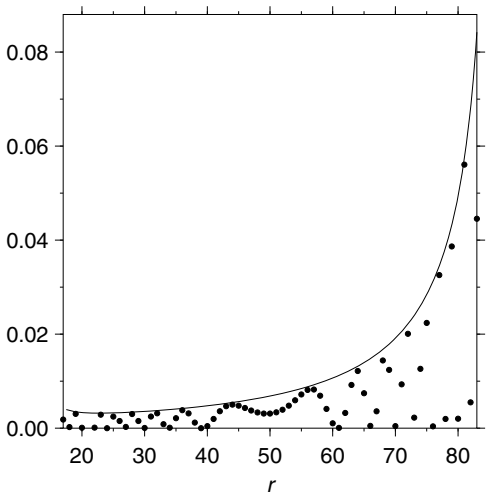


FIGURE 1. The probability profile for the Hadamard QRW at time $n = 100$ on $17 \leqslant r \leqslant 83$, i.e., $\hat{r} = r/n \in [0.17, 0.83] \subset J$. The curve is an upper envelope, computed by aligning the phases of the summands, while the dots are actual squared magnitudes.

3.2. Experimental data with three or more chiralities. When the number of chiralities is allowed to exceed two, new phenomena emerge. The possibility of a bound state arises. This means that for some fixed location r, the amplitude $a(i, j, r, n)$ does not go to zero as $n \to \infty$. This was first shown to occur in [**BCA03, IKS05**]. From a generating function viewpoint, bound states occur when the denominator Q of the generating function factors. The occurrence of bound states appears to be a non-generic phenomenon.

In 2007, two freshman undergraduates, Torin Greenwood and Rajarshi Das, investigated one-dimensional quantum walks with three and four chiralities and more general choices of U and $\{v^{(j)}\}$. Their empirical findings are catalogued in [**GD07**]. The probability profile shown in Figure 2 is typical of what they found, and is the basis for an example running throughout this section. In this example,

$$U = \frac{1}{27}\begin{bmatrix} 17 & 6 & 20 & -2 \\ -20 & 12 & 13 & -4 \\ -2 & -15 & 4 & -22 \\ -6 & -18 & 12 & 15 \end{bmatrix} \tag{3.1}$$

and $v^{(j)} = 1, -1, 0, 2$ for $j = 1, 2, 3, 4$ respectively. The profile shown in the figure is a plot of $|a(1, 1, r, 1000)|^2$ against r, for integers r in the interval $[-1000, 2000]$. The values were computed exactly by recursion, and then plotted. The most obvious new feature, compared to Figure 1, is the existence of a number of peaks in the interior of what is clearly the feasible region. The phase factor is somewhat more chaotic as well, which turns out to be due to a greater number of summands in the amplitude function asymptotics. (See Theorem 2.4.) Our aim is to use the theory described in Section 2 to establish the locations of these peaks, that is to say, the values of $\hat{r}$ for which $n^{1/2}a(i, j, r, n)$ is unbounded as $n \to \infty$, for r sufficiently near $n\hat{r}$.

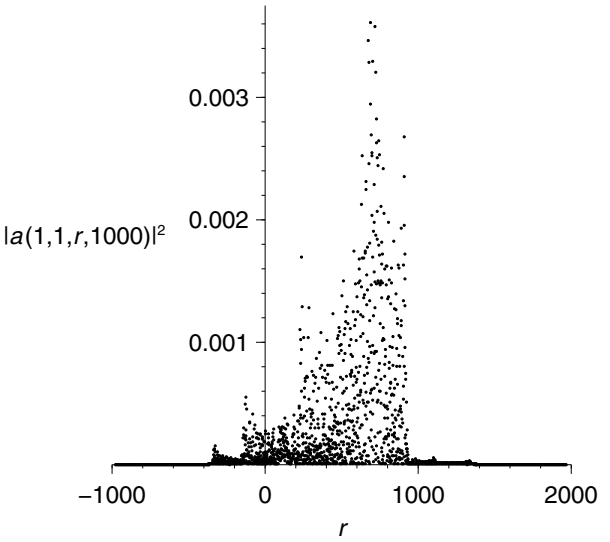


FIGURE 2. The probability profile for the four-chirality QRW at time $n = 1000$.

3.3. Results and conjectures. The results of Section 2 may be summarized informally in the case of one-dimensional QRW as follows. Provided the quantities ∇Q and κ do not vanish for the points z associated with a velocity $\hat{\mathbf{r}}$, the amplitude profile will be a sum of terms whose phase factors may be somewhat chaotic, but whose magnitudes are proportional to $|\kappa|^{-1/2}/\,|\nabla_{\log} Q|$. In practice the magnitude of the amplitude will vary between zero and the sum of the magnitudes of the terms, depending on the behavior of the phases. In the two-chirality case, with only two summands, it is easy to relate empirical data (say, the data in Figure 1) to this asymptotic result. However, in the multi-chirality case, the empirical data in Figure 2 are not easily reconciled with the asymptotic result, firstly because the predicted asymptotics are not trivial to compute, and secondly because the computation appears at first to be at odds with the data. In the remainder of Section 3, we show how the theoretical computations may be executed in a computer algebra system, and then compare the asymptotic result with the empirical data in Figure 2. The first step is to verify some of the hypotheses of Theorems 2.3–2.4. The second step, reconciling the theory and the data, will be done in Section 3.4.

PROPOSITION 3.1. *Let $Q(\mathbf{x}, y)$ be the denominator of the generating function for any QRW (in any dimension d) that satisfies the smoothness hypothesis* (2.8). *Let $\pi\colon \mathcal{V}_1 \to T^d$ be the projection from $\mathcal{V}_1 \subset T^{d+1} \subset \mathbb{C}^{d+1}$ to the d-torus $T^d \subset \mathbb{C}^d$ that forgets the last coordinate. Then the following properties hold.*

(i) *$\partial Q/\partial y$ does not vanish on $\mathcal{V}_1$;*
(ii) *$\mathcal{V}_1$ is a compact d-manifold;*
(iii) *$\pi\colon \mathcal{V}_1 \to T^d$ is smooth and nonsingular;*
(iv) *$\mathcal{V}_1$ is in fact homeomorphic to a union of some number s of d-tori, each mapping smoothly to T^d under π, with the j'th d-torus covering T^d some number n_j times, for $1 \leqslant j \leqslant s$;*

(v) $\kappa\colon \mathcal{V}_1 \to \mathbb{R}$ *vanishes exactly where the determinant of the Jacobian of the map* $\mu\colon \mathcal{V}_1 \to \mathbb{R}^d$ *vanishes;*
(vi) κ *vanishes on* $\mu^{-1}\left[\partial\left\{\mu[\mathcal{V}_1]\right\}\right]$*, the pre-image of the boundary of the image of* $\mathcal{V}_1$ *under* μ*.*

PROOF. The first two conclusions are shown as [**BBBP08**, Proposition 2.2]. The map π is smooth on T^{d+1}, hence on $\mathcal{V}_1$, and nonsingularity follows from the nonvanishing of the partial derivative with respect to y. The fourth conclusion follows from the classification of compact d-manifolds covering the d-torus. For the fifth conclusion, recall that the Gauss–Kronecker curvature of a real hypersurface is defined as the determinant of the Jacobian of the map taking p to the unit normal at p. We have identified projective space with the slice $z_{d+1} = 1$ rather than with the slice $|\mathbf{z}| = 1$, but these are locally diffeomorphic, so the Jacobian of μ still vanishes exactly when κ vanishes. Finally, if an interior point of a manifold maps to a boundary point of the image of the manifold under a smooth map, then the Jacobian vanishes there; hence the last conclusion follows from the fifth. □

An empirical fact is that in all of the several dozen quantum random walks we have investigated, the number of components of $\mathcal{V}_1$ and the degrees of the map π on each component depend on the dimension d and the vector of chiralities, but not on the unitary matrix U.

CONJECTURE 3.2. If $d, k, \mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(k)}$ are fixed and U varies over unitary matrices, then the number of components of $\mathcal{V}_1$ and the degrees of the map π on each component are constant, except for a set of matrices of positive co-dimension.

REMARK. The unitary group is connected, so if the conjecture fails then a transition occurs at which $\mathcal{V}_1$ is not smooth. We know that this happens, resulting in a bound state [**IKS05**]; however in the three-chirality case, the degeneracy does not seem to mark a transition in the topology of $\mathcal{V}_1$.

In the one-dimensional case, the manifold $\mathcal{V}_1$ is a union of topological circles. The map $\mu\colon \mathcal{V}_1 \to \mathbb{R}$ is evidently smooth, so it maps $\mathcal{V}_1$ to a union of intervals. In all catalogued cases, in fact the range of μ is an interval, so we have the following open question:

QUESTION 3.3. Is it possible for the image of μ to be disconnected?

Because μ smoothly maps a union of circles to the real line, the Jacobian of the map μ must vanish at least twice on each circle. Let $\mathcal{W}$ denote the set of $(x, y) \in \mathcal{V}_1$ for which $\kappa(x, y) = 0$. The cardinality of $\mathcal{W}$ is not an invariant (compare, for example, the example in Section 3.4 with the first 4-chirality example of [**GD07**]). This has the following interesting consequence. Again, because the unitary group $\mathcal{U}_k$ is connected, by interpolation there must be some U for which there is a *double* degeneracy in the Jacobian of μ, at some $(x, y) \in \mathcal{V}_1$. This means that the associated Taylor series for $\log y$ on $\mathcal{V}_1$ as a function of $\log x$ will be missing not only its quadratic term, but its cubic term as well. In a scaling window of size $n^{1/2}$ near any peak, it is shown in [**BP07**] that each amplitude is asymptotic to an Airy function. However, with a double degeneracy, the same method yields a quartic-Airy limit instead of the usual cubic-Airy limit. This may be the first combinatorial example of such a limit, and will be discussed in forthcoming work.

Let $W = \{\mathbf{w}^{(s)}\}_{s=1}^{t}$ be a set of vectors in $\mathbb{R}^{d+1}$. Say that W is *rationally degenerate* if when $\mathbf{r}^*$ varies over $\mathbb{Z}^{d+1}$, the t-tuples $(\mathbf{r}^* \cdot \mathbf{w}^{(1)}, \ldots, \mathbf{r}^* \cdot \mathbf{w}^{(t)})$, mod 2π

coordinatewise, are not dense in $(\mathbb{R} \bmod 2\pi)^t$. Generic sets W are rationally nondegenerate because degeneracy requires a number of linear relations to hold over $2\pi\mathbb{Q}$. If W is rationally nondegenerate, then the distribution on t-tuples when $\mathbf{r}^*$ is distributed uniformly over any cube of side $M \geqslant 1$ in $\mathbb{Z}^{d+1}$ will converge weakly as $M \to \infty$ to the uniform distribution on $(\mathbb{R} \bmod 2\pi)^t$. Let $\chi(\alpha_1, \ldots, \alpha_t)$ denote the distribution of the squared modulus of the sum of t complex numbers chosen independently at random with moduli $\alpha_1, \ldots, \alpha_t$ and arguments uniform on $[-\pi, \pi]$.

The result below follows from the preceding discussion, Theorems 2.3 and 2.4, and Proposition 3.1.

PROPOSITION 3.4. *For any d-dimensional QRW, let Q, $Z(\hat{\mathbf{r}})$, and κ be as above. Let $J \subset \mathbb{R}^d$ be the image of $\mathcal{V}_1 \subset T^{d+1}$ under μ. Let $\hat{\mathbf{r}}$ be any point of J such that $\kappa(\mathbf{z}) \neq 0$ for all $\mathbf{z} \in Z(\hat{\mathbf{r}}) \subset \mathcal{V}_1$ and such that the set of $|Z(\hat{\mathbf{r}})|$ $(d+1)$-vectors $W := (1/\mathrm{i}) \log Z(\hat{\mathbf{r}})$, each logarithm being computed coordinatewise, is rationally nondegenerate. Let $\mathbf{r}^*(n) = (\mathbf{r}(n), n)$, $n \geqslant 1$, be a sequence of integer $(d+1)$-vectors with $\mathbf{r}(n)/n \to \hat{\mathbf{r}}$. Then for any $\epsilon > 0$ and any interval $I \subset \mathbb{R}$ there exists an $M \geqslant 1$ such that for sufficiently large n and for each (i, j), the empirical distribution of $[2\pi n \, \|(\hat{\mathbf{r}}, 1)\|]^d$ times the squared moduli of the amplitudes*

$$\left\{ a(i, j, \mathbf{r}(n) + \boldsymbol{\xi}, n + \eta_{d+1}) : \boldsymbol{\eta} = (\boldsymbol{\xi}, \eta_{d+1}) \in \{0, \ldots, M-1\}^{d+1} \right\}$$

gives a weight to the interval I that is within ϵ of the weight given to I by the distribution $\chi(\alpha_1, \ldots, \alpha_t)$, where $t = |Z(\hat{\mathbf{r}})|$, $\{\mathbf{z}^{(s)}\}_{s=1}^t$ enumerates $Z(\hat{\mathbf{r}})$, and

$$\alpha_s = \frac{|P_{ij}(\mathbf{z}_s)|}{|\nabla_{\log} Q(\mathbf{z}_s)|} \, |\kappa(\mathbf{z}_s)|^{-1/2} .$$

That is, the fraction of these M^{d+1} (normalized) squared amplitude moduli that lie in the interval I will differ by less than ϵ from $Pr_\chi(I)$.

If on the other hand $\hat{\mathbf{r}} \notin \overline{J}$, then the empirical distribution for any fixed $M \geqslant 1$ will converge as $n \to \infty$ to the point mass at zero.

REMARK. Rational nondegeneracy becomes more difficult to check as t, the size of $Z(\hat{\mathbf{r}})$, increases, which happens when the number of chiralities increases. If one weakens the conclusion to convergence to some nondegenerate distribution with support in $\tilde{J} := \left[0, \sum_{s=1}^t \left|P_{ij}(\mathbf{z}_s)\kappa(\mathbf{z}_s)^{-1/2}/\nabla_{\log} Q(\mathbf{z}_s)\right|^2\right]$, then one needs only that not all components of all differences $(1/\mathrm{i}) \log \mathbf{z} - (1/\mathrm{i}) \log \mathbf{z}'$ be in $2\pi\mathbb{Q}$, for $\mathbf{z}, \mathbf{z}' \in Z(\hat{\mathbf{r}})$. For the purpose of qualitatively explaining the plots, this is good enough, although the upper envelope may be strictly less than the upper endpoint of $\tilde{J}$ (and the lower envelope be strictly greater than zero), if there is rational degeneracy.

Comparing the $d = 1$ case of this theoretical result to Figure 2, we see that $J \subset \mathbb{R}$ appears to be a proper subinterval of $[-1, 2]$, and that there appear to be up to seven peaks which are local maxima of the probability profile. These include the endpoints of J (cf. the last conclusion of Proposition 3.1) as well as several interior points, which we now understand to be places where the map μ folds back on itself. We now turn our attention to corroborating our understanding of this picture, by computing the number and locations of the peaks.

3.4. Computations. Much of our computation is carried out symbolically in MAPLE. Symbolic computation is significantly faster when the entries of U

are rational, than when they are, say, quadratic algebraic numbers. Also, MAPLE sometimes incorrectly simplifies or fails to simplify expressions involving radicals. It is easy to generate quadratically algebraic orthogonal or unitary matrices via the Gram–Schmidt procedure. For rational matrices, however, we turn to a result we found in [**LO91**].

PROPOSITION 3.5 (Cayley correspondence). *The map* $S \mapsto (I+S)(I-S)^{-1}$ *takes the skew symmetric matrices over a field to the orthogonal matrices over the same field. To generate unitary matrices instead, use skew-hermitian matrices* S.

The map in the proposition is rational, so choosing S to be rational, we obtain orthogonal matrices with rational entries. In our running example,

$$S = \begin{bmatrix} 0 & -3 & -1 & 3 \\ 3 & 0 & 1 & -2 \\ 1 & -1 & 0 & 2 \\ -3 & 2 & -2 & 0 \end{bmatrix},$$

leading to the matrix U of equation (3.1).

The example shows amplitudes for the transition from chirality 1 to chirality 1, so we need the polynomials P_{11} and Q:

$$\begin{aligned} P_{11}(x,y) &= \left(27\,x - 15\,yx^3 - 4\,yx + 12\,y^2x^3 - 12\,y + 4\,y^2x^2 + 9\,y^2 - 17\,y^3x^2\right)x \\ Q(x,y) &= -17\,y^3x^2 + 9\,y^2 + 27\,x - 12\,y + 12\,y^2x^3 + 8\,y^2x^2 - 15\,yx^3 - 4\,y^3x^3 \\ &\quad - 15\,y^3x + 12\,y^2x - 4\,yx - 17\,yx^2 + 9\,y^2x^4 - 12\,y^3x^4 + 27\,y^4x^3\,. \end{aligned}$$

The curvature $\kappa = \kappa(x,y)$ is proportional to

$$(-x\,Q_x - y\,Q_y)\,x\,Q_x\,y\,Q_y - x^2\,y^2\,(Q_y^2\,Q_{xx} + Q_x^2\,Q_{yy} - 2\,Q_x\,Q_y\,Q_{xy})\,,$$

where subscripts denote partial derivatives. Evaluating this in MAPLE 11 leads to xy times a polynomial $K(x,y)$ that occupies about half a page. The command

```
Basis([Q,K],plex(y,x));
```

gives a Gröbner basis, the first element of which is an elimination polynomial, vanishing at precisely those x-values for which there is a pair $(x,y) \in \mathcal{V}$ for which $\kappa(x,y) = 0$. It equals a power of x, times a degree-52 polynomial $p(x)$. We may verify that $\mathcal{V}$, i.e., the curve $Q(x,y) = 0$, is smooth, by computing that the ideal generated by Q, Q_x, Q_y has as basis the trivial basis, [1].

To pass to the subset of the 52 roots of $p(x)$ that are on the unit circle, i.e., that correspond to pairs $(x,y) \in \mathcal{V}_1$, one trick is as follows. If $\tilde{x} = x + 1/x$ then x is on the unit circle if and only if $\tilde{x}$ is in the real interval $[-2,2]$. The polynomial of which the possible $\tilde{x}$ are roots is the elimination polynomial $q(\tilde{x})$ for the basis $[p, 1 - \tilde{x}x + x^2]$, which has degree 26. Applying MAPLE's built-in Sturm sequence evaluator to q shows symbolically that it has exactly six roots $\tilde{x}$ in $[-2,2]$. They lead to six conjugate pairs of x values. The second Gröbner basis element is a polynomial linear in y, so each x value has precisely one corresponding y value. The y value for $\overline{x}$ is the conjugate of the y value for x, and the function μ takes the same value at both points of a conjugate pair. Evaluating the μ function at all six places leads to approximate floating point expressions for $\hat{r} = \mu(x,y)$, namely

$$\hat{r} \approx -0.346306,\ -0.143835,\ 0.229537,\ 0.929248,\ 1.126013,\ 1.362766\,. \tag{3.2}$$

Drawing vertical lines corresponding to these six velocities $\hat{r} = r/n$ yields Figure 3.

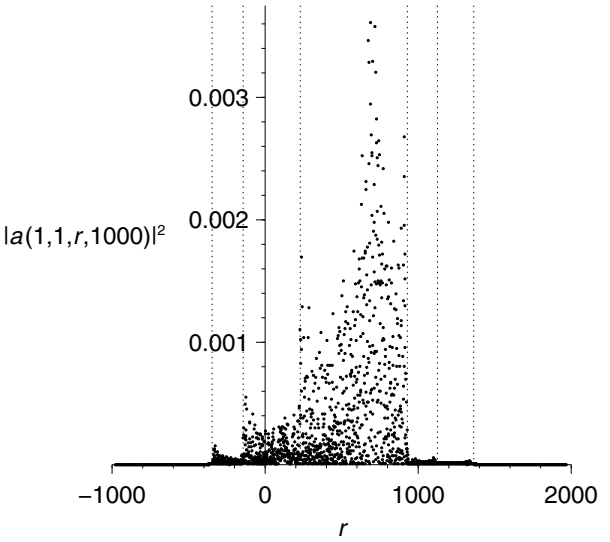


FIGURE 3. The probability profile for the four-chirality QRW at time $n = 1000$, with dotted vertical lines at peak locations.

Surprisingly, the largest peak appearing in the data plot appears to be missing from the set of analytically computed peak velocities. Simultaneously, some of the analytically computed peaks appear quite small, and it seems implausible that the probability profile blows up there. Indeed, this had us puzzled for quite a while. In order to double-check our work, we plotted y against x, resulting in the plot in Figure 4(a), which should be interpreted as having periodic boundary conditions because each of x, y ranges over the unit circle. This shows $\mathcal{V}_1$ to be the union of two topological circles, with the projection of each onto x having degree 2. (Note: each projection onto y has degree 1, and the homology class of each circle is $(2, -1)$ in the basis generated by the x and y axes.) We also plotted $\hat{r} = \mu(x, y)$ against x. To facilitate computation, we used Gröbner bases to eliminate y from the pair $Q = 0$, $xQ_x - \mu y Q_y = 0$, the latter equation being the condition that $(x, y) \in Z(\hat{r})$. This gives a single polynomial equation $S(x, \hat{r}) = 0$ of degree 20 in x and degree 4 in $\hat{r}$, the solution of which is a complex algebraic curve. The intersection of this curve with $|x| = 1$, amounting to a graph of the multivalued map $\hat{r} = \mu(x)$, is shown in Figure 4(b). It shows nicely how the six 'peak' values of $\hat{r}$ given in (3.2), which are indicated by dotted lines, occur at values where the map μ backtracks. By computing the discriminant of $S(x, \hat{r})$ with respect to x, one can readily compute a degree-26 polynomial $P(\hat{r})$, six of the (real) roots of which (i.e., the ones corresponding to $|x| = 1$, $|y| = 1$) are these six values of $\hat{r}$. In full,

$$
\begin{aligned}
P(\hat{r}) := {} & 9938622113821456697978437593 60\,\hat{r}^{26} - 1337213584954984592616036523 7760\,\hat{r}^{25} \\
& + 8066487931437402605871426304512 0\,\hat{r}^{24} - 3169695469803414513463854490245 12\,\hat{r}^{23} \\
& + 1050448354442761227341649604817760\,\hat{r}^{22} - 3170336649899764448701673508335616\,\hat{r}^{21} \\
& + 7568611292278835396888667677296272\,\hat{r}^{20} - 1193053000861081967513176586398795 2\,\hat{r}^{19} \\
& + 9065796993280522601291964101806929\,\hat{r}^{18} + 4759976690500340006135895402266070\,\hat{r}^{17} \\
& - 1951633768753299843090698552226727 1\,\hat{r}^{16} + 1996845044432606007538495311580882 3\,\hat{r}^{15}
\end{aligned}
$$

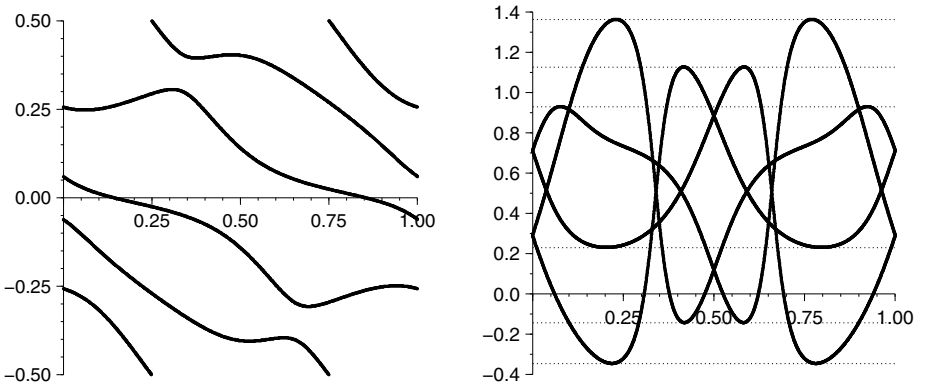


FIGURE 4. Two interleaved circles and their images under the Gauss map. (a) y versus x; (b) $\hat{r}$ (i.e., μ) versus x. As x and y lie on the unit circle, they are represented by $\arg(x)/2\pi$, $\arg(y)/2\pi$.

$$
\begin{aligned}
&- 5443538460557148059355813843071037\,\hat{r}^{14} - 9252724590678726335406645199911997\,\hat{r}^{13}\\
&+ 1191765967443169827579122813077202 1\,\hat{r}^{12} - 4695455477768378466223049515143717\,\hat{r}^{11}\\
&- 1933992724620309233522773490366759\,\hat{r}^{10} + 2691806123752000961762772824527445\,\hat{r}^{9}\\
&- 7782272341402738258511413154541 35\,\hat{r}^{8} - 1549551803567046582527789694383 67\,\hat{r}^{7}\\
&+ 1148504379941690376585059323189 82\,\hat{r}^{6} - 1184727132025417473266193057087 7\,\hat{r}^{5}\\
&- 1148046968669991845399464878870\,\hat{r}^{4} + 1998372452019029124159724934 48\,\hat{r}^{3}\\
&- 2332931429485848868622591028 8\,\hat{r}^{2} + 9678295619028857598464339 04\,\hat{r}\\
&- 1855904649425894505416419 2.
\end{aligned}
$$

This polynomial $P(\hat{r})$ is one of the four irreducible factors of $\mathsf{disc}(S(x,\hat{r}),x)$.

The explanation of the appearance of the extra peak at $\hat{r} = r/n \approx 0.7$ becomes clear if we compare plots at $n = 1000$ and $n = 10000$. (See Figures 5ab.) At first glance, it looks as if the extra peak is still quite prominent in the latter plot, but

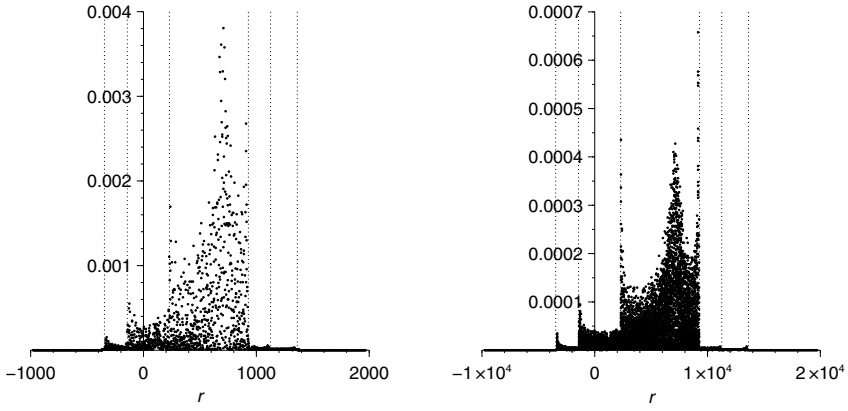


FIGURE 5. As time $n \to \infty$, the extra peak scales down more rapidly than the others. (a) $n = 1000$. (b) $n = 10000$.

in fact it has been lowered with respect to the other peaks. To be precise, the extra peak has gone down by a factor of 10, from 0.004 to 0.0004, indicating that its height scales as n^{-1}. (Its width has remained the same, indicating convergence to a finite probability profile.) The six peaks with $\hat{r}$ values given in (3.2), however, have gone down by factors of $10^{2/3}$, as is known to occur in the Airy scaling windows near velocities $\hat{\mathbf{r}}$ where $\kappa(\mathbf{z}) = 0$ for some $\mathbf{z} \in Z(\hat{\mathbf{r}})$ [**BP07**]. If Figure 5(b) is vertically scaled so that the highest peak has the same height as in Figure 5(a), its width at half the maximum height will shrink somewhat, as must occur in an Airy scaling window, which has width $\sqrt{n}$. The behavior of the extra peak is clearly anomalous.

The extra peak comes from the relatively flat spots on the curve of Figure 4(b), at height $\hat{r} \approx 0.7$. Being nearly horizontal, they generate the extra peak and spread it over a macroscopic rescaled region.

4. Two-dimensional QRW

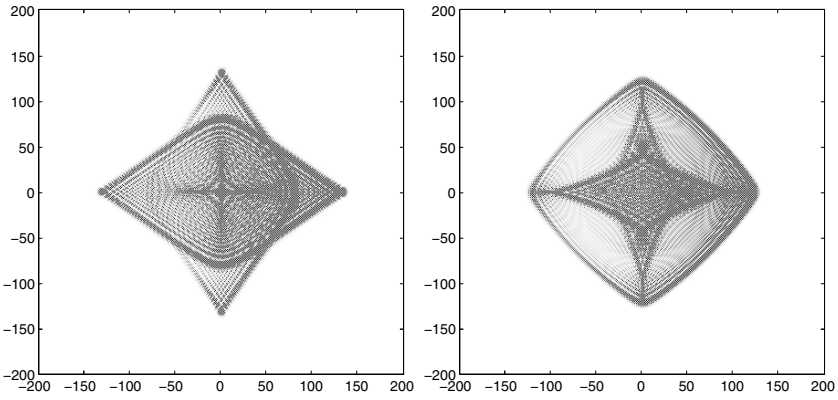
In this section we consider two examples of QRW with $d = 2$, $k = 4$, and steps $\mathbf{v}^{(1)} = (1, 0)$, $\mathbf{v}^{(2)} = (-1, 0)$, $\mathbf{v}^{(3)} = (0, 1)$, and $\mathbf{v}^{(4)} = (0, -1)$. To complete the specification of the two examples, we give the two unitary matrices:

$$U_1 := \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}, \tag{4.1}$$

$$U_2 := \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & -1 & 1 & 1 \end{bmatrix}. \tag{4.2}$$

Note that these are both Hadamard matrices; neither is the Hadamard matrix with the bound state considered in [**Moo04**], nor is either in the two-parameter family referred to as "Grover walks" in [**WKKK08**]. The second differs from the first in that the signs in the third row are reversed. Both are members of one-parameter families analyzed in [**BBBP08**], in Sections 4.1 and 4.3 respectively. The (arbitrary) names given to these matrices in [**Bra07, BBBP08**] are respectively $S(1/2)$ and $B(1/2)$. Intensity plots at time 200 for these two quantum walks, given in Figure 6, reproduce those taken from [**BBBP08**] but with different parameter values (1/2 each time, instead of 1/8 and 2/3 respectively). For the case of U_1 it is shown in [**BBBP08**, Lemma 4.3] that $\mathcal{V}_1$ is smooth. Asymptotics follow, as in Theorem 2.4 of the present paper, and an intensity plot of the asymptotics is generated that matches the empirical time 200 plot quite well. In the case of U_2, $\mathcal{V}_1$ is not smooth but [**BBBP08**, Theorem 3.5] shows that the singular points do not contribute to the asymptotics. Again, a limiting intensity plot follows from Theorem 2.4 of the present paper, and matches the time 200 profile quite well.

It follows from Proposition 3.4 that the union of darkened curves on which the intensity blows up is the algebraic curve where κ vanishes, and that this includes the boundary of the feasible region. The main result of this section is the identification of the algebraic curve. While this result is only computational, it is one of the first examples of computation of such a curve, the only similar prior example being the computation of the "Octic circle" boundary of the feasible region for so-called



(a) U_1 (b) U_2

FIGURE 6. The probability profiles for two QRWs in two dimensions at time $n = 200$: the darkness at (r, s) corresponds to the squared amplitude $|a(1, 1, r, s, 200)|^2$.

diabolo tilings, identified without proof by Cohn and Pemantle and first proved by [**KO07**] (see also [**BP10**]). The perhaps somewhat comical statement of the result is as follows.

THEOREM 4.1. *For the quantum walk with unitary coin flip U_2, the curvature $\kappa = \kappa(\mathbf{z})$ of the variety $\arg \mathcal{V}_1$ vanishes at some $\mathbf{z} = (x, y, z) \in Z(\hat{\mathbf{r}})$ if and only if $\hat{\mathbf{r}} = (\hat{r}, \hat{s})$ is a (real) zero of the polynomial P_2 and satisfies $|\hat{r}| + |\hat{s}| < 3/4$, where*

$$\begin{aligned}
P_2(\hat{r},\hat{s}) := {} & 1 + 14(\hat{r}^2+\hat{s}^2) - 3126(\hat{r}^4+\hat{s}^4) + 97752(\hat{r}^6+\hat{s}^6) - 1445289(\hat{r}^8+\hat{s}^8) \\
& + 12200622(\hat{r}^{10}+\hat{s}^{10}) - 64150356(\hat{r}^{12}+\hat{s}^{12}) + 220161216(\hat{r}^{14}+\hat{s}^{14}) \\
& - 504431361(\hat{r}^{16}+\hat{s}^{16}) + 774608490(\hat{r}^{18}+\hat{s}^{18}) - 785130582(\hat{r}^{20}+\hat{s}^{20}) \\
& + 502978728(\hat{r}^{22}+\hat{s}^{22}) - 184298359(\hat{r}^{24}+\hat{s}^{24}) + 29412250(\hat{r}^{26}+\hat{s}^{26}) \\
& - 1284\,\hat{r}^2\hat{s}^2 - 113016(\hat{r}^2\hat{s}^4+\hat{r}^4\hat{s}^2) + 5220612(\hat{r}^2\hat{s}^6+\hat{r}^6\hat{s}^2) \\
& - 96417162(\hat{r}^2\hat{s}^8+\hat{r}^8\hat{s}^2) + 924427224(\hat{r}^2\hat{s}^{10}+\hat{r}^{10}\hat{s}^2) \\
& - 4865103360(\hat{r}^2\hat{s}^{12}+\hat{r}^{12}\hat{s}^2) + 14947388808(\hat{r}^2\hat{s}^{14}+\hat{r}^{14}\hat{s}^2) \\
& - 27714317286(\hat{r}^2\hat{s}^{16}+\hat{r}^{16}\hat{s}^2) + 30923414124(\hat{r}^2\hat{s}^{18}+\hat{r}^{18}\hat{s}^2) \\
& - 19802256648(\hat{r}^2\hat{s}^{20}+\hat{r}^{20}\hat{s}^2) + 6399721524(\hat{r}^2\hat{s}^{22}+\hat{r}^{22}\hat{s}^2) \\
& - 721963550(\hat{r}^2\hat{s}^{24}+\hat{r}^{24}\hat{s}^2) + 7942218\,\hat{r}^4\hat{s}^4 - 68684580(\hat{r}^4\hat{s}^6+\hat{r}^6\hat{s}^4) \\
& - 666538860(\hat{r}^4\hat{s}^8+\hat{r}^8\hat{s}^4) + 15034322304(\hat{r}^4\hat{s}^{10}+\hat{r}^{10}\hat{s}^4) \\
& - 86727881244(\hat{r}^4\hat{s}^{12}+\hat{r}^{12}\hat{s}^4) + 226469888328(\hat{r}^4\hat{s}^{14}+\hat{r}^{14}\hat{s}^4) \\
& - 296573996958(\hat{r}^4\hat{s}^{16}+\hat{r}^{16}\hat{s}^4) + 183616180440(\hat{r}^4\hat{s}^{18}+\hat{r}^{18}\hat{s}^4) \\
& - 32546593518(\hat{r}^4\hat{s}^{20}+\hat{r}^{20}\hat{s}^4) - 8997506820(\hat{r}^4\hat{s}^{22}+\hat{r}^{22}\hat{s}^4) \\
& + 3243820496\,\hat{r}^6\hat{s}^6 - 25244548160(\hat{r}^6\hat{s}^8+\hat{r}^8\hat{s}^6) + 59768577720(\hat{r}^6\hat{s}^{10}+\hat{r}^{10}\hat{s}^6) \\
& - 147067477144(\hat{r}^6\hat{s}^{12}+\hat{r}^{12}\hat{s}^6) + 458758743568(\hat{r}^6\hat{s}^{14}+\hat{r}^{14}\hat{s}^6) \\
& - 749675452344(\hat{r}^6\hat{s}^{16}+\hat{r}^{16}\hat{s}^6) + 435217945700(\hat{r}^6\hat{s}^{18}+\hat{r}^{18}\hat{s}^6)
\end{aligned}$$

$$
\begin{aligned}
&- 16479111716(\hat{r}^6\hat{s}^{20} + \hat{r}^{20}\hat{s}^6) + 194515866042\,\hat{r}^8\hat{s}^8 \\
&- 421026680628(\hat{r}^8\hat{s}^{10} + \hat{r}^{10}\hat{s}^8) + 611623295476(\hat{r}^8\hat{s}^{12} + \hat{r}^{12}\hat{s}^8) \\
&- 331561483632(\hat{r}^8\hat{s}^{14} + \hat{r}^{14}\hat{s}^8) + 7820601831(\hat{r}^8\hat{s}^{16} + \hat{r}^{16}\hat{s}^8) \\
&+ 72391117294(\hat{r}^8\hat{s}^{18} + \hat{r}^{18}\hat{s}^8) + 421043188488\,\hat{r}^{10}\hat{s}^{10} \\
&- 1131276050256(\hat{r}^{10}\hat{s}^{12} + \hat{r}^{12}\hat{s}^{10}) - 196657371288(\hat{r}^{10}\hat{s}^{14} + \hat{r}^{14}\hat{s}^{10}) \\
&+ 151002519894(\hat{r}^{10}\hat{s}^{16} + \hat{r}^{16}\hat{s}^{10}) + 586397171964\,\hat{r}^{12}\hat{s}^{12} \\
&- 231584205720(\hat{r}^{12}\hat{s}^{14} + \hat{r}^{14}\hat{s}^{12})\,.
\end{aligned}
$$

This polynomial $P_2(\hat{r},\hat{s})$ *has degree-*26 *univariate specializations*

$$
\begin{aligned}
P_2(\hat{r},0) &= (1-\hat{r})^4(1+\hat{r})^4(1-2\hat{r}-7\hat{r}^2)^3(1+2\hat{r}-7\hat{r}^2)^3(1+72\hat{r}^2-291\hat{r}^4+250\hat{r}^6)\,, \\
P_2(\hat{r},\hat{r}) &= (1-8\hat{r}^2)(1-4\hat{r}^2+32\hat{r}^4)^3(1+24\hat{r}^2-3696\hat{r}^4+512\hat{r}^6)^2\,.
\end{aligned}
$$

We may check visually that the zero set of P_2 does indeed coincide with the curves of peak intensity for the U_2 QRW. (See Figure 7.) Before embarking on

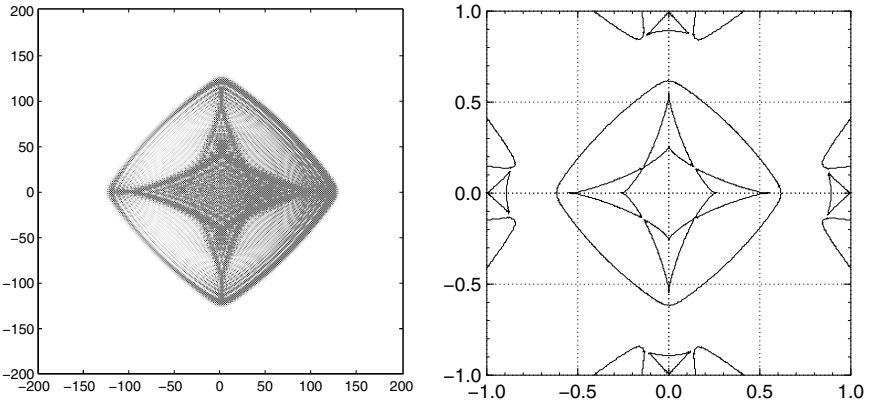


FIGURE 7. A large-time probability profile for the U_2 QRW alongside the graph of the zero set of P_2. (a) The probabilities for $n = 200$ at $\mathbf{r} = (r, s)$. (b) The zero set of P_2 in $\hat{\mathbf{r}} = (r/n, s/n)$.

the proof of Theorem 4.1, let us be clear about what is required. If $\hat{\mathbf{r}}$ is in the boundary of the feasible region G, then κ must vanish at the pre-images of $\hat{\mathbf{r}}$ in the unit torus. The boundary ∂G of the feasible region is therefore a component of a real algebraic variety, W. The variety W is the image under the logarithmic Gauss map μ of the points of the unit torus $T^3 = \{|x| = |y| = |y| = 1\}$ where Q and κ both vanish. Computing this variety is easy in principle: two algebraic equations in $(x, y, z, \hat{r}, \hat{s})$ give the conditions for $\mu(x, y, z) = (\hat{r}, \hat{s})$, and two more give conditions for $Q(x, y, z) = 0$ and $\kappa(x, y, z) = 0$; algebraically eliminating $\{x, y, z\}$ then gives the defining polynomial $P_2(\hat{r}, \hat{s})$ for W. In fact, due to the number of variables and the degree of the polynomials, a straightforward Gröbner basis computation does not work, and we need to use iterated resultants in order to get the computation to halt. The final step is to discard extraneous real zeros of P_2, namely those in the interior of G or G^c, so as to arrive at a precise description of ∂G.

PROOF. The condition for $\mathbf{z} = (x, y, z) \in Z(\hat{r}, \hat{s})$ is given by the vanishing of two polynomials H_1 and H_2 in $(x, y, z, \hat{r}, \hat{s})$, where

$$H_1(x, y, z, \hat{r}, \hat{s}) := xQ_x - \hat{r}zQ_z\,;$$
$$H_2(x, y, z, \hat{r}, \hat{s}) := yQ_y - \hat{s}zQ_z\,.$$

The curvature of $\arg \mathcal{V}_1$ also vanishes when a single polynomial, which we shall call $L(x, y, z)$, vanishes. While explicit formulae for L may be well known in some circles, we include a brief derivation. For $\mathbf{z} = (x, y, z) \in \mathcal{V}_1$, write $x = e^{\mathrm{i}X}, y = e^{\mathrm{i}Y}$ and $z = e^{\mathrm{i}Z}$, so that $\arg \mathbf{z} = (X, Y, Z) \in \arg \mathcal{V}_1$. By Proposition 3.1 we know that $Q_z \neq 0$ on $\mathcal{V}_1$, hence the parametrization of $\mathcal{V}_1$ by X and Y near a point (x, y, z) is smooth and the partial derivatives $Z_X, Z_Y, Z_{XX}, Z_{XY}, Z_{YY}$ are well defined. Implicitly differentiating $Q(e^{\mathrm{i}X}, e^{\mathrm{i}Y}, e^{\mathrm{i}Z(X,Y)}) = 0$ with respect to X and Y, we obtain

$$Z_X = -\frac{xQ_x}{zQ_z} \quad \text{and} \quad Z_Y = -\frac{yQ_y}{zQ_z}\,,$$

and differentiating again yields

$$Z_{XX} = \frac{-\mathrm{i}\,xz}{(zQ_z)^3}\left[Q_xQ_z(zQ_z - 2xzQ_{xz} + xQ_x) + xz(Q_x^2Q_{zz} + Q_z^2Q_{xx})\right]\,;$$
$$Z_{YY} = \frac{-\mathrm{i}\,yz}{(zQ_z)^3}\left[Q_yQ_z(zQ_z - 2yzQ_{yz} + yQ_y) + yz(Q_y^2Q_{zz} + Q_z^2Q_{yy})\right]\,;$$
$$Z_{XY} = \frac{-\mathrm{i}\,xyz}{(zQ_z)^3}\left[zQ_z(Q_zQ_{xy} - Q_xQ_{yz} - Q_yQ_{xz}) + Q_xQ_yQ_z + zQ_xQ_yQ_{zz}\right]\,.$$

In $\mathbb{R}^{d+1}$, the Gaussian curvature of a surface vanishes exactly where the determinant of the Hessian, of any parametrization of the surface as a graph over d variables, vanishes. In particular, the curvature of $\arg \mathcal{V}_1 \subset \mathbb{R}^3$ vanishes where

$$\det \begin{pmatrix} Z_{XX} & Z_{XY} \\ Z_{XY} & Z_{YY} \end{pmatrix}$$

vanishes, and plugging in the computed values yields the polynomial

$$\begin{aligned} L(x, y, z) := {} & xQ_yQ_zQ_x^2 + yQ_xQ_zQ_y^2 + zQ_xQ_yQ_z^2 \\ & + xy(Q_zQ_x^2Q_{yy} + Q_zQ_y^2Q_{xx} - 2Q_xQ_yQ_zQ_{xy}) \\ & + yz(Q_xQ_y^2Q_{zz} + Q_xQ_z^2Q_{yy} - 2Q_xQ_yQ_zQ_{yz}) \\ & + xz(Q_yQ_x^2Q_{zz} + Q_yQ_z^2Q_{xx} - 2Q_xQ_yQ_zQ_{xz}) \\ & + xyz\big[(Q_x^2Q_{yy}Q_{zz} + Q_y^2Q_{zz}Q_{xx} + Q_z^2Q_{xx}Q_{yy}) \\ & \qquad - (Q_x^2Q_{yz}^2 + Q_y^2Q_{xz}^2 + Q_z^2Q_{xy}^2) \\ & \qquad + 2(Q_xQ_yQ_{yz}Q_{xz} + Q_xQ_zQ_{xy}Q_{yz} + Q_yQ_zQ_{xz}Q_{xy}) \\ & \qquad - 2(Q_xQ_yQ_{zz}Q_{xy} + Q_xQ_zQ_{yy}Q_{xz} + Q_yQ_zQ_{xx}Q_{yz})\big]. \end{aligned}$$

It follows that the curvature of $\arg \mathcal{V}_1$ vanishes at $\arg \mathbf{z}$ for some $\mathbf{z} = (x, y, z) \in Z(\hat{r}, \hat{s})$ if and only if the four polynomials Q, H_1, H_2 and L all vanish at some point $(x, y, z, \hat{r}, \hat{s})$ with $(x, y, z) \in T^3$. Ignoring the condition $(x, y, z) \in T^3$ for the moment, we see that we need to eliminate the variables (x, y, z) from the four equations, leading to a one-dimensional ideal in $\hat{r}$ and $\hat{s}$. Unfortunately Gröbner basis computations can have very long run times, with published examples showing for example that the number of steps can be doubly exponential in the number

of variables. Indeed, we were unable to get MAPLE to halt on this computation (indeed, on much smaller computations). The method of resultants, however, led to a quicker elimination computation.

DEFINITION 4.2 (resultant). Let $f(x) := \sum_{j=0}^{\ell} a_j x^j$ and $g(x) := \sum_{j=0}^{m} b_j x^j$ be two polynomials in the single variable x, with coefficients in a field K. Define the resultant $\mathsf{result}(f, g, x)$ to be the determinant of the $(\ell + m) \times (\ell + m)$ matrix

$$\begin{pmatrix}
a_0 & & & & b_0 & & & \\
a_1 & a_0 & & & b_1 & b_0 & & \\
a_2 & a_1 & \ddots & & b_2 & b_1 & \ddots & \\
\vdots & a_2 & \ddots & a_0 & \vdots & b_2 & \ddots & b_0 \\
a_l & \vdots & \ddots & a_1 & b_m & \vdots & \ddots & b_1 \\
 & a_l & \vdots & a_2 & & b_m & \vdots & b_2 \\
 & & \ddots & \vdots & & & \ddots & \vdots \\
 & & & a_l & & & & b_m
\end{pmatrix}.$$

The crucial fact about resultants is the following fact, whose proof may be found in a number of places such as [**CLO98, GKZ94**]:

$$\mathsf{result}(f, g, x) = 0 \iff \exists x : f(x) = g(x) = 0\,. \tag{4.3}$$

Iterated resultants are not quite as nice. For example, if f, g, h are polynomials in x and y, they may be viewed as polynomials in y with coefficients in the field of rational functions, $K(x)$. Then $\mathsf{result}(f, h, y)$ and $\mathsf{result}(g, h, y)$ are polynomials in x, vanishing respectively when the pairs (f, h) and (g, h) have common roots. The quantity

$$R := \mathsf{result}\left(\mathsf{result}(f, h, y), \mathsf{result}(g, h, y), x\right)$$

will then vanish if and only if there is a value of x for which $f(x, y_1) = h(x, y_1) = 0$ and $g(x, y_2) = h(x, y_2) = 0$. It follows that if $f(x, y) = g(x, y) = 0$ then $R = 0$, but the converse does not in general hold. A detailed discussion of this may be found in [**BM09**].

For our purposes, it will suffice to compute iterated resultants and then pass to a subvariety where a common root indeed occurs. We may eliminate repeated factors as we go along. Accordingly, we compute

$$\begin{aligned}
R_{12} &:= \mathcal{R}\mathrm{ad}(\mathsf{result}(Q, L, x))\,, \\
R_{13} &:= \mathcal{R}\mathrm{ad}(\mathsf{result}(Q, H_1, x))\,, \\
R_{14} &:= \mathcal{R}\mathrm{ad}(\mathsf{result}(Q, H_2, x))\,,
\end{aligned}$$

where $\mathcal{R}\mathrm{ad}(P)$ denotes the product of the first powers of each irreducible factor of P. MAPLE is kind to us because we have used the shortest of the four polynomials, Q, in each of the three first-level resultants. Next, we eliminate y via

$$\begin{aligned}
R_{124} &:= \mathcal{R}\mathrm{ad}(\mathsf{result}(R_{12}, R_{14}, y))\,, \\
R_{134} &:= \mathcal{R}\mathrm{ad}(\mathsf{result}(R_{13}, R_{14}, y))\,.
\end{aligned}$$

Polynomials R_{124} and R_{134} each have several small univariate factors, as well as one large multivariate factor which is irreducible over the rationals. Denote the

large factors by f_{124} and f_{134}. Clearly the univariate factors do not contribute to the set we are looking for, so we eliminate z by defining

$$R_{1234} := \mathcal{R}\mathrm{ad}(\mathsf{result}(f_{124}, f_{134}, z)) .$$

MAPLE halts, and we obtain a single polynomial in the variables $(\hat{r}, \hat{s})$ whose zero set contains the set we are after. Let Ω denote the set of $(\hat{r}, \hat{s})$ such that $\kappa(x, y, z) = 0$ for some $(x, y, z) \in \mathcal{V}$ with $\mu(x, y, z) = (\hat{r}, \hat{s})$ [note: this definition uses $\mathcal{V}$ instead of $\mathcal{V}_1$.] It follows from the symmetries of the problem that Ω is symmetric under $\hat{r} \mapsto -\hat{r}$ as well as $\hat{s} \mapsto -\hat{s}$ and the interchange of $\hat{r}$ and $\hat{s}$. Computing iterated resultants, as we have observed, leads to a large zero set Ω'; the set Ω' may not possess $\hat{r}$-$\hat{s}$ symmetry, as this is broken by the choice of order of iteration. Factoring the iterated resultant, we may eliminate any component of Ω' whose image under transposition of $\hat{r}$ and $\hat{s}$ is not in Ω'. Doing so yields the irreducible polynomial P_2. Because the set Ω is algebraic and known to be a subset of the zero set of the irreducible polynomial P_2, we see that Ω is equal to the zero set of P_2.

Let $\Omega_0 \subseteq \Omega$ denote the subset of those $(\hat{r}, \hat{s})$ for which at least one $(x, y, z) \in \mu^{-1}((\hat{r}, \hat{s}))$ with $\kappa(x, y, z) = 0$ lies on the unit torus. It remains to check that Ω_0 consists of those $(\hat{r}, \hat{s}) \in \Omega$ with $|\hat{r}| + |\hat{s}| < 3/4$.

The locus of points in $\mathcal{V}$ at which κ vanishes is a complex algebraic curve γ given by the simultaneous vanishing of Q and L. It is nonsingular as long as ∇Q and ∇L are not parallel, in which case its tangent vector is parallel to $\nabla Q \times \nabla L$. Let $\rho := xQ_x/(zQ_z)$ and $\sigma := yQ_y/(zQ_z)$ be the coordinates of the map μ under the identification of $\mathbb{CP}^2$ with $\{(\hat{r}, \hat{s}, 1) : \hat{r}, \hat{s} \in \mathbb{C}\}$. The image of γ under μ (and this identification) is a nonsingular curve in the plane, provided that γ is nonsingular and either $\mathrm{d}\rho$ or $\mathrm{d}\sigma$ is nonvanishing on the tangent. For this it is sufficient that one of the two determinants $\det M_\rho, \det M_\sigma$ not vanish, where the columns of M_ρ are $\nabla Q, \nabla L, \nabla\rho$ and the columns of M_σ are $\nabla Q, \nabla L, \nabla\sigma$.

Let (x_0, y_0, z_0) be any point in $\mathcal{V}_1$ at which one of these two determinants does not vanish. It follows from Lemma 2.2 that the tangent vector to γ at (x_0, y_0, z_0) in logarithmic coordinates is real; therefore the image of γ near (x_0, y_0, z_0) is a nonsingular real curve. Removing singular points from the zero set of P_2 leaves a union $\mathcal{U}$ of connected components, each of which therefore lies in Ω_0 or is disjoint from Ω_0. The proof of the theorem is now reduced to listing the components, checking that none crosses the boundary $|\hat{r}| + |\hat{s}| = 3/4$, and checking $Z(\hat{r}, \hat{s})$ for a single point $(\hat{r}, \hat{s})$ on each component (note: any component intersecting $\{|\hat{r}| + |\hat{s}| > 1\}$ need not be checked as we know the coefficients to be identically zero here). □

We close by stating a result for U_1, analogous to Theorem 4.1. The proof is entirely analogous as well, and will be omitted.

THEOREM 4.3. *For the quantum walk with unitary coin flip U_1, the curvature $\kappa = \kappa(\mathbf{z})$ of the variety $\arg \mathcal{V}_1$ vanishes at some $\mathbf{z} = (x, y, z) \in Z(\hat{\mathbf{r}})$ if and only if $\hat{\mathbf{r}} = (\hat{r}, \hat{s})$ is a (real) zero of the polynomial P_1 and satisfies $|\hat{r}| + |\hat{s}| \leqslant 2/3$, where*

$$\begin{aligned} P_1(\hat{r}, \hat{s}) := {} & 16(\hat{r}^6 + \hat{s}^6) - 56(\hat{r}^8 + \hat{s}^8) - 1543(\hat{r}^{10} + \hat{s}^{10}) + 14793(\hat{r}^{12} + \hat{s}^{12}) \\ & - 59209(\hat{r}^{14} + \hat{s}^{14}) + 132019(\hat{r}^{16} + \hat{s}^{16}) - 176524(\hat{r}^{18} + \hat{s}^{18}) \\ & + 141048(\hat{r}^{20} + \hat{s}^{20}) - 62208(\hat{r}^{22} + \hat{s}^{22}) + 11664(\hat{r}^{24} + \hat{s}^{24}) + 256\,\hat{r}^2\hat{s}^2 \\ & - 1472(\hat{r}^2\hat{s}^4 + \hat{r}^4\hat{s}^2) - 23060(\hat{r}^2\hat{s}^6 + \hat{r}^6\hat{s}^2) + 291173(\hat{r}^2\hat{s}^8 + \hat{r}^8\hat{s}^2) \end{aligned}$$

$$
\begin{aligned}
&- 1449662(\hat{r}^2\hat{s}^{10} + \hat{r}^{10}\hat{s}^2) + 4140257(\hat{r}^2\hat{s}^{12} + \hat{r}^{12}\hat{s}^2) \\
&- 7492584(\hat{r}^2\hat{s}^{14} + \hat{r}^{14}\hat{s}^2) + 8790436(\hat{r}^2\hat{s}^{16} + \hat{r}^{16}\hat{s}^2) \\
&- 6505200(\hat{r}^2\hat{s}^{18} + \hat{r}^{18}\hat{s}^2) + 2763072(\hat{r}^2\hat{s}^{20} + \hat{r}^{20}\hat{s}^2) \\
&- 513216(\hat{r}^2\hat{s}^{22} + \hat{r}^{22}\hat{s}^2) - 19343\,\hat{r}^4\hat{s}^4 + 72718(\hat{r}^4\hat{s}^6 + \hat{r}^6\hat{s}^4) \\
&+ 1647627(\hat{r}^4\hat{s}^8 + \hat{r}^8\hat{s}^4) - 12711677(\hat{r}^4\hat{s}^{10} + \hat{r}^{10}\hat{s}^4) + 39759700(\hat{r}^4\hat{s}^{12} + \hat{r}^{12}\hat{s}^4) \\
&- 67173440(\hat{r}^4\hat{s}^{14} + \hat{r}^{14}\hat{s}^4) + 64689624(\hat{r}^4\hat{s}^{16} + \hat{r}^{16}\hat{s}^4) \\
&- 33614784(\hat{r}^4\hat{s}^{18} + \hat{r}^{18}\hat{s}^4) + 7363872(\hat{r}^4\hat{s}^{20} + \hat{r}^{20}\hat{s}^4) + 3183044\,\hat{r}^6\hat{s}^6 \\
&- 13374107(\hat{r}^6\hat{s}^8 + \hat{r}^8\hat{s}^6) + 2503464(\hat{r}^6\hat{s}^{10} + \hat{r}^{10}\hat{s}^6) + 72282208(\hat{r}^6\hat{s}^{12} + \hat{r}^{12}\hat{s}^6) \\
&- 153035200(\hat{r}^6\hat{s}^{14} + \hat{r}^{14}\hat{s}^6) + 128187648(\hat{r}^6\hat{s}^{16} + \hat{r}^{16}\hat{s}^6) \\
&- 40374720(\hat{r}^6\hat{s}^{18} + \hat{r}^{18}\hat{s}^6) + 18664050\,\hat{r}^8\hat{s}^8 - 10639416(\hat{r}^8\hat{s}^{10} + \hat{r}^{10}\hat{s}^8) \\
&+ 92321584(\hat{r}^8\hat{s}^{12} + \hat{r}^{12}\hat{s}^8) - 197271552(\hat{r}^8\hat{s}^{14} + \hat{r}^{14}\hat{s}^8) \\
&+ 121508208(\hat{r}^8\hat{s}^{16} + \hat{r}^{16}\hat{s}^8) + 14725472\,\hat{r}^{10}\hat{s}^{10} \\
&+ 100227200(\hat{r}^{10}\hat{s}^{12} + \hat{r}^{12}\hat{s}^{10}) - 227481984(\hat{r}^{10}\hat{s}^{14} + \hat{r}^{14}\hat{s}^{10}) \\
&+ 279234496\,\hat{r}^{12}\hat{s}^{12}\,.
\end{aligned}
$$

This polynomial $P_1(\hat{r},\hat{s})$ has degree-24 univariate specializations

$$
\begin{aligned}
P_1(\hat{r},0) &= \hat{r}^6(1-\hat{r})^3(1+\hat{r})^3(2-3\hat{r})^2(2+3\hat{r})^2(1+4\hat{r}^2-88\hat{r}^4+208\hat{r}^6-144\hat{r}^8)\,, \\
P_1(\hat{r},\hat{r}) &= \hat{r}^4(1-2\hat{r})^2(1+2\hat{r})^2(16-27\hat{r}^2-2416\hat{r}^4-144\hat{r}^6-128\hat{r}^8)^2\,.
\end{aligned}
$$

5. Summary

We have stated an asymptotic amplitude theorem for general one-dimensional quantum walk with an arbitrary number of chiralities and shown how the theoretical result corresponds, not always in an obvious way, to data generated at times of order several hundred to several thousand. We have stated a general shape theorem for two-dimensional quantum walks. The boundary is a part of an algebraic curve, and we have shown how this curve may be computed, both in principle and in a MAPLE computation that halts before running out of memory.

Acknowledgements. The authors are indebted to the editors of this volume for improvements to the manuscript going far beyond the call of duty. These include not only significant corrections to the statements of Theorem 2.4 and Proposition 3.4, but also reverse engineered computations to recover the correct initial vectors and unitary matrix in Section 3.2. The univariate polynomials presented in Theorems 4.1 and 4.3 and at the end of Section 3.4 are also due to the editors.

References

[ABN+01] A. Ambainis, E. Bach, A. Nayak, A. Vishwanath, and J. Watrous, *One-dimensional quantum walks*, pp. 37–49 in *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing*, ACM Press, New York, 2001.

[ADZ93] Y. Aharonov, L. Davidovich, and N. Zagury, *Quantum random walks*, Phys. Rev. A **48** (1993), no. 2, 1687–1690.

[BBBP08] Y. Baryshnikov, W. Brady, A. Bressler, and R. Pemantle, *Two-dimensional quantum random walk*, preprint, 34 pp., 2008. Available as arXiv:0810.5495.

[BCA03] T. A. Brun, H. A. Carteret, and A. Ambainis, *Quantum walks driven by many coins*, Phys. Rev. A **67** (2003), no. 5, article 052317, 17 pp.

[BM09] L. Busé and B. Mourrain, *Explicit factors of some iterated resultants and discriminants*, Math. Comp. **78** (2009), no. 265, 345–386.

[BP07] A. Bressler and R. Pemantle, *Quantum random walks in one dimension via generating functions*, pp. 403–414 in *2007 Conference on Analysis of Algorithms, AofA 07*, Assoc. Discrete Mathematics and Theoretical Computer Science (DMTCS), Nancy, France, 2007.

[BP10] Y. Baryshnikov and R. Pemantle, *The "octic circle" theorem for diabolo tilings: A generating function with a quartic singularity*, manuscript in progress, 2010.

[Bra07] W. Brady, *Quantum random walks on* $\mathbb{Z}^2$, M.Phil. thesis, The University of Pennsylvania, 2007.

[CIR03] H. A. Carteret, M. E. H. Ismail, and B. Richmond, *Three routes to the exact asymptotics for the one-dimensional quantum walk*, J. Phys. A **36** (2003), no. 33, 8775–8795.

[CLO98] D. A. Cox, J. B. Little, and D. O'Shea, *Using Algebraic Geometry*, no. 185 in Graduate Texts in Mathematics, Springer-Verlag, New York/Berlin, 1998.

[GKZ94] I. Gel'fand, M. Kapranov, and A. Zelevinsky, *Discriminants, Resultants and Multidimensional Determinants*, Birkhäuser, Boston/Basel, 1994.

[GD07] T. Greenwood and R. Das, *Combinatorics of Quantum Random Walks on Integer Lattices in One or More Dimensions*, on-line repository, 2007. Currently available at `http://www.math.upenn.edu/~pemantle/Summer2007/First_Page.html`.

[IKK04] N. Inui, Y. Konishi, and N. Konno, *Localization of two-dimensional quantum walks*, Phys. Rev. A **69** (2004), no. 5, article 052323, 9 pp.

[IKS05] N. Inui, N. Konno, and E. Segawa, *One-dimensional three-state quantum walk*, Phys. Rev. E **72** (2005), no. 5, article 056112, 7 pp.

[Kem03] J. Kempe, *Quantum random walks: An introductory overview*, Contemp. Phys. **44** (2003), no. 4, 307–327.

[Ken07] V. Kendon, *Decoherence in quantum walks: A review*, Math. Structures Comput. Sci. **17** (2007), no. 6, 1169–1220.

[KO07] R. Kenyon and A. Okounkov, *Limit shapes and the complex Burgers equation*, Acta Math. **199** (2007), no. 2, 263–302.

[Kon08] N. Konno, *Quantum walks for computer scientists*, pp. 309–452 in *Quantum Potential Theory* (M. Schürmann, U. Franz, and P. Biane, eds.), no. 1954 in Lecture Notes in Mathematics, Springer-Verlag, New York/Berlin, 2008.

[LO91] H. Liebeck and A. Osborne, *The generation of all rational orthogonal matrices*, Amer. Math. Monthly **98** (1991), no. 2, 131–133.

[MBSS02] T. D. Mackay, S. D. Bartlett, L. T. Stephenson, and B. Sanders, *Quantum walks in higher dimensions*, J. Phys. A **35** (2002), no. 12, 2745–2754.

[Mey96] D. Meyer, *From quantum cellular automata to quantum lattice gases*, J. Statist. Phys. **85** (1996), no. 5–6, 551–574.

[Moo04] Cristopher Moore, unpublished e-mail on the subject of quantum walks, 2004. Available in the Domino Archive.

[VA08] S. E. Venegas-Andraca, *Quantum Walks for Computer Scientists*, in *Synthesis Lectures on Quantum Computing*, Morgan and Claypool, San Rafael, CA, 2008.

[WKKK08] K. Watabe, N. Kobayashi, M. Katori, and N. Konno, *Limit distributions of two-dimensional quantum walks*, Phys. Rev. A **77** (2008), no. 6, article 062331, 9 pp.

Department of Mathematics, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104, USA
E-mail address: andrewbr@math.upenn.edu
URL: http://www.math.upenn.edu/~andrewbr

Department of Mathematics, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104, USA
E-mail address: toringr@sas.upenn.edu

Department of Mathematics, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104, USA
E-mail address: pemantle@math.upenn.edu
URL: http://www.math.upenn.edu/~pemantle

Department of Mathematics, University of Ljubljana, Jadranska 19, SI-1000 Ljubljana, Slovenia
E-mail address: marko.petkovsek@fmf.uni-lj.si
URL: http://www.fmf.uni-lj.si/~petkovsek

Contemporary Mathematics
Volume **520**, 2010

A case study in bivariate singularity analysis

Timothy DeVries

Abstract. The multivariate singularity analysis of Pemantle and Wilson is explored, and used to derive an asymptotic expression for the number of bicolored supertrees, counted by the number of nodes. This sequence can be realized as the diagonal coefficient sequence of a bivariate rational function F. Its asymptotics have been obtained previously by univariate methods, but the analysis contained herein serves as a case study for the general multivariate method. The analysis relies heavily on the structure of a height function h along the pole set $\mathcal{V}$ of F. What makes this example interesting is the geometry of h on $\mathcal{V}$, namely that h has a degenerate saddle point away from the boundary of the domain of analyticity of F, which contributes to the asymptotics of the coefficients. Due to the geometry of h at this point, the asymptotic expression cannot be computed directly from the standard formulas of multivariate singularity analysis. Performing the analysis in this case represents a first step towards understanding more general cases of this geometric type.

1. Introduction

Let $F(x,y) = P(x,y)/Q(x,y) = \sum_{r,s\geqslant 0} a_{r,s}x^r y^s$ be the rational generating function defined by

$$\begin{aligned} P(x,y) &= 2x^2y\left(2x^5y^2 - 3x^3y + x + 2x^2y - 1\right), \\ Q(x,y) &= x^5y^2 + 2x^2y - 2x^3y + 4y + x - 2. \end{aligned} \tag{1.1}$$

In this paper we use multivariate singularity analysis to derive the asymptotic estimate

$$a_{n,n} \sim \frac{4^n}{8\Gamma(3/4)\, n^{5/4}}. \tag{1.2}$$

While this result has been obtained previously by univariate methods, the multivariate analysis of this sequence proves to be interesting to the study of multivariate singularity analysis itself. Specifically, this example is interesting because it does not lend itself to automatic computation by the formulas set forth in [**PW02**], the main reference for the multivariate techniques we shall use. In what follows, we briefly discuss how this affects the ultimate goal of producing automatic asymptotic analyses for multivariate rational generating functions. Before proceeding, however, it is important to discuss how the function F originates.

2000 *Mathematics Subject Classification.* Primary 41A60; Secondary 41A63, 05A16.

The coefficients $a_{n,n}$ count *bicolored supertrees*, which are defined as follows. First, denote by $\mathcal{G}$ the class of *Catalan trees*, i.e., rooted, unlabelled, planar trees, counted by the number of nodes. The class $\mathcal{G}$ has generating function

$$G(x) = \frac{1}{2}\left(1 - \sqrt{1-4x}\right),$$

whose coefficients are the Catalan numbers. Denote by $\tilde{\mathcal{G}}$ the class of *bicolor-planted* Catalan trees: Catalan trees with an extra red or blue node attached to the root (likewise counted by the number of nodes). The class $\tilde{\mathcal{G}}$ has generating function

$$\tilde{G}(x) = 2xG(x).$$

Then the class of bicolored supertrees, denoted by $\mathcal{K}$, is defined by the combinatorial substitution $\mathcal{K} = \mathcal{G} \circ \tilde{\mathcal{G}}$. That is, the elements of $\mathcal{K}$ are Catalan trees with each node replaced by a bicolor-planted Catalan tree. The class $\mathcal{K}$ has algebraic generating function $K(x) = G(\tilde{G}(x))$. More explicitly,

$$K(x) = \frac{1}{2} - \frac{1}{2}\sqrt{1 - 4x + 4x\sqrt{1-4x}} = 2x^2 + 2x^3 + 8x^4 + 18x^5 + 64x^6 + O(x^7),$$

with coefficients from [**Slo10**, seq. no. `A168506`]. Denote by k_n the coefficient of x^n in the expansion of $K(x)$ above, i.e., the number of bicolored supertrees having n nodes. An asymptotic estimate for k_n has been obtained by univariate analysis of $K(x)$ [**FS09**, examples VI.10 and VII.20].

As it turns out, $a_{n,n} = k_n$, and $F(x, y)$ was produced by Alex Raichev and Mark Wilson to have this exact property (see [**RW08**]). This was done using Safonov's algorithm, which is roughly a procedure for realizing multivariate algebraic generating functions as so-called diagonals of rational generating functions in one higher variable [**Saf00**]. The idea proposed by Raichev and Wilson was to use Safonov's algorithm to reduce asymptotic computations on *algebraic* generating functions to those on *rational* generating functions, albeit in one higher dimension. In the case of bicolored supertrees, however, the multivariate analysis is not straightforward, for reasons we shall presently discuss.

The technique of multivariate singularity analysis has the following basic structure: begin with Cauchy's Integral Formula, manipulate the integral/integrand, and end with saddle point integration. To be more explicit, let $F = \eta/Q$ be the generating function of the coefficients to be analyzed, with $\eta\colon \mathbb{C}^d \to \mathbb{C}$ entire and $Q \in \mathbb{Q}[x_1, \ldots, x_d]$. Cauchy's Integral Formula expresses the coefficients as an integral of a particular d-form. By appropriately adjusting this integral, we can rewrite it as the integral of a related $(d-1)$-form defined on the variety $\mathcal{V}_Q = \{\mathbf{x} : Q(\mathbf{x}) = 0\}$, along a cycle $C \subseteq \mathcal{V}_Q$. We define a height function h on the variety $\mathcal{V}_Q$ related to the rate of decay of this new integrand. We then 'push down' the cycle C along $\mathcal{V}_Q$, minimizing the maximum of h along C at critical points of the function h. Under the right conditions, the coefficients can finally be approximated as saddle point integrals along C in small neighborhoods of a finite set of these critical points, known as the *contributing points*. In the case $d = 2$, all critical points of h are actually *saddle points* of h, i.e., critical points of h which are not local extrema.

This technique has been well studied (see [**PW02**] and [**PW08**]), to the point that bivariate rational generating function asymptotics can be computed automatically in many cases. Specifically, when all the contributing points are *minimal*—that is, on the generating function's boundary of convergence—then an explicit

algorithm exists for determining which critical points contribute and computing the saddle point integral near these points (in the bivariate case). And when the generating function is *combinatorial*, i.e., when all its coefficients are non-negative, then the contributing points will all be minimal (under the standing assumption of [**PW08**, Assumption 3.6]). In the case examined presently, however, the generating function is *not* combinatorial. Thus there is no guarantee that any contributing points will be minimal. In fact, we shall show that the contributing point is *not* minimal.

Thus, [**PW02**] does not provide us with the locations of the contributing points. Worse than that, however, is that even once the contributing points have been found, there is no formula automatically producing the correct saddle point computation in a neighborhood of these points. This is because the contour followed by C in a neighborhood of the contributing points is not automatically known. (However, for minimal contributing points, an explicit path for C near these points is known; see [**PW02**]). This is particularly bad when the contributing point is a degenerate saddle point for the height function. Since the height on C is locally maximized at the contributing point, it must locally approach and depart along ascent and descent paths. A greater degree of degeneracy means more ascent/descent paths, hence more possibilities for the local path followed by C. And indeed in the case presently studied, the contributing point *is* a degenerate saddle point of the height function.

Understanding the saddle point integration near these degenerate saddles is particularly important because degenerate saddles arise frequently in combinatorial applications (despite the fact that they are nongeneric). A careful analysis of [**PW02**] reveals that, in the absence of such degenerate saddles, one obtains leading term asymptotics only of the form $cA^n n^{p/2}$ (for constants c, A and integer p). By Safonov's algorithm, any univariate algebraic generating function can be realized as the diagonal of a bivariate rational generating function. But by univariate asymptotic methods, we know that the coefficients of such univariate functions can produce leading term asymptotics of the form $cA^n n^{p/q}$ for *arbitrary* $q \in \mathbb{N}$ (see [**FS09**, § VII.7]), and so a multivariate analysis of the corresponding bivariate rational function should turn up a degenerate saddle whenever $q > 2$.

Understanding this example represents a first step towards understanding what to do when the contributing saddle point is not on the boundary of convergence, and when the height function has a degenerate saddle at this point. The rest of this paper is concerned with examining this example. In Section 2 we shall present the theory needed to reduce Cauchy's Integral Formula to a series of saddle point integrals, as outlined above. In Section 3 we shall apply these methods to the function $F(x, y)$ defined earlier, obtaining our asymptotic estimate.

2. Multivariate singularity analysis

2.1. Coefficient representation. For the duration of this paper, let $F\colon \mathbb{C}^d \to \mathbb{C}$ be a function analytic in a neighborhood of the origin, having series representation

$$F(\mathbf{x}) = \sum_{\mathbf{r} \in \mathbb{N}^d} a_{\mathbf{r}} \mathbf{x}^{\mathbf{r}},$$

where $\mathbf{x}^{\mathbf{r}}$ is shorthand notation for $x_1^{r_1} \cdot \ldots \cdot x_d^{r_d}$. The goal is to obtain an asymptotic expansion for the coefficients $a_{\mathbf{r}}$ given F, and the main tool for this is Cauchy's Integral Formula.

THEOREM 2.1 (Cauchy's Integral Formula). *Let F be as above, analytic in a polydisc $D_0 = \{\mathbf{x} : |x_j| < \varepsilon_j \; \forall j\}$, for some positive, real ε_j. Assume further that F is continuous on the boundary torus $T_0 = \partial D_0$, a product of loops around the origin in each coordinate, each one positively oriented with respect to the complex orientation of its respective plane. Then*

$$a_{\mathbf{r}} = \int_{T_0} \omega_F,$$

where

$$\omega_F = \frac{1}{(2\pi \mathrm{i})^d} \cdot \frac{F(\mathbf{x})}{x_1 \cdot \ldots \cdot x_d} \mathbf{x}^{-\mathbf{r}} \, \mathrm{d}\mathbf{x}.$$

Cauchy's Integral Formula can be found in most textbooks presenting complex analysis in a multivariable setting, and follows easily as an iterated form of the single variable formula. See, for example, [**Sha92**, p. 19].

We wish to use the structure of Cauchy's formula to obtain an asymptotic formula for $a_{\mathbf{r}}$ as $\mathbf{r} \to \infty$, but first we need to be more precise about what is meant by "$\mathbf{r} \to \infty$." There are many ways to send the vector $\mathbf{r}$ to infinity, but one of the most natural ways is to fix a direction in the positive d-hyperoctant and send $\mathbf{r}$ to infinity along this direction. Specifically, define the $(d-1)$-simplex Δ^{d-1} by

$$\Delta^{d-1} = \left\{ (\hat{r}_1, \ldots, \hat{r}_d) : \hat{r}_j \geqslant 0 \; \forall j, \; \sum_{j=1}^{d} \hat{r}_j = d \right\},$$

where we choose the convention that the $\hat{r}_j$ sum to d for later notational convenience. Then any $\mathbf{r}$ in the positive d-hyperoctant can be written uniquely as $\mathbf{r} = |\mathbf{r}| \, \hat{\mathbf{r}}$, where $|\mathbf{r}| \in \mathbb{R}^+$ and $\hat{\mathbf{r}} \in \Delta^{d-1}$. We examine $\mathbf{r}$ as $|\mathbf{r}| \to \infty$ and $\hat{\mathbf{r}} \to \hat{\mathbf{r}}_0$ for some fixed direction $\hat{\mathbf{r}}_0 \in \Delta^{d-1}$.

Now we turn to the structure of the integrand ω_F, specifically $\mathbf{x}^{-\mathbf{r}}$ (the portion that changes as $\mathbf{r}$ varies). With an eye on the end goal of reducing our computation to a saddle integral, we use the following representation (away from the coordinate axes):

$$\mathbf{x}^{-\mathbf{r}} = \exp\left(-\sum_{j=1}^{d} r_j \ln x_j \right) = \exp\left(|\mathbf{r}| \, H_{\hat{\mathbf{r}}}(\mathbf{x}) \right),$$

where

$$H_{\hat{\mathbf{r}}}(\mathbf{x}) = -\sum_{j=1}^{d} \hat{r}_j \ln x_j. \tag{2.1}$$

When no confusion exists, we shall simply refer to the function $H_{\hat{\mathbf{r}}}$ as H. The overall magnitude of the integrand will be an important factor in computing an asymptotic expansion for $a_{\mathbf{r}}$, so we next examine the magnitude of $\exp\left(|\mathbf{r}| \, H_{\hat{\mathbf{r}}}(\mathbf{x}) \right)$. We have

$$\left| \exp\left(|\mathbf{r}| H_{\hat{\mathbf{r}}}(\mathbf{x}) \right) \right| = \exp\left(|\mathbf{r}| \, \Re H_{\hat{\mathbf{r}}}(\mathbf{x}) \right) = \exp\left(|\mathbf{r}| h_{\hat{\mathbf{r}}}(\mathbf{x}) \right),$$

where

$$h_{\hat{\mathbf{r}}}(\mathbf{x}) = \Re H_{\hat{\mathbf{r}}} = -\sum_{j=1}^{d} \hat{r}_j \ln |x_j|. \tag{2.2}$$

When no confusion exists, we shall simply refer to the function $h_{\hat{\mathbf{r}}}$ as h. The geometry of the height function h will play an important role in our analysis.

As $|\mathbf{r}| \to \infty$, the above equations show that the magnitude of the integrand grows at an exponentially slower rate along points farther away from the origin (where the height function h is smaller). This motivates pushing the domain of integration out towards infinity, reducing the growth rate of the integrand on the domain over which it is integrated. Of course if F has poles they will present an obstruction, but we can still try push the domain of integration around these poles. In the end we obtain an integral over two domains: one near the pole set of F (obtained by pushing the original domain around the poles), and one past the pole set of F (far away from the origin). This idea is formalized in the theorem below.

THEOREM 2.2. *Let $F = P/Q$, with $P, Q \colon \mathbb{C}^d \to \mathbb{C}$ entire, where the vanishing set $\mathcal{V}_Q$ of Q is smooth. Let T_0 be a torus as in Cauchy's Integral Formula. Let $T_1 \subseteq \mathbb{C}^d$ be a torus homotopic to T_0 under a homotopy*

$$K \colon T \times [0,1] \to \mathbb{C}^d, \qquad \text{with } T_0 = T \times \{0\},\ T_1 = T \times \{1\},$$

passing through $\mathcal{V}_Q$ transversally. Identifying K with its image in $\mathbb{C}^d$, assume further that K does not intersect the coordinate axes, and that $\partial K \cap \mathcal{V}_Q = \varnothing$. Define

$$C = K \cap \mathcal{V}_Q.$$

Then for any tubular neighborhood ν of C in K, we have

$$a_{\mathbf{r}} = \int_{T_0} \omega_F = \int_{\partial \nu} \omega_F + \int_{T_1} \omega_F,$$

given the proper orientation of $\partial \nu$.

Note: when we say $\mathcal{V}_Q$ is smooth we mean that $\mathcal{V}_Q$ has the structure of a smooth manifold (see [**Bre93**, p. 68]). And when we say that K passes through $\mathcal{V}_Q$ transversally we mean that the image of K intersects with $\mathcal{V}_Q$ transversally as (real) submanifolds of $\mathbb{C}^d$ (see [**Bre93**, p. 84]).

PROOF. Counting (real) dimensions, $\dim \mathcal{V}_Q = 2d - 2$ and $\dim K = d + 1$. Hence their transverse intersection C is a $d - 1$ real-dimensional subspace of K.

Now take any tubular neighborhood ν of C in K. As ν is a full-dimensional submanifold of the orientable manifold K, ν is orientable and hence its boundary $\partial \nu$ is orientable too. Given the proper orientation of $\partial \nu$, we have that

$$\partial(K \setminus \nu) = T_1 - T_0 + \partial \nu.$$

Note that ω_F is holomorphic on $K \setminus \nu$. By Stokes' Theorem [**Bre93**, p. 267] and the fact that ω_F is an exact form we get

$$\int_{T_1 - T_0 + \partial \nu} \omega_F = \int_{K \setminus \nu} \mathrm{d}\omega_F = \int_{K \setminus \nu} 0 = 0,$$

leading to the equality of the theorem. □

When T_1 is far enough away from the origin, $\int_{T_1} \omega_F$ is negligible (possibly even 0), and so the asymptotic analysis of the coefficients $a_{\mathbf{r}}$ reduces to an integral near the pole set of F. In the next section, we reduce this further to an integral *on* the pole set of F.

2.2. The Residue Theorem. In this subsection we present a theory generalizing the theory of residues of the complex analysis of one variable. The theory was developed by Jean Leray in 1959, and more details regarding the construction can be found in [**AY83**, § 16]. The main result we obtain is Theorem 2.8 below, an analogue of the Cauchy Residue Theorem in one variable. Its application to coefficient analysis is found in Corollary 2.9.

We restrict our attention to a limited part of Leray's theory, focusing on meromorphic d-forms in $\mathbb{C}^d$.

Definition 2.3. Let η be a meromorphic d-form, represented as

$$\eta = \frac{P}{Q}\,\mathrm{d}\mathbf{x} \qquad \text{on a domain} \quad U \subseteq \mathbb{C}^d,$$

where P and Q are holomorphic on U. Denote by $\mathcal{V}_Q$ the zero set of Q on U, and assume that η has a simple pole everywhere on $\mathcal{V}_Q$. Denote by $\iota\colon \mathcal{V}_Q \to U$ the inclusion map. Then we define the *residue of* η *on* $\mathcal{V}_Q$ by

$$\mathrm{Res}(\eta) = \iota^*\theta,$$

where ι^* denotes pullback by ι (see [**Bre93**, p. 263]), and where θ is any solution to

$$\mathrm{d}Q \wedge \theta = P\,\mathrm{d}\mathbf{x}.$$

Before delving into the existence and uniqueness of the residue, we do a few example computations.

Example 2.4. For $\eta = P/Q\,\mathrm{d}\mathbf{x}$ as above, wherever $Q_i = \frac{\partial Q}{\partial x_i}$ does not vanish we have the representation

$$\mathrm{Res}(\eta) = (-1)^{i-1}\frac{P}{Q_i}\,\mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_{i-1} \wedge \mathrm{d}x_{i+1} \wedge \ldots \wedge \mathrm{d}x_d.$$

As a special case, note that for $Q = x_1$ we obtain

$$\mathrm{Res}(\eta) = P(0, x_2, \ldots, x_d)\,\mathrm{d}x_2 \wedge \ldots \wedge \mathrm{d}x_d.$$

In the case where $d = 1$, this reduces to $\mathrm{Res}(P(x)/x) = P(0)$, which is precisely the ordinary residue of $P(x)/x$ at $x = 0$. This motivates the above definition as a genuine extension of the single variable residue.

Example 2.5. As the most pertinent case of the Example 2.4, we examine $\mathrm{Res}(\omega_F)$ where $F = P/Q$ is meromorphic. Away from the coordinate axes, ω_F can be written as

$$\omega_F = \frac{\frac{1}{(2\pi\mathrm{i})^d} \cdot \frac{P(\mathbf{x})}{x_1 \cdot \ldots \cdot x_d}\exp(|\mathbf{r}|H(\mathbf{x}))}{Q(\mathbf{x})}\,\mathrm{d}\mathbf{x},$$

where the numerator and denominator are holomorphic functions. So wherever Q_d and the x_j do not vanish (for all j), we have

$$\mathrm{Res}(\omega_F) = \frac{(-1)^{d-1}}{(2\pi\mathrm{i})^d} \cdot \frac{P(\mathbf{x})}{x_1 \cdot \ldots \cdot x_d\, Q_d(\mathbf{x})} e^{|\mathbf{r}|H(\mathbf{x})}\,\mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_{d-1}.$$

We now show existence and uniqueness of the residue form along the simple pole set $\mathcal{V}_Q$.

PROPOSITION 2.6. *Let η be as in Definition 2.3. Then for any point $\mathbf{p} \in \mathcal{V}_Q$, there is a neighborhood $V \subseteq U$ of $\mathbf{p}$ and a holomorphic $(d-1)$-form θ on V solving the equation*

$$\mathrm{d}Q \wedge \theta = P \,\mathrm{d}\mathbf{x}. \tag{2.3}$$

Furthermore, the restriction $\iota^*\theta$ induced by the inclusion $\iota\colon \mathcal{V}_Q \cap V \to V$ is unique.

PROOF. First, we prove the existence of a solution θ to (2.3) in a neighborhood of $\mathbf{p}$. As Q has a simple zero at $\mathbf{p}$, the implicit function theorem implies that for some neighborhood V of $\mathbf{p}$ there is a biholomorphic function $\psi\colon \mathbb{C}^d \to V$ such that $Q(\psi(\mathbf{x})) = x_1$. Define the form θ_0 by

$$\theta_0 = (P \circ \psi)\,|J|\,\mathrm{d}x_2 \wedge \ldots \wedge \mathrm{d}x_d,$$

where J is the Jacobian of the function ψ. The claim is that $\theta = (\psi^{-1})^*\theta_0$ is a solution to (2.3).

Indeed, by definition of θ_0 we have that $\mathrm{d}x_1 \wedge \theta_0 = (P \circ \psi)\,|J|\,\mathrm{d}\mathbf{x}$. Pulling back both sides of this equation by ψ^{-1} yields

$$\mathrm{d}(\psi^{-1}(\mathbf{x})_1) \wedge (\psi^{-1})^*\theta_0 = P \cdot (\psi^{-1})^*(|J|\,\mathrm{d}\mathbf{x}),$$

which simplifies to $\mathrm{d}Q \wedge \theta = P\,\mathrm{d}\mathbf{x}$, as desired.

To prove uniqueness, assume that we have two $(d-1)$-forms θ and $\tilde{\theta}$ such that $\mathrm{d}Q \wedge \theta = P\,\mathrm{d}\mathbf{x}$ and $\mathrm{d}Q \wedge \tilde{\theta} = P\,\mathrm{d}\mathbf{x}$. Then $\mathrm{d}Q \wedge (\theta - \tilde{\theta}) = 0$, which implies

$$\psi^*(\mathrm{d}Q \wedge (\theta - \tilde{\theta})) = \mathrm{d}x_1 \wedge \psi^*(\theta - \tilde{\theta}) = 0.$$

But this means that $\psi^*(\theta - \tilde{\theta})$ is a multiple of $\mathrm{d}x_1$. Pulling back by $(\psi^{-1})^*$, this implies that $\theta - \tilde{\theta}$ is a multiple of $\mathrm{d}Q$. Finally, pulling back by ι^*, this implies that $\iota^*(\theta - \tilde{\theta})$ is a multiple of $\mathrm{d}(Q \circ \iota) = 0$. Thus $\iota^*(\theta - \tilde{\theta})$ vanishes, and so $\iota^*\theta = \iota^*\tilde{\theta}$. □

REMARK 2.7. Let η be as in the definition of the residue form, and let $\psi\colon V \to U$ be a biholomorphic function. Then

(1) The residue form is natural, i.e., $\mathrm{Res}(\eta)$ does not depend on the particular P and Q chosen to represent η as $(P/Q)\,\mathrm{d}\mathbf{x}$.
(2) The residue form is functorial, i.e., $\mathrm{Res}(\psi^*\eta) = \psi^* \mathrm{Res}(\eta)$ (where on the right side of the equation, ψ is restricted to the domain $\psi^{-1}(\mathcal{V}_Q) = \mathcal{V}_{Q\circ\psi}$).

THEOREM 2.8 (Cauchy–Leray Residue Theorem). *Let η be a meromorphic d-form on domain $U \subseteq \mathbb{C}^d$, with pole set $\mathcal{V} \subseteq U$ along which η has only simple poles. Let N be a d-chain in U, locally the product of a $(d-1)$-chain C on $\mathcal{V}$ with a circle γ in the normal slice to $\mathcal{V}$, oriented positively with respect to the complex structure of the normal slice. Then*

$$\int_N \eta = 2\pi\mathrm{i} \int_C \mathrm{Res}(\eta).$$

PROOF. We proceed by examining the structure of the integral locally. So fix an arbitrary $\mathbf{p} \in C$. In a neighborhood $V \subseteq \mathbb{C}^d$ of $\mathbf{p}$, the surrounding space looks

like a direct product of $\mathcal{V} \cap V$ (isomorphic to $\mathbb{C}^{d-1}$ for V small) and the normal space to $\mathcal{V} \cap V$ (isomorphic to $\mathbb{C}$). Hence there is a biholomorphic function

$$\begin{aligned} \varphi \colon V &\to \mathbb{C} \times \mathbb{C}^{d-1} \\ \mathbf{x} &\mapsto (\varphi_1(\mathbf{x}), \varphi_2(\mathbf{x})), \end{aligned}$$

where the map φ_2^{-1} is a parametrization of $\mathcal{V} \cap V$, and

$$\begin{aligned} \varphi(\mathcal{V} \cap V) &= \{0\} \times \varphi_2(\mathcal{V} \cap V), \\ \varphi(N \cap V) &= \gamma \times \varphi_2(C \cap V), \end{aligned}$$

where $\gamma \subseteq \mathbb{C}$ is a loop around the origin, positively oriented. Furthermore, if V is chosen small enough, we can guarantee that the meromorphic form $(\varphi^{-1})^*\eta$ has a *global* representation as $P/Q \, \mathrm{d}\mathbf{x}$. Note that, by the structure of η and definition of φ, Q must vanish on the set

$$\varphi(\mathcal{V} \cap V) = \{\mathbf{x} \in \mathbb{C}^d : x_1 = 0\},$$

where it has only simple zeroes.

We claim that if we can prove the equality stated in the residue theorem restricted to V, we shall be done with the theorem. This is due to the additivity of integration and the compactness of C: we can split up a tubular neighborhood of C (containing N) into finitely many such neighborhoods on which the theorem holds, then prove the theorem by breaking the integral into a sum over these pieces.

So without loss of generality, we may assume that this local structure holds globally on C and that the domain of the map φ is all of $\mathbb{C}^d$. By changing variables, we get

$$\int_N \eta = \int_{\gamma \times \varphi_2(C)} \frac{P}{Q} \, \mathrm{d}\mathbf{x} = \int_{\mathbf{p} \in \varphi_2(C)} \left(\int_{\gamma \times \{\mathbf{p}\}} \frac{P}{Q} \, \mathrm{d}x_1 \right) \mathrm{d}x_2 \wedge \ldots \wedge \mathrm{d}x_d. \tag{2.4}$$

the upshot being the ability to split the above into an iterated integral, by the product structure of $\gamma \times \varphi_2(C)$.

The next step is to compute the inner integral from (2.4) by the ordinary residue theorem, but doing so will require a change of variables. To that end, define the function $\psi \colon \mathbb{C}^d \to \mathbb{C}^d$ by

$$\psi(\mathbf{x}) = (Q(\mathbf{x}), x_2, x_3, \ldots, x_d),$$

and fix some $\mathbf{p} \in \mathbb{C}^{d-1}$. The claim is that ψ is biholomorphic in a neighborhood $W \subseteq \mathbb{C}^d$ of $(0, \mathbf{p})$. By the inverse function theorem, this is true if and only if $|J(\mathbf{p})| = Q_1(\mathbf{p}) \neq 0$, where J is the Jacobian of ψ. As Q has a simple zero at $\mathbf{p}$, it cannot be true that $Q_i(\mathbf{p}) = 0$ for all i. But $Q_i(\mathbf{p}) = 0$ for all $i \neq 1$, because Q is constant (equal to 0) on the entire plane $x_1 = 0$. Thus $Q_1(\mathbf{p}) \neq 0$, as desired. Note that ψ^{-1} must have the form

$$\psi^{-1}(\mathbf{x}) = (f(\mathbf{x}), x_2, x_3, \ldots, x_d)$$

for some function f, and that $Q \circ \psi^{-1} = x_1$.

We should like to perform a change of variables and compute the inner integral from (2.4) over the domain $\psi(\gamma \times \{\mathbf{p}\})$. The only problem with this is that there is no guarantee that $\gamma \times \varphi_2(C) \subseteq W$. But we *can* make this guarantee by shrinking N, i.e., shrinking the loop γ closer to the origin, and by (potentially) restricting our attention to a small portion of C. Note that shrinking N has no effect on the original integral (the new N will differ from the old N by a boundary, and we are

integrating a closed form), and that, as we have already stated, we need only prove the residue theorem locally. Thus we may assume without loss of generality that $\gamma \times \varphi_2(C)$ is contained entirely within the domain of ψ.

After the suggested change of variables, we obtain

$$\int_N \eta = \int_{\mathbf{p}\in\varphi_2(C)} \left(\int_{\psi(\gamma\times\{\mathbf{p}\})} \frac{P\circ\psi^{-1}}{x_1} \frac{\partial f}{\partial x_1}\,\mathrm{d}x_1 \right) \mathrm{d}x_2 \wedge \ldots \wedge \mathrm{d}x_n.$$

By the form of ψ, $\psi(\gamma \times \{\mathbf{p}\})$ is simply a loop around the origin in the plane $\{x \in \mathbb{C}^d : (x_2, \ldots, x_d) = \mathbf{p}\}$. So by the ordinary residue theorem we can compute

$$\int_{\psi(\gamma\times\{\mathbf{p}\})} \frac{P\circ\psi^{-1}}{x_1} \frac{\partial f}{\partial x_1}\,\mathrm{d}x_1 = 2\pi\mathrm{i}\cdot P(\psi^{-1}(0,\mathbf{p}))\frac{\partial f}{\partial x_1}(0,\mathbf{p}).$$

Substituting back into (2.4) yields

$$\begin{aligned}\int_N \eta &= 2\pi\mathrm{i}\int_{\mathbf{p}\in\varphi_2(C)} P(\psi^{-1}(0,\mathbf{p}))\frac{\partial f}{\partial x_1}(0,\mathbf{p})\,\mathrm{d}x_2\wedge\ldots\wedge\mathrm{d}x_n \\ &= 2\pi\mathrm{i}\int_{\{0\}\times\varphi_2(C)} \operatorname{Res}\left(\frac{P\circ\psi^{-1}\cdot\frac{\partial f}{\partial x_1}}{x_1}\,\mathrm{d}\mathbf{x}\right),\end{aligned}$$

where the second equality comes from the residue computation of Example 2.4.

But note that

$$\begin{aligned}(\psi^{-1})^*\left(\frac{P}{Q}\,\mathrm{d}\mathbf{x}\right) &= \frac{P\circ\psi^{-1}}{x_1}\left(\sum_{j=1}^{d}\frac{\partial f}{\partial x_j}\,\mathrm{d}x_j\right)\wedge\mathrm{d}x_2\wedge\ldots\wedge\mathrm{d}x_d \\ &= \frac{P\circ\psi^{-1}}{x_1}\frac{\partial f}{\partial x_1}\,\mathrm{d}\mathbf{x},\end{aligned}$$

and so the integral equation becomes

$$\begin{aligned}\int_N \eta &= 2\pi\mathrm{i}\int_{\{0\}\times\varphi_2(C)} \operatorname{Res}\left((\psi^{-1})^*\left(\frac{P}{Q}\,\mathrm{d}\mathbf{x}\right)\right) \\ &= 2\pi\mathrm{i}\int_{\{0\}\times\varphi_2(C)} \operatorname{Res}\big((\psi^{-1})^*(\varphi^{-1})^*\eta\big).\end{aligned}$$

Finally, by the functoriality of the residue form, we obtain

$$\int_N \eta = 2\pi\mathrm{i}\int_{\{0\}\times\varphi_2(C)} (\psi^{-1})^*(\varphi^{-1})^*\operatorname{Res}(\eta) = 2\pi\mathrm{i}\int_C \operatorname{Res}(\eta).$$

□

The residue theorem applies directly to the coefficient analysis of the previous subsection by the following corollary.

COROLLARY 2.9. *Under the assumptions and notation of Theorem* 2.2,

$$a_{\mathbf{r}} = 2\pi\mathrm{i}\int_C \operatorname{Res}(\omega_F) + \int_{T_1}\omega_F,$$

given the proper orientation of C.

PROOF. By the residue theorem, $\int_{\partial\nu}\omega_F = 2\pi\mathrm{i}\int_C \operatorname{Res}(\omega_F)$. The result follows by substituting this equality into the conclusion of Theorem 2.2. □

And thus the asymptotic coefficient analysis reduces to the integration of a $(d-1)$-form along a cycle on the pole set of the coefficient generating function. The final step is to compute this integral by means of the saddle point method.

2.3. Critical points of the height function. The goal is to obtain an asymptotic expansion for $2\pi\mathrm{i}\int_C \mathrm{Res}(\omega_F)$, where $F = P/Q$ for some entire functions P and Q, F is analytic in a neighborhood of the origin, and $\mathcal{V}_Q$ is smooth. By Example 2.5 we can expect $\mathrm{Res}(\omega_F)$ to take the form

$$\mathrm{Res}(\omega_F) = \frac{(-1)^{d-1}}{(2\pi\mathrm{i})^d} \cdot \frac{P(\mathbf{x})}{x_1 \cdot \ldots \cdot x_d\, Q_d(\mathbf{x})} e^{|\mathbf{r}|H(\mathbf{x})} \mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_{d-1}$$

(where Q_d does not vanish), and as before we see that the exponential growth of this form is governed by the height function h. This motivates a deformation of the cycle C along $\mathcal{V}_Q$, pushing C down to a homologous cycle $\tilde{C}$ on which the maximum modulus of h is minimized. This procedure is obstructed when the cycle gets trapped on a saddle point of h on $\mathcal{V}_Q$, and the idea is to arrange $\tilde{C}$ so that the local maxima of h along $\tilde{C}$ are *all* achieved at such saddle points. Away from the highest saddle points (the *contributing points*), the integral will contribute asymptotically negligible quantities, and near the contributing points the integral will be amenable to the saddle point method.

Thus the first task is to identify the location of the critical points of $h_{\hat{\mathbf{r}}}|_{\mathcal{V}_Q}$. These points can be realized as the zero set of d equations, as exhibited below.

THEOREM 2.10 (Location of Critical Points). *Assume $\hat{r}_d \neq 0$. Then the critical points of $h_{\hat{\mathbf{r}}}$ restricted to $\mathcal{V}_Q$ are precisely those points $\mathbf{p} \in \mathbb{C}^d$ satisfying the following d equations:*

$$Q(\mathbf{p}) = 0,$$
$$\hat{r}_d p_j Q_j(\mathbf{p}) - \hat{r}_j p_d Q_d(\mathbf{p}) = 0 \quad \forall j \neq d.$$

In the case $d = 2$, these critical points are actually saddle points of $h_{\hat{\mathbf{r}}}|_{\mathcal{V}_Q}$.

For the purposes of computation it should be noted that when Q is a polynomial, the above set of critical points is generically finite and can be found algorithmically by the method of Gröbner bases (see [**CLO05**, § 1.3]).

PROOF. The equation $Q(\mathbf{p}) = 0$ is clear: any critical point of $h|_{\mathcal{V}_Q}$ will have to be on $\mathcal{V}_Q$. So we turn to the remaining $d-1$ equations.

Fix a point $\mathbf{p} \in \mathcal{V}_Q$ (not on the coordinate axes). By the Cauchy–Riemann equations, $\mathbf{p}$ is a critical point of $\Re\left(H|_{\mathcal{V}_Q}\right)$ if and only if it is a critical point of $\Im\left(H|_{\mathcal{V}_Q}\right)$. Thus $\mathbf{p}$ is a critical point of $h|_{\mathcal{V}_Q}$ exactly when

$$\nabla(H|_{\mathcal{V}_Q})(\mathbf{p}) = 0.$$

But $\nabla(H|_{\mathcal{V}_Q})(\mathbf{p})$ is simply the projection of $\nabla H(\mathbf{p})$ onto the tangent space $T_{\mathbf{p}}\mathcal{V}_Q$. Hence the previous equation is true if and only if

$$\nabla H(\mathbf{p}) \parallel \nabla Q(\mathbf{p}),$$

as $\nabla Q(\mathbf{p})$ is a vector normal to the tangent space to $\mathcal{V}_Q$ at $\mathbf{p}$. This condition reduces to the equation

$$\left(\frac{-\hat{r}_1}{p_1}, \ldots, \frac{-\hat{r}_d}{p_d}\right) = \lambda\left(Q_1(\mathbf{p}), \ldots, Q_d(\mathbf{p})\right)$$

for some scalar λ, which is captured by the remaining $d-1$ equations of the theorem.

For the $d = 2$ case, let $\mathbf{p}$ be any critical point of $h|_{\mathcal{V}_Q}$ (hence a critical point of $H|_{\mathcal{V}_Q}$ by the above). In a chart map in a neighborhood of the origin, we can write

$$H|_{\mathcal{V}_Q}(z) = c_0 + c_k z^k (1 + O(z)),$$

for some constants c_0 and c_k and $k \geqslant 2$. As $h = \Re(H)$, it follows that $h|_{\mathcal{V}_Q}$ has a k^{th} order saddle at $\mathbf{p}$. □

After deforming the domain of integration so that h is locally maximized at the critical points located above, the final step is to obtain an asymptotic expansion by applying the saddle point method near these points. In the case where $d = 2$, this results in a single variable saddle integral. Specifically, we shall make use of the following theorem.

THEOREM 2.11. *Let A and ϕ be holomorphic functions on a neighborhood of $0 \in \mathbb{C}$, with*

$$A(z) = \sum_{j=l}^{\infty} b_j z^j, \qquad \phi(z) = \sum_{j=k}^{\infty} c_j z^j,$$

where $l \geqslant 0$, $k \geqslant 2$ and $b_l \neq 0$, $c_j \neq 0$. Let $\gamma\colon [-\varepsilon, \varepsilon] \to \mathbb{C}$ be any smooth curve with $\gamma(0) = 0$, $\gamma'(0) \neq 0$, and assume that $\Re\phi(\gamma(t)) \geqslant 0$ with equality only at $t = 0$. Denote by γ^+ the image of γ restricted to the domain $[0, \varepsilon]$. Then for some coefficients a_j we have a full asymptotic expansion

$$\int_{\gamma^+} A(z) e^{-\lambda\phi(z)}\, \mathrm{d}z = \sum_{j=l}^{\infty} \frac{a_j}{k} \Gamma\left(\frac{1+j}{k}\right) (c_k\lambda)^{-(1+j)/k}$$

as $\lambda \to \infty$, where the choice of k^{th} root in $(c_k\lambda)^{-(1+j)/k}$ is made by taking the principal root of $v^{-1}(c_k\lambda v^k)^{1/k}$ where $v = \gamma'(0)$. The leading two coefficients a_j are given by

$$a_l = b_l, \qquad a_{l+1} = b_{l+1} - \frac{2+l}{k} \cdot \frac{c_{k+1}}{c_k}.$$

For the purposes of computation it should be noted that each coefficient a_j can be effectively computed from the values $b_l, \ldots, b_j$ and $c_k, \ldots, c_{k+j-l}$.

See [**Pem09**] for the proof, or [**Hen91**, § 11.8] for a treatment from which the above may be derived. It should be noted that, while the saddle point method is a very well known and well understood technique, it is often presented only as a method for solving a general class of problems; theorems are usually only given for limited, special case applications. Theorem 2.11 is stated in a generality not easily found in the literature.

3. Application to bicolored supertrees

The purpose of this section is to apply the methods of Section 2 to the example presented in the introduction. Throughout this section we use the notation of

Section 2, and in the case of bicolored supertrees this means

$$\begin{aligned} F &= \frac{P}{Q},\ P \text{ and } Q \text{ defined as in (1.1)},\\ |\mathbf{r}| &= n,\ \hat{\mathbf{r}} = \hat{\mathbf{r}}_0 = (1,1),\\ H(x,y) &= -\ln x - \ln y,\\ h(x,y) &= -\ln|x| - \ln|y|. \end{aligned}$$

Then as outlined in Section 2, the procedure will be as follows.

(1) Reduce the asymptotic computation to an integral on the variety $\mathcal{V}_Q$ using Corollary 2.9.
(2) Locate the critical points of $h|_{\mathcal{V}_Q}$ and deform the contour of integration so as to minimize the maximum of h at such points.
(3) Compute an asymptotic expansion for this integral by applying Theorem 2.11 near these maxima, and bounding the order away from these maxima.

These three steps will be carried out in the subsections that follow. Thanks to all the work laid out in the previous section, many of these steps will be automatic. The most difficult step will be step (2), finding the new saddle point contour and proving that it possesses the right properties (Lemma 3.6). The rest will be a matter of applying the theorems, when appropriate.

Before jumping into computations, however, we shall need to do some initial work on describing the variety $\mathcal{V}_Q$.

3.1. Describing the variety. Because Q is quadratic in the variable y, we can explicitly solve $Q = 0$ for y as a function of x. This will allow us to parametrize $\mathcal{V}_Q$ by x where possible. So, define

$$\begin{aligned} y_1(x) &= \frac{-x^2 + x^3 - 2 + \sqrt{x^4 + 4x^2 - 4x^3 + 4}}{x^5},\\ y_2(x) &= \frac{-x^2 + x^3 - 2 - \sqrt{x^4 + 4x^2 - 4x^3 + 4}}{x^5}, \end{aligned}$$

where in each case the principal root is chosen. Then by the quadratic formula,

$$\mathcal{V}_Q = \{(x, y_j(x)) : x \in \mathbb{C} \setminus \{0\},\ j = 1,2\} \cup \{(0, 1/2)\}$$

(though note that we may write $(0,1/2) = (0, y_1(0))$ by analytically continuing y_1 to $x = 0$). To parametrize $\mathcal{V}_Q$ by x, we define the parametrization functions

$$\iota_1(x) = (x, y_1(x)), \qquad \iota_2(x) = (x, y_2(x)).$$

For the purposes of later computation, it will be nice to know a domain on which these parametrization functions are holomorphic.

LEMMA 3.1. *ι_1 and ι_2 are holomorphic on the punctured strip*

$$\left\{x \in \mathbb{C} \setminus 0 : \Im\left(1 \pm \sqrt{1 - 2\mathrm{i}}\right) < \Im x < \Im\left(1 \pm \sqrt{1 + 2\mathrm{i}}\right)\right\}.$$

PROOF. By definition of the functions y_1 and y_2, the only points where ι_1 and ι_2 may fail to be holomorphic have $x = 0$ or $f(x) = x^4 + 4x^2 - 4x^3 + 4 \leqslant 0$ (by the choice of principal square root). Thus we examine when $f(x)$ is a nonpositive real number.

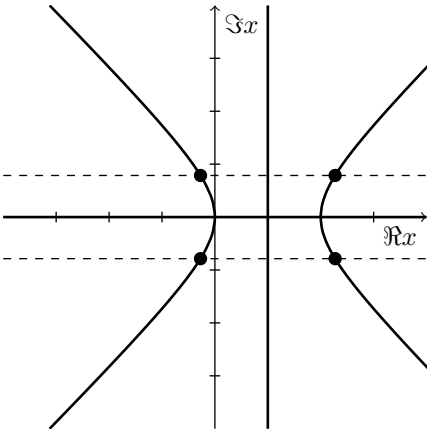


FIGURE 1. Zero sets of $\Im f$ and f.

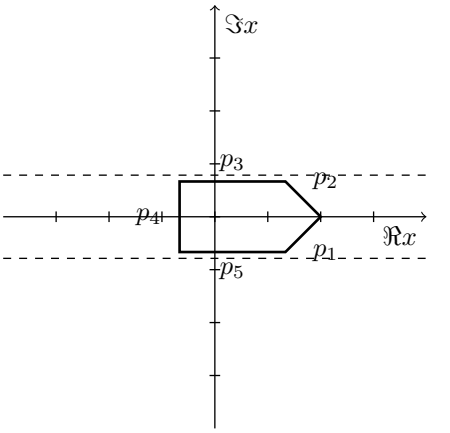


FIGURE 2. The pentagonal path p.

Let $a = \Re(x)$ and $b = \Im(x)$. We are interested in when $f(a + \mathrm{i}b) \leqslant 0$, so we first examine the equation $\Im f(a + \mathrm{i}b) = 0$, or

$$4b(a-1)(a^2 - 2a - b^2) = 0.$$

The solution set of the above equation is the union of the lines $a = 1$, $b = 0$ and the hyperbola $a^2 - 2a - b^2 = 0$. The points $x = 1 \pm \sqrt{1 \pm 2\mathrm{i}}$ where $f(x) = 0$ partition the set $\Im f(x) = 0$ into five components on which $\Re f(x)$ is either all positive or all negative (by continuity of f on the connected set $\Im f(x) = 0$). See Figure 1 (the dashed lines denoting the boundary of the strip).

Only one of these five components enters the strip $\Im\left(1 \pm \sqrt{1 - 2\mathrm{i}}\right) < \Im x < \Im\left(1 \pm \sqrt{1 + 2\mathrm{i}}\right)$, and by plugging a sample point of this component into f (say $f(0) = 4$) we see that f is positive on this component. Thus at no point within this strip is $f(x) \leqslant 0$, and so ι_1 and ι_2 are holomorphic on the claimed domain. □

Finally, as evidenced by Example 2.5, it will be useful for representing the residue form to know where $Q_y = \frac{\partial Q}{\partial y}$ is nonzero along $\mathcal{V}_Q$. Computing a Gröbner basis of the ideal $\langle Q, Q_y \rangle$ in MAPLE [**Wat08**] via the command

```
Basis([Q,diff(Q,y)],plex(y,x));
```

we obtain the univariate polynomial $x^4 + 4x^2 - 4x^3 + 4$ as the first basis element. Hence the x coordinate of any point where Q and Q_y simultaneously vanish must be a root of this polynomial. This justifies the following remark.

REMARK 3.2. Along $\mathcal{V}_Q$, Q_y is nonzero whenever $x \neq 1 \pm \sqrt{1 \pm 2\mathrm{i}}$ (the roots of the equation $x^4 + 4x^2 - 4x^3 + 4 = 0$).

3.2. Integral on the pole variety. The following lemma accounts for the first step of the analysis: using Corollary 2.9 to reduce the computation of $a_{n,n}$ to an integral on $\mathcal{V}_Q$.

LEMMA 3.3. *For $\varepsilon > 0$, define*

$$C_\varepsilon = \{x \in \mathbb{C} : |x| = \varepsilon\},$$

the circle of radius ε about $0 \in \mathbb{C}$, oriented counterclockwise. Then for sufficiently small $\varepsilon > 0$,

$$a_{n,n} = 2\pi \mathrm{i} \int_{\iota_1(C_\varepsilon)} \mathrm{Res}(\omega_F) + 2\pi \mathrm{i} \int_{\iota_2(C_\varepsilon)} \mathrm{Res}(\omega_F). \tag{3.1}$$

PROOF. We first verify that the variety $\mathcal{V}_Q$ is smooth. This is true only if Q, Q_x and Q_y do not simultaneously vanish, which is true if and only if the variety $I = \langle Q, Q_x, Q_y \rangle$ is trivial (the whole polynomial ring). We check this algorithmically, using Gröbner bases. In MAPLE, we compute the Gröbner basis of I with the command

```
Basis([Q,diff(Q,x),diff(Q,y)],plex(y,x));
```

MAPLE returns the basis `[1]` for I, so the ideal is indeed trivial.

Now, let $\varepsilon > 0$, $\delta > 0$ be sufficiently small so that

$$a_{n,n} = \int_{T_0} \omega_F, \qquad \text{where} \quad T_0 = \{(x,y) \in \mathbb{C}^2 : |x| = \varepsilon, |y| = \delta\},$$

by Cauchy's Integral Formula. Define the quantities

$$m_0 = \inf\{|y_j(x)| : x \in C_\varepsilon,\ j = 1,2\},$$
$$M_0 = \sup\{|y_j(x)| : x \in C_\varepsilon,\ j = 1,2\}.$$

For ε sufficiently small, note that $M_0 < \infty$ (by continuity of the y_j; see Lemma 3.1) and $m_0 > 0$ (the x-axis intersects $\mathcal{V}_Q$ only at the point $(2,0)$).

Assume δ is chosen small enough so that $\delta < m_0$. Fix any $M > M_0$. Then define the homotopy

$$\begin{aligned} K : T_0 \times [0,1] &\to \mathbb{C}^2 \\ (x,y,t) &\mapsto \left(x, y\left(1 + t\left(\tfrac{M}{\delta} - 1\right)\right)\right), \end{aligned}$$

expanding T_0 in the y direction past $\mathcal{V}_Q$. Then K intersects $\mathcal{V}_Q$ in the set $C = \iota_1(C_\varepsilon) \cup \iota_2(C_\varepsilon)$ and avoids the coordinate axes. Furthermore, K intersects $\mathcal{V}_Q$ transversally (as K expands in the y direction, intersecting $\mathcal{V}_Q$ where it is a graph of x). Thus, by Corollary 2.9 we obtain

$$a_{n,n} = 2\pi \mathrm{i} \int_{\iota_1(C_\varepsilon)} \mathrm{Res}(\omega_F) + 2\pi \mathrm{i} \int_{\iota_2(C_\varepsilon)} \mathrm{Res}(\omega_F) + \int_{T_1} \omega_F, \tag{3.2}$$

where C_ε is oriented counterclockwise (determined by examination of Theorem 2.2 and the Residue Theorem).

Now fix n large and let M vary. As the rest of the terms in (3.2) have no M dependence, $\int_{T_1} \omega_F$ must be a constant function of M. But by trivial bounds, we can show that

$$\int_{T_1} \omega_F = O(M^{1-n}) \qquad \text{as} \quad M \to \infty,$$

as $\frac{P}{(2\pi \mathrm{i})^2 xyQ} = O(1)$, $\exp(nH) = O(M^{-n})$ and the area of T_1 is $O(M)$. For $n > 1$, $M^{1-n} \to 0$ as $M \to \infty$. Hence the only constant $\int_{T_1} \omega_F$ can be equal to is 0. $\square$

3.3. Saddle points and contour deformation. Step (2) in the analysis is to locate the saddle points of $h|_{\mathcal{V}_Q}$ and deform the contour of integration appropriately, using this information. The saddle points can be found automatically, as follows.

LEMMA 3.4. *$h|_{\mathcal{V}_Q}$ has three saddle points, located at*

$$\begin{aligned}\left(2, \tfrac{1}{8}\right) &= \iota_1(2),\\ \left(1-\sqrt{5}, \tfrac{3+\sqrt{5}}{16}\right) &= \iota_1(1-\sqrt{5}),\\ \left(1+\sqrt{5}, \tfrac{3-\sqrt{5}}{16}\right) &= \iota_2(1+\sqrt{5}).\end{aligned}$$

PROOF. By Theorem 2.10, the critical points of $h|_{\mathcal{V}_Q}$ are those points where Q and $xQ_x - yQ_y$ simultaneously vanish. We can compute these points algorithmically by computing the Gröbner basis for the ideal $I = \langle Q, xQ_x - yQ_y\rangle$. This is done in MAPLE with the command

```
Basis([Q,x*diff(Q,x)-y*diff(Q,y)],plex(y,x));
```

which returns a basis consisting of the following two polynomials:

$$32 - 8x^2 - 32x + 20x^3 - 8x^4 + x^5, \qquad x^4 - 48 - 6x^3 + 8x^2 + 128y + 16.$$

The first polynomial factors as $(x^2-2x-4)(x-2)^3$, with roots $x = 2$ and $x = 1\pm\sqrt{5}$. Substituting these values of x into the second polynomial and solving for y yields the critical points claimed in the lemma. □

We note here the interesting geometry near the critical point $(2, 1/8)$, which will turn out to be the sole contributing point. Expanding $H(\iota_1(x))$ near $x = 2$, we obtain

$$H(\iota_1(x)) = H(\iota_1(2)) + \frac{1}{16}(x-2)^4 + O\left((x-2)^6\right),$$

and hence $h|_{\mathcal{V}_Q}$ has a degenerate saddle (of order 4) near this critical point, with steepest descent directions emanating from $x = 2$ at angles $\pi/4 + j(\pi/2)$ radians $(j = 1, 2, 3, 4)$. We also see that along the path $|x| = 2$, $h(\iota_1(x))$ is locally *minimized* at $x = 2$, as this path passes through the critical point along ascent directions. Hence $x = 2$ is a local maximum for $|y_1(x)|$ along this path, and so there are points $(x, y) \in \mathcal{V}_Q$ near $(2, 1/8)$ such that $|x| = 2$ and $|y| < 1/8$. Because $\mathcal{V}_Q$ cuts in toward the origin near $\iota_1(2)$, this critical point is not on the boundary of the domain of convergence of F. In the terminology of the introduction, this critical point is *not* minimal.

Knowing where the saddle points of h are, the next task for us is to deform the contour of integration in (3.3) so as to minimize the maximum modulus of h along the new contour at said saddle points. The integral over domain $\iota_2(C_\varepsilon)$ will be shown to vanish, while the domain $\iota_1(C_\varepsilon)$ will be pushed to a "pentagonal" path through the critical point $(2, 1/8)$.

The specific path to which $\iota_1(C_\varepsilon)$ will be deformed is $\iota_1(p)$, where p is the pentagonal path depicted in Figure 2, with vertices at the points

$$\left\{\tfrac{4}{3} - \mathrm{i}\tfrac{2}{3},\ 2,\ \tfrac{4}{3} + \mathrm{i}\tfrac{2}{3},\ -\tfrac{2}{3} + \mathrm{i}\tfrac{2}{3},\ -\tfrac{2}{3} - \mathrm{i}\tfrac{2}{3}\right\}.$$

Denote by $p_1, \ldots, p_5$ the edges of p, as denoted in the figure.

Performing the suggested deformation results in the following lemma.

LEMMA 3.5.

$$a_{n,n} = 2\pi \mathrm{i} \int_{\iota_1(p)} \mathrm{Res}(\omega_F), \tag{3.3}$$

where p is oriented counterclockwise.

PROOF. For $\delta < \varepsilon$, let K be a homotopy shrinking the circle C_ε to the circle C_δ. By holomorphicity of ι_2 (Lemma 3.1), $\iota_2 \circ K$ is a homotopy from $\iota_2(C_\varepsilon)$ to $\iota_2(C_\delta)$ along $\mathcal{V}_Q$, and $\mathrm{Res}(\omega_F)$ is holomorphic along this homotopy. By Stokes' Theorem we obtain

$$\int_{\iota_2(C_\varepsilon)} \mathrm{Res}(\omega_F) = \int_{\iota_2(C_\delta)} \mathrm{Res}(\omega_F).$$

Now fix n large and let δ vary. Note that as the left hand side of the above equation has no δ dependence, neither does the right.

By the fact that $y_2(x) = -4x^{-5}(1 + O(x))$ as $x \to 0$, we get that $\frac{-P}{(2\pi \mathrm{i})^2 xyQy} = O(\delta^{-4})$, $\exp(nH) = O(\delta^{4n})$ and the area of $\iota_2(C_\delta)$ is $O(\delta^{-4})$ as $\delta \to 0$. This implies that

$$\int_{\iota_2(C_\delta)} \mathrm{Res}(\omega_F) = \int_{\iota_2(C_\delta)} \frac{1}{(2\pi \mathrm{i})^2} \cdot \frac{-P}{xyQ_y} e^{nH} \, \mathrm{d}x = O(\delta^{4n-8})$$

as $\delta \to 0$ (note that this representation of the residue is valid by Remark 3.2). For $n > 2$, $\delta^{4n-8} \to 0$ as $\delta \to 0$. Thus we must have that this integral is equal to 0.

As for the integral over $\iota_1(C_\varepsilon)$ in (3.3), let K now be a homotopy expanding the circle C_ε to the pentagonal path p. Then by Lemma 3.1, $\iota_1 \circ K$ is a homotopy from $\iota_1(C_\varepsilon)$ to $\iota_2(p)$ along $\mathcal{V}_Q$, and $\mathrm{Res}(\omega_F)$ is likewise holomorphic along the image of this homotopy. Then by Stokes' Theorem,

$$\int_{\iota_1(C_\varepsilon)} \mathrm{Res}(\omega_F) = \int_{\iota_1(p)} \mathrm{Res}(\omega_F),$$

where p is oriented counterclockwise. The theorem follows. □

Now we show that h is indeed maximized on $\iota_1(p)$ uniquely at the point $(2, 1/8)$. That this is true local to the saddle point $(2, 1/8)$ is clear from the form of H near this point, as explored following the proof of Lemma 3.4. To show that this is true globally will require more effort.

LEMMA 3.6. $h(\iota_1(x)) < h(\iota_1(2)) = \ln 4, \quad \forall x \in p \setminus \{2\}$.

PROOF. Because $h(\iota_1(x))$ is continuous on the connected set p, we need only show that $h(\iota_1(x)) \neq \ln 4$ for all $x \in p \setminus \{2\}$, and that $h(\iota_1(x)) < \ln 4$ for some $x \in p \setminus \{2\}$. The latter condition can be easily checked by plugging some arbitrary point into $h(\iota_1(x))$. As for the former condition, the idea will be to cook up some polynomial equations that must be satisfied in order for it to be true that $h(\iota_1(x)) = \ln 4$. We then use techniques from computational algebra to show that these equations cannot be satisfied for any (x, y) with $x \in p \setminus \{2\}$ and $y = y_1(x)$.

The conditions from which we shall derive our polynomial equations are as follows:

(1) $x \in p_j$ for some $j \in \{1, \ldots, 5\}$.
(2) y such that $(x, y) \in \mathcal{V}_Q$.
(3) $h(x, y) = \ln 4$, or $e^{h(x,y)} = 4$.

Each of these conditions implies a (set of) polynomial equations in the variables $\Re(x)$, $\Im(x)$, $\Re(y)$ and $\Im(y)$, as we shall show shortly. Note that we are throwing away some important information in condition 2 above, namely that we want $y = y_1(x)$, not $y = y_2(x)$. This will be important later in the proof.

We examine first the case where $x \in p_3$. Denote $a = \Re(x)$, $b = \Im(x)$, $c = \Re(y)$ and $d = \Im(y)$. Then condition 1 implies the polynomial constraint:

$$P_1 = b - \frac{2}{3} = 0.$$

Note: condition 1 implies the additional constraint $a \in [-2/3, 4/3]$, which we shall make use of shortly.

Condition 2 implies the following two polynomial constraints:

$$P_2 = \Re(Q(a + \mathrm{i}b, c + \mathrm{i}d)) = 0,$$
$$P_3 = \Im(Q(a + \mathrm{i}b, c + \mathrm{i}d)) = 0.$$

Finally, condition 3 translates to $4\,|x|\,|y| = 1$, or

$$P_4 = 16(a^2 + b^2)(c^2 + d^2) - 1 = 0.$$

We are interested in whether these four polynomial equations have a common *real-valued* solution, and we shall use Gröbner bases and Sturm sequences to answer this question. Since we expect the variety generated by $I = \langle P_1, P_2, P_3, P_4\rangle$ to be finite—I is generated by four polynomials in four unknowns—we hope to use Gröbner bases to eliminate variables and produce a univariate polynomial $B(a) \in I$. Any point (a, b, c, d) solving $P_j = 0$ for all j will likewise solve $B = 0$. Then we try to use Sturm sequences to that such a B has no real roots $a \in [-2/3, 4/3]$, proving that $h(\iota_1(x)) \neq \ln 4$ for $x \in p_3$.

We compute the Gröbner basis with the command

```
Basis([P1,P2,P3,P4],plex(d,c,b,a))
```

and find that the first element B of the basis *is* univariate in the variable a, a polynomial of degree 16. We can check that $B(-2/3) \neq 0$ and $B(4/3) \neq 0$ by direct computation in Maple. To check whether or not B has any roots on the interval $(-2/3, 4/3)$ we employ Sturm's Theorem (see [**BPR06**, p. 52]).

To employ Sturm's Theorem, we must verify that B is squarefree. This is true if and only if the ideal $\langle B, B'\rangle$ is equal to the trivial ideal $\langle 1\rangle$. Indeed, computing the Gröbner basis for $\langle B, B'\rangle$ by

```
Basis([B,diff(B,a)],plex(a));
```

returns the trivial basis `[1]`, i.e., B is squarefree.

Then to count the number of roots in $(-2/3, 4/3)$ via Sturm's Theorem, we enter the command

```
sturm(sturmseq(B,a),a,-2/3,4/3)
```

and Maple returns that there are 0 real roots on the interval $(-2/3, 4/3)$.

Computations are similar for p_4 and p_5, but things are a bit more complicated along p_1 and p_2. Let us look at p_2. The first polynomial equation becomes

$$P_1 = a + b - 2 = 0,$$

with $a \in [4/3, 2]$, while the rest of the polynomial equations remain the same. Going through the same procedure as before, we can produce a Gröbner basis for $\langle P_1, P_2, P_3, P_4 \rangle$ with an element $B(a)$ univariate in a. $B(a)$ factors as

$$B(a) = (a-2)^4 \tilde{B}(a),$$

where by direct computation we see that $\tilde{B}$ is nonzero at $a = 4/3$ and $a = 2$. Note: we expected that B would have a root at $a = 2$, corresponding to the fact that $h(\iota_1(2)) = \ln 4$.

The next step would be to attempt to show that $\tilde{B}$ has no roots on the interval $(4/3, 2)$, but this is not true. Using Sturm sequences, one can show that $\tilde{B}$ has exactly one root $a_0 \in (4/3, 2)$, and this is because there *is* a pair x, y with $x \in p_2 \backslash \{2\}$ and $h(x, y) = \ln 4$. The claim is that this corresponds to a point where $y = y_2(x)$, not where $y = y_1(x)$.

To see that there must be such a pair, note that $y_2(x) \to 0$ as $x \to 2$. Hence $h(\iota_2(x)) \to \infty$ as $x \to 2$. But by direct computation we can show $h(\iota_2(4/3)) < \ln 4$. As $h(\iota_2(x))$ is continuous on $p_2 \setminus \{2\}$, there must be some $x \in p_2 \setminus \{2\}$ such that $h(\iota_2(x)) = \ln 4$. This pair x, $y = y_2(x)$ satisfies the polynomial equations $P_j = 0$.

Now assume by way of contradiction that $h(\iota_1(x)) = \ln 4$ for some $x \in p_2 \setminus \{2\}$. Because $\tilde{B}$ has just one root $a_0 \in (4/3, 2)$, it must be that this occurs at the same x value for which $h(\iota_2(x)) = \ln 4$, specifically $x_0 = a_0 + (2 - a_0)\mathrm{i}$. Hence we have

$$|x_0|\,|y_1(x_0)| = |x_0|\,|y_2(x_0)| = \frac{1}{4},$$

which implies that $|y_1| = |y_2|$ at the point x_0. So at this value of x we have

$$c^2 + d^2 = |y|^2 = |y_1 y_2| = \frac{|x-2|}{|x|^5}.$$

The preceding equation implies that $|x|^{10}(c^2+d^2)^2 = |x-2|^2$, which translates into the polynomial equation

$$P_5 = (a^2+b^2)^5(c^2+d^2)^2 - ((a-2)^2+b^2) = 0.$$

We now have a new polynomial equation that must be satisfied in order that $h(\iota_1(x)) = \ln 4$ on $p_2 \setminus \{2\}$. But if we compute a Gröbner basis for $\langle P_1, \ldots, P_5 \rangle$, we get the trivial basis `[1]`, meaning that the polynomials have no common solution. Hence $h(\iota_1(x)) \neq \ln 4$ for $x \in p_2 \setminus \{2\}$. Analogous methods can be used to handle the case of p_1. □

3.4. Saddle point integration. The final step in the analysis is to use saddle point techniques and order bounds to prove (1.2).

THEOREM 3.7.

$$k_n = a_{n,n} \sim \frac{4^n}{8\Gamma(3/4)\, n^{5/4}}.$$

PROOF. We proceed from Lemma 3.5. The theorem will be proved in two steps: bounding the integral in (3.3) outside a neighborhood of the critical point, then applying saddle point techniques near that critical point.

For any neighborhood N of $x = 2$, we look at $\int_{\iota_1(p \setminus N)} \mathrm{Res}(\omega_F)$, which can be written as

$$\int_{\iota_1(p \setminus N)} \frac{1}{(2\pi\mathrm{i})^2} \cdot \frac{-P}{xy\, Q_y} e^{nH}\, \mathrm{d}x$$

(note that this representation is valid by Remark 3.2). As $h \circ \iota_1$ is continuous on the compact set $p \setminus N$, $h \circ \iota_1$ achieves an upper bound M on $p \setminus N$. By Lemma 3.6, $M < \ln 4$. Thus by trivial bounds we have

$$\int_{\iota_1(p\setminus N)} \operatorname{Res}(\omega_F) = O(e^{Mn}) = o((4-\delta)^n)$$

for sufficiently small $\delta > 0$, as $n \to \infty$. Hence

$$a_{n,n} = 2\pi\mathrm{i} \int_{\iota_1(p\cap N)} \operatorname{Res}(\omega_F) + o((4-\delta)^n) \tag{3.4}$$

for *any* neighborhood N of $x = 2$, provided δ is sufficiently small.

For N small enough, $p \cap N = (p_1 \cap N) \cup (p_2 \cap N)$. We examine the integral over $\iota_1(p_1 \cap N)$ and $\iota_1(p_2 \cap N)$ separately, starting with $\iota_1(p_2 \cap N)$. By using the aforementioned representation of the residue form (and changing variables), we obtain

$$2\pi\mathrm{i} \int_{\iota_1(p_2\cap N)} \operatorname{Res}(\omega_F) = \int_{p_2\cap N} \frac{1}{2\pi\mathrm{i}} \cdot \frac{-P(\iota_1(x))}{x y_1(x) Q_y(\iota_1(x))} e^{nH(\iota_1(x))}\,\mathrm{d}x.$$

After another change of variables ($x \mapsto x + 2$) and a suitable choice of neighborhood N, the above integral can be rewritten as

$$4^n \int_{\gamma^+} A(x) e^{-n\phi(x)}\,\mathrm{d}x,$$

where we have, for some fixed $\varepsilon > 0$,

$$\begin{aligned}
\gamma(x) &= (\mathrm{i}-1)x; \quad x \in [-\varepsilon, \varepsilon],\\
A(x) &= \frac{1}{2\pi\mathrm{i}} \cdot \frac{-P(\iota_1(x+2))}{(x+2) y_1(x+2) Q_y(\iota_1(x+2))},\\
\phi(x) &= \ln 4 - H(\iota_1(x+2)),
\end{aligned}$$

and we recall that γ^+ is the restriction of the image of γ to the domain $[0, \varepsilon]$. The series expansion of A and ϕ at $x = 0$ begin

$$\begin{aligned}
A(x) &= \frac{\mathrm{i}}{16\pi} x^3 + \frac{\mathrm{i}}{32\pi} x^4 + O(x^5),\\
\phi(x) &= \frac{-1}{16} x^4 + O(x^6),
\end{aligned}$$

and $\Re\phi(x)$ is uniquely minimized on γ^+ at $x = 0$ where we have $\phi(0) = 0$, as a consequence of Lemma 3.6. Thus this is exactly the situation where the saddle point technique of Theorem 2.11 can be applied. The values of b_j and c_j are as in the expansions above. Then $v = \gamma'(0) = \mathrm{i} - 1$, and we compute the principal root

$$\frac{(c_k n v^k)^{1/k}}{v} = \frac{((-1/16) n (\mathrm{i}-1)^4)^{1/4}}{\mathrm{i}-1} = \frac{-1-\mathrm{i}}{2\sqrt{2}} n^{1/4}.$$

The conclusion of Theorem 2.11 is then

$$2\pi\mathrm{i} \int_{\iota_1(p_2\cap N)} \operatorname{Res}(\omega_F) = 4^n \left(\frac{-\mathrm{i}}{4\pi} n^{-1} + \frac{(1+\mathrm{i})\sqrt{2}\,\Gamma(5/4)}{8\pi} n^{-5/4} + O(n^{-3/2}) \right)$$

As for the integral over $\iota_1(p_1 \cap N)$, the same argument yields

$$2\pi\mathrm{i} \int_{\iota_1(p_1\cap N)} \operatorname{Res}(\omega_F) = -4^n \int_{\gamma^+} A(x) e^{-n\phi(x)}\,\mathrm{d}x,$$

where A and ϕ are the same but γ is defined by $\gamma(x) = (-\mathrm{i}-1)x$ (and the negative sign in front comes from a reversal of orientation). For $v = \gamma'(0) = -\mathrm{i} - 1$, we compute the principal root

$$\frac{(c_k n v^k)^{1/k}}{v} = \frac{((-1/16)n(-\mathrm{i}-1)^4)^{1/4}}{-\mathrm{i}-1} = \frac{-1+\mathrm{i}}{2\sqrt{2}} n^{1/4}.$$

Then by Theorem 2.11 we obtain

$$2\pi\mathrm{i} \int_{\iota_1(p_1 \cap N)} \mathrm{Res}(\omega_F) = 4^n \left(\frac{\mathrm{i}}{4\pi} n^{-1} + \frac{(1-\mathrm{i})\sqrt{2}\,\Gamma(5/4)}{8\pi} n^{-5/4} + O(n^{-3/2}) \right).$$

Adding up the contribution of each piece and plugging into (3.4) yields

$$a_{n,n} = 4^n \left(\frac{\sqrt{2}\,\Gamma(5/4)}{4\pi} n^{-5/4} + O(n^{-3/2}) \right) + o((4-\delta)^n) \sim \frac{4^n \sqrt{2}\,\Gamma(5/4)}{4\pi} n^{-5/4}.$$

By using the identity $\Gamma(5/4)\Gamma(3/4) = \pi/(2\sqrt{2})$, the theorem follows. □

References

[AY83] Lev A. Aĭzenberg and Aleksandr P. Yuzhakov, *Integral Representations and Residues in Multidimensional Complex Analysis*, no. 58 in Translations of Mathematical Monographs, American Mathematical Society, Providence, RI, 1983. MR735793 (85a:32006)

[BPR06] Saugata Basu, Richard Pollack, and Marie-Françoise Roy, *Algorithms in Real Algebraic Geometry*, 2nd edition, no. 10 in Algorithms and Computation in Mathematics, Springer-Verlag, New York/Berlin, 2006. MR2248869 (2007b:14125)

[Bre93] Glen E. Bredon, *Topology and Geometry*, no. 139 in Graduate Texts in Mathematics, Springer-Verlag, New York/Berlin, 1993. MR1224675 (94d:55001)

[CLO05] David A. Cox, John B. Little, and Donal O'Shea, *Using Algebraic Geometry*, 2nd edition, no. 185 in Graduate Texts in Mathematics, Springer-Verlag, New York/Berlin, 2005. MR2122859 (2005i:13037)

[FS09] Philippe Flajolet and Robert Sedgewick, *Analytic Combinatorics*, Cambridge Univ. Press, Cambridge, UK, 2009. MR2483235

[Hen91] Peter Henrici, *Applied and Computational Complex Analysis. Vol. 2. Special Functions—Integral Transforms—Asymptotics—Continued Fractions*, Wiley, New York, 1991. MR1164865 (93b:30001)

[Pem09] Robin Pemantle, *Analytic combinatorics in more than one variable: Chapter 5*, University of Pennsylvania preprint, Dec. 2009. Currently available at `http://www.math.upenn.edu/~pemantle/581-html/chapter05.pdf`.

[PW02] Robin Pemantle and Mark C. Wilson, *Asymptotics of multivariate sequences. I. Smooth points of the singular variety*, J. Combin. Theory Ser. A **97** (2002), no. 1, 129–161. MR1879131 (2003a:05015)

[PW08] ———, *Twenty combinatorial examples of asymptotics derived from multivariate generating functions*, SIAM Rev. **50** (2008), no. 2, 199–272. MR2403050 (2009b:05023)

[RW08] Alexander Raichev and Mark C. Wilson, *A new approach to asymptotics of Maclaurin coefficients of algebraic functions*, Report CDMTCS-322, Centre for Discrete Mathematics and Theoretical Computer Science (CDMTCS), University of Auckland, Auckland, New Zealand, April 2008. Currently available at `http://www.cs.auckland.ac.nz/CDMTCS/researchreports/322alexmcw.pdf`.

[Saf00] Konstantin V. Safonov, *On power series of algebraic and rational functions in* $\mathbf{C}^n$, J. Math. Anal. Appl. **243** (2000), no. 2, 261–277. MR1741523 (2001f:32001)

[Sha92] Boris V. Shabat, *Introduction to Complex Analysis. Part II. Functions of Several Variables*, no. 110 in Translations of Mathematical Monographs, American Mathematical Society, Providence, RI, 1992. MR1192135 (93g:32001)

[Slo10] Neil J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, published electronically at `http://www.research.att.com/~njas/sequences`, 2010.

[Wat08] Waterloo Maple Inc., Maple, *version 12*, Waterloo, Ontario, Canada, 2008.

Department of Mathematics, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104, USA

E-mail address: `tdevries@math.upenn.edu`

URL: `http://www.math.upenn.edu/~tdevries`

This page intentionally left blank

Contemporary Mathematics
Volume **520**, 2010

Asymptotic normality of statistics on permutation tableaux

Paweł Hitczenko and Svante Janson

Abstract. We use a probabilistic approach to give new derivations of the expressions for the probability generating functions of basic statistics defined on permutation tableaux. Since our expressions are exact, we can identify the distributions of such basic statistics as the number of unrestricted rows, the number of rows, and the number of 1s in the first row. All three distributions were known, via involution on permutation tableaux and bijections between permutation tableaux and permutations, to be asymptotically normal after suitable normalizations. Our technique gives direct proofs of these results, as it allows us to work directly with permutation tableaux. We also establish the asymptotic normality of the number of superfluous 1s. This last result relies on a bijection between permutation tableaux and permutations, and on a rather general sufficient condition for the central limit theorem for the sums of random variables, in terms of a dependency graph of the summands.

1. Introduction

Permutation tableaux are relatively new objects that are in bijection with permutations [**Bu, CoN, SW**]. They were introduced in the context of enumeration of totally positive Grassmannian cells [**P, W**]. More recently, permutation tableaux generated additional research activity when they were connected in [**Co, CoW, CoW1**] to a particle model in statistical physics called the Partially ASymmetric Exclusion Process (PASEP); see [**BrE, BrCPRE, DDM, DEHP, DS, S, USW**] for more information on PASEP.

A *permutation tableau* [**SW**] is a Ferrers diagram of a partition of a positive integer into non-negative parts, whose boxes are filled with 0s and 1s according to the following rules:

(1) Each column of the diagram contains at least one 1.
(2) There is no 0 which has a 1 above it in the same column *and* a 1 to its left in the same row.

An example is given in Figure 1.

2000 *Mathematics Subject Classification.* Primary 60F05; Secondary 05A16, 05E10, 60E10.
Key words and phrases. Permutation tableau, central limit theorem.
The first author was supported in part by NSA Grant #H98230-09-1-0062.
The research of the second author was partially done while visiting Institut Mittag-Leffler, Djursholm, Sweden.

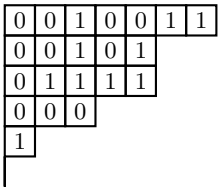


FIGURE 1. Example of a permutation tableau.

The size parameter of a permutation tableau is its *length*, defined as the number of rows plus the number of columns. For example, the tableau in Figure 1 has 6 rows and 7 columns, so its length is 13.

Various statistics on permutation tableaux were defined in [**CoW1, SW**]. We recall that a 0 in a permutation tableau is *restricted* if there is a 1 above it in the same column. A row is *unrestricted* if it does not contain a restricted 0. A 1 is *superfluous* if it has a 1 above itself in the same column. We shall be interested in the number of unrestricted rows, the number of superfluous 1s, as well as the number of 1s in the top row, and the number of rows. For example, the tableau in Figure 1 has three superfluous 1s, six rows, three 1s in the top row, and five unrestricted rows.

Most of the past research on statistics of permutation tableaux was based on bijections between permutation tableaux of length n and permutations of $[n] := \{1, \dots, n\}$, and used known properties of permutations. This was not always easy, as sometimes it is not that easy to see into what parameter a given statistic is mapped by a bijection. In [**CoH**] a direct approach based on a probabilistic consideration was proposed. It enabled the authors to compute the expected values of these statistics in a simple and unified way. In this work we go one step further: we use the approach proposed in [**CoH**] to compute the probability generating functions of the quantities in question. This will allow us, in particular, to obtain new and direct proofs of several results concerning the limiting distributions of (properly normalized) statistics on permutation tableaux. The original proofs of these results were based on bijective methods. As a matter of fact, since the approach proposed in [**CoH**] allows for exact (and not only asymptotic) computation, in most cases we shall be able to identify the distribution of a given statistic, exactly as was the case with the bijective approach. As consequences, we shall see, in particular, that:

- the number of unrestricted rows in a random permutation tableau of length n has the same distribution as the number of cycles, or the number of records, in a random permutation of $[n]$ (a record in a permutation $\sigma = (\sigma_i)$ is any σ_i such that $\sigma_i > \sigma_j$ for $j < i$);
- the number of rows has the same distribution as the number of descents in a random permutation counted by Eulerian numbers (a descent in (σ_i) is any pair (σ_i, σ_{i+1}) such that $\sigma_i > \sigma_{i+1}$);
- the number of 1s in the first row is, in distribution, one less than the number of unrestricted rows.

Earlier proofs of the first two results follow from the analysis of known bijections between permutations and permutation tableaux (see [**CoN, SW**]), while the third is a consequence of an involution on permutation tableaux presented in [**CoW1**]. However, neither the involution from [**CoW1**] nor any of the bijections from [**CoN, SW**] are very straightforward. In addition, the first two results given above required

differing bijections in [**SW**]. So, while these bijections carry more information, they turn out to be rather cumbersome to work with. In contrast, our approach allows us to work directly with permutation tableaux without the disadvantage of having to appeal to the bijections and the involution given in [**CoN, SW, CoW1**]. Looked at from this perspective, our approach provides a streamlined and unified derivation of the above results.

As for the number of superfluous 1s, the situation is a bit more complicated. Although we do derive its probability generating function, we do not deduce directly from this the asymptotic normality of its distribution. Instead, we shall rely on the fact that one of the two bijections between permutation tableaux and permutations described in [**SW**] sends the number of superfluous 1s to the number of occurrences of the generalized pattern $31{-}2$, and we shall prove the central limit theorem for the number of such occurrences in a random permutation of $[n]$ (an occurrence of the generalized pattern $31{-}2$ in a permutation (σ_i) is any pair (i,j), $1<i<j$, with $\sigma_{i-1}>\sigma_j>\sigma_i$). Our proof is based on a rather general sufficient condition for the central limit theorem developed in [**J**] and [**JŁR**, Section 6.1]. Bóna [**Bó**] has used the same method to obtain similar results on asymptotic normality for generalized permutation patterns, but as far as we know the central limit theorem for the number of occurrences of a generalized pattern such as $31{-}2$ is new. Furthermore, the number of permutation tableaux of length n with k superfluous 1s is equal to the number of permutations of $[n]$ with k crossings (a crossing [**Co**] in (σ_i) is a pair (i,j) such that $i<j\leqslant\sigma_i<\sigma_j$ or $i>j>\sigma_i>\sigma_j$). It follows, therefore, that the number of crossings in a random permutation of $[n]$ is asymptotically normal.

2. Basic facts

Let $\mathcal{T}_n$ be the set of all permutation tableaux of length n. We denote the uniform probability measure on $\mathcal{T}_n$ by P_n, and E_n will denote the expectation with respect to P_n. We denote by R_n, C_n, U_n, F_n, and S_n the random variables representing the numbers of rows, columns, unrestricted rows, 1s in the top row, and superfluous 1s, respectively, in a random tableau of length n. (We may allow $n=0$; there is a single empty permutation tableau of length 0, and $U_0=F_0=R_0=C_0=S_0=0$. It is sometimes convenient to start inductions with $n=0$, but for simplicity we often treat $n\geqslant 1$ only.)

If $\mathcal{F}$ is a σ-algebra then $\mathsf{E}(\,\cdot\,|\mathcal{F})$ denotes the conditional expectation given $\mathcal{F}$. If X is non-negative integer valued, for example one of the permutation tableaux statistics just defined, we let $g_X(z)=\mathsf{E}\,z^X$ be its probability generating function (in general defined at least for $|z|\leqslant 1$; for the variables consider here, z can be any complex number). We shall often omit the subscript X.

The arguments in [**CoH**] were based on a construction of tableaux of size k from tableaux of size $k-1$ by extending the latter from the south-west (SW) corner either to the south (creating a new empty row) or to the west (creating and then filling a new column). Note that each permutation tableau in $\mathcal{T}_k$ is an extension of a unique permutation tableau in $\mathcal{T}_{k-1}$. We refer the reader to [**CoH**, Section 2] for details. We let M_k indicate the direction of the move at the kth step (i.e., when the length is increased from $k-1$ to k). We refer to M_k as the kth move, and write $M_k=\mathsf{S}$ or $M_k=\mathsf{W}$ to indicate its direction.

The following simple observations were crucial for the arguments in [**CoH**] and are crucial here as well (see [**CoH**] for some details that are omitted here).

- When extending a tableau by a W move, the new column must be filled with 0s and 1s; each restricted row must be filled with 0s, but the unrestricted rows can be filled arbitrarily except that there must be at least one 1. Hence, a tableau with U unrestricted rows can be extended by a W move in $2^U - 1$ ways; since there always is a unique S move, the total number of extensions is 2^U.
- We use T_n to denote a generic element of $\mathcal{T}_n$, and use T_{n-1} to denote the corresponding element of $\mathcal{T}_{n-1}$ (i.e., such that T_n is an extension of T_{n-1}); we let $U_n = U_n(T_n)$ be the number of unrestricted rows in T_n and $U_{n-1} = U_{n-1}(T_{n-1})$ the number of unrestricted rows in T_{n-1}, and similarly for the other variables that we study. In this way, P_n induces a probability measure (also denoted P_n) on $\mathcal{T}_{n-1}$, namely each element of $\mathcal{T}_{n-1}$ is assigned a measure that is proportional to the number of tableaux from $\mathcal{T}_n$ that it generates when its length is increased from $n-1$ to n. Note that this differs from the uniform distribution P_{n-1} on $\mathcal{T}_{n-1}$. Since $|\mathcal{T}_n| = n!$ and a tableau T from $\mathcal{T}_{n-1}$ generates $2^{U_{n-1}(T)}$ tableaux of length n, the relationship between these two measures is easy to find. One way to state this relationship is that if X is any random variable on $\mathcal{T}_{n-1}$ then

$$\mathsf{E}_n X = \frac{1}{n}\,\mathsf{E}_{n-1}\left(2^{U_{n-1}} X\right), \tag{2.1}$$

  where integration on the left is with respect to the measure induced on $\mathcal{T}_{n-1}$ by P_n, and integration on the right is with respect to P_{n-1}. Equivalently, if we let $\mathcal{F}_{n-1}$ be the σ-algebra on $\mathcal{T}_n$ generated by the mapping $T_n \mapsto T_{n-1}$, then $\mathrm{d}\,\mathsf{P}_n / \mathrm{d}\,\mathsf{P}_{n-1} = 2^{U_{n-1}}/n$ on $\mathcal{F}_{n-1}$.
- The sequence of distributions of the number of unrestricted rows is given as follows: $U_0 \equiv 0$, $U_1 \equiv 1$, and for $n \geqslant 1$ the conditional distribution of U_n given $\mathcal{F}_{n-1}$ (i.e., given T_{n-1}) is (under P_n)

$$\mathcal{L}(U_n \mid \mathcal{F}_{n-1}) = 1 + \mathrm{Bin}(U_{n-1}), \tag{2.2}$$

  where $\mathrm{Bin}(m)$ denotes a binomial random variable with parameters m and $1/2$.

To illustrate how these facts are put together, consider the number of unrestricted rows, U_n. For its probability generating function we have:

$$\begin{aligned} g_{U_n}(z) = \mathsf{E}_n\, z^{U_n} &= \mathsf{E}_n\, \mathsf{E}_n\left(z^{U_n} \mid \mathcal{F}_{n-1}\right) = \mathsf{E}_n\, \mathsf{E}_n\left(z^{1+\mathrm{Bin}(U_{n-1})} \mid \mathcal{F}_{n-1}\right) \\ &= z\,\mathsf{E}_n\Big(\frac{z+1}{2}\Big)^{U_{n-1}} = \frac{z}{n}\,\mathsf{E}_{n-1}\left(2^{U_{n-1}}\Big(\frac{z+1}{2}\Big)^{U_{n-1}}\right) \\ &= \frac{z}{n}\,\mathsf{E}_{n-1}\,(z+1)^{U_{n-1}}, \end{aligned} \tag{2.3}$$

where we have used (in order) conditioning, (2.2), the obvious fact that $g_{\mathrm{Bin}(m)}(z) = \mathsf{E}\, z^{\mathrm{Bin}(m)} = \left(\frac{z+1}{2}\right)^m$, and (2.1). It follows by induction, and the fact $g_{U_1}(z) = z$, that

$$g_{U_n}(z) = \frac{\Gamma(z+n)}{\Gamma(z)\,n!} = \binom{z+n}{n} = \prod_{j=0}^{n-1} \frac{z+j}{j+1} = \prod_{k=1}^{n}\Big(1 - \frac{1}{k} + \frac{z}{k}\Big). \tag{2.4}$$

The factor on the right-hand side is the probability generating function of a random variable which is 1 with probability $1/k$ and 0 with probability $1 - 1/k$. Since the

product of probability generating functions corresponds to summing independent random variables, we obtain the following statement

THEOREM 2.1. *The number of unrestricted rows U_n is distributed like*

$$U_n \stackrel{\mathrm{d}}{=} \sum_{k=1}^{n} J_k,$$

where J_1, $J_2, \ldots,$ are independent indicators with $\mathsf{P}(J_k = 1) = 1/k$. In particular, if $h_n = \sum_{k=1}^{n} k^{-1}$ and $h_n^{(2)} = \sum_{k=1}^{n} k^{-2}$ are the nth harmonic and generalized harmonic numbers, respectively, then

$$\mathsf{E}_n U_n = \sum_{k=1}^{n} \mathsf{P}(J_k) = h_n, \quad \mathsf{var}_n(U_n) = \sum_{k=1}^{n} \mathsf{var}(J_k) = \sum_{k=1}^{n} \frac{1}{k}\left(1 - \frac{1}{k}\right) = h_n - h_n^{(2)},$$

and

$$\frac{U_n - \ln n}{\sqrt{\ln n}} \xrightarrow{\mathrm{d}} N(0,1),$$

where $N(0,1)$ denotes a standard normal random variable.

REMARK 2.2. It is seen from the above statement that the distribution of U_n coincides with the distribution of the number of cycles in a random permutation of $[n]$ (see for example [**F**, Chapter X, Section 6(b)] or [**R**, Chapter 4, Section 3]), or, equivalently, the number of RL minima in a random permutation (which is known to be equidistributed with the number of cycles). Indeed, the bijections between permutation tableaux and permutations described in [**CoN**] map unrestricted rows onto RL minima in the corresponding permutation.

REMARK 2.3. More generally, we may define on $\mathcal{T}_n$ permutation tableaux $T_k \in \mathcal{T}_k$ for every $k \leqslant n$, such that T_{k+1} is an extension of T_k, and let $\mathcal{F}_k$ be the σ-algebra generated by T_k. It can be seen by induction, similar to (2.3), that a permutation tableau T_k of length k can be extended to $(n-k)!\,(n-k+1)^{U_k(T_k)}$ permutation tableaux in $\mathcal{T}_n$, and hence $\mathrm{d}\,\mathsf{P}_n / \mathrm{d}\,\mathsf{P}_k = (n-k+1)^{U_k} / \binom{n}{k}$ on $\mathcal{F}_k$, $1 \leqslant k \leqslant n$. Furthermore, (2.4) can be generalized to

$$\mathsf{E}_n(z^{U_n} \mid \mathcal{F}_k) = \frac{\Gamma(z+n-k)k!}{\Gamma(z)\,n!}(z+n-k)^{U_k}.$$

We can study further the sequence $(U_k)_{k=1}^{n}$ as a stochastic process under P_n; similar calculations show that this is an inhomogeneous Markov process with transitions given by $\mathcal{L}(U_{k+1} \mid \mathcal{F}_k) = 1 + \mathrm{Bin}(U_k, \frac{n-k}{n-k+1})$, which generalizes (2.2).

3. The number of 1s in the first row

Recall that F_n denotes the number of 1s in the top row of a random permutation tableau of length n. The following fact is a consequence of an involution on permutation tableaux given in [**CoW1**]. Since this involution is not too easy to describe, we provide a more direct justification of (3.1) based on the arguments given in the preceding section.

THEOREM 3.1. *For every $n \geqslant 1$, we have*

$$F_n \stackrel{\mathrm{d}}{=} U_n - 1. \tag{3.1}$$

In particular,

$$\mathsf{E}_n F_n = h_n - 1, \qquad \mathsf{var}_n(F_n) = h_n - h_n^{(2)},$$

and

$$\frac{F_n - \ln n}{\sqrt{\ln n}} \xrightarrow{\mathrm{d}} N(0,1).$$

To establish (3.1), it will be convenient to prove some auxiliary lemmas. Define G_k to be the position (counting only unrestricted rows) of the topmost 1 on the kth move, provided that this move is W. Then

$$F_n = \sum_{k=1}^{n} I_{G_k=1}.$$

Note that G_k is undefined if the kth move is S. While it is inconsequential in this section as we shall be interested in the event $\{G_k = 1\}$, for the purpose of the subsequent sections it is convenient to set $G_k = U_{k-1} + 1$ if the kth move is S. We have

LEMMA 3.2. *For all $k \geqslant 2$,*

$$\mathsf{P}_k(G_k = 1 \mid \mathcal{F}_{k-1}) = \frac{1}{2}.$$

PROOF. A given tableau $T \in \mathcal{T}_{k-1}$ has $2^{U_{k-1}(T)}$ extensions. The number of extensions by a W move with a 1 in the topmost row (which always is unrestricted) is $2^{U_{k-1}(T)-1}$ since the other $U_{k-1}(T) - 1$ unrestricted rows can be filled arbitrarily. Hence the probability that $G_k = 1$ in a random extension of T is $2^{U_{k-1}(T)-1}/2^{U_{k-1}(T)} = 1/2$. □

LEMMA 3.3. *For any complex numbers z, w and every $k \geqslant 2$,*

$$\mathsf{E}_k(z^{I_{G_k}} w^{U_k} \mid \mathcal{F}_{k-1}) = w\,\frac{z+w}{w+1}\left(\frac{w+1}{2}\right)^{U_{k-1}}.$$

PROOF. Let $T \in \mathcal{T}_{k-1}$. If T is extended with $G_k = 1$, i.e., by a W move with 1 added in the topmost row, then we may add 0 or 1 arbitrarily to all other unrestricted rows, and the rows with 1 added remain unrestricted. Hence, conditionally given $G_k = 1$, the extension has $1 + \mathrm{Bin}(U_{k-1}(T) - 1)$ unrestricted rows. Recall from (2.2) that without further conditioning, the number of unrestricted rows is $1 + \mathrm{Bin}(U_{k-1}(T))$. Consequently, using also Lemma 3.2,

$$\begin{aligned}
\mathsf{E}_k(z^{I_{G_k}} w^{U_k} \mid \mathcal{F}_{k-1}) &= z\,\mathsf{E}_k(w^{U_k} I_{G_k=1} \mid \mathcal{F}_{k-1}) + \mathsf{E}_k(w^{U_k} I_{G_k\neq 1} \mid \mathcal{F}_{k-1}) \\
&= (z-1)\,\mathsf{E}_k(w^{U_k} I_{G_k=1} \mid \mathcal{F}_{k-1}) + \mathsf{E}_k(w^{U_k} \mid \mathcal{F}_{k-1}) \\
&= (z-1)\,\mathsf{E}_k(w^{1+\mathrm{Bin}(U_{k-1}-1)} I_{G_k=1} \mid \mathcal{F}_{k-1}) + \mathsf{E}_k(w^{1+\mathrm{Bin}(U_{k-1})} \mid \mathcal{F}_{k-1}) \\
&= (z-1)w\left(\frac{w+1}{2}\right)^{U_{k-1}-1} \mathsf{P}_k(G_k = 1 \mid \mathcal{F}_{k-1}) + w\left(\frac{w+1}{2}\right)^{U_{k-1}} \\
&= (z-1)\frac{w}{w+1}\left(\frac{w+1}{2}\right)^{U_{k-1}} + w\left(\frac{w+1}{2}\right)^{U_{k-1}} \\
&= w\,\frac{z+w}{w+1}\left(\frac{w+1}{2}\right)^{U_{k-1}}.
\end{aligned}$$

□

LEMMA 3.4. *The joint probability generating function of F_n and U_n is given by*

$$\mathsf{E}_n\big(z^{F_n}w^{U_n}\big) = w\frac{\Gamma(z+w+n-1)}{\Gamma(z+w)\,n!} = w\prod_{k=2}^{n}\frac{z+w+k-2}{k}.$$

PROOF. By Lemma 3.3 and (2.1),

$$\begin{aligned}\mathsf{E}_n\big(z^{F_n}w^{U_n}\big) &= \mathsf{E}_n\,\mathsf{E}_n\big(z^{F_{n-1}+I_{G_n=1}}w^{U_n}\mid\mathcal{F}_{n-1}\big)\\ &= \mathsf{E}_n\left(z^{F_{n-1}}w\frac{z+w}{w+1}\Big(\frac{w+1}{2}\Big)^{U_{n-1}}\right)\\ &= w\,\frac{z+w}{w+1}\,\frac{1}{n}\,\mathsf{E}_{n-1}\left(z^{F_{n-1}}(w+1)^{U_{n-1}}\right),\end{aligned}$$

and the formula follows by induction. □

We can now complete the proof of Theorem 3.1.

PROOF OF THEOREM 3.1. Taking $w=1$ in Lemma 3.4 we obtain

$$\mathsf{E}_n\,z^{F_n} = \frac{\Gamma(z+n)}{\Gamma(z+1)\,n!},$$

which equals $\mathsf{E}_n\,z^{U_n-1}$ by (2.4). □

Note that we recover (2.4) by taking $z=1$ in Lemma 3.4. We can also describe the joint distribution of F_n and U_n.

THEOREM 3.5. *For every $n\geqslant 1$, the joint distribution of F_n and U_n is given by*

$$(F_n,U_n)\stackrel{\mathrm{d}}{=}\sum_{k=1}^{n}(J_k,I_k),$$

where the random vectors (J_k,I_k) are independent and $J_1=0$, $I_1=1$, and, for $k\geqslant 2$, $\mathsf{P}(I_k=1,\,J_k=0)=\mathsf{P}(I_k=0,\,J_k=1)=1/k$, $\mathsf{P}(I_k=0,\,J_k=0)=1-2/k$.

PROOF. We have $\mathsf{E}(z^{J_k}w^{I_k})=(k-2+z+w)/k$ if $k\geqslant 2$ and w if $k=1$, and thus the joint probability generating function of the right-hand side equals the product in Lemma 3.4. □

COROLLARY 3.6. *The covariance of F_n and U_n is*

$$\mathsf{cov}_n(F_n,U_n)=\mathsf{E}_n(F_nU_n)-\mathsf{E}_n\,F_n\,\mathsf{E}_n\,U_n=-(h_n^{(2)}-1).$$

PROOF. We have $\mathsf{cov}(J_k,I_k)=-\,\mathsf{E}\,J_k\,\mathsf{E}\,I_k=-k^{-2}$ for $k\geqslant 2$, and $\mathsf{cov}(J_1,I_1)=0$, and thus by Theorem 3.5, $\mathsf{cov}(F_n,U_n)=\sum_{k=2}^{n}k^{-2}$. □

It easily follows further from the central limit theorem for vector-valued random variables, e.g., with Lyapunov's condition, that the normal limits in Theorems 2.1 and 3.1 hold jointly, with the joint limit being a pair of two independent standard normal variables. We omit the details.

4. The number of rows

In this section we consider the number of rows R_n in a random permutation tableau of length n. Since the rows of a permutation tableau correspond to S steps in the process of its construction, we can write

$$R_n = \sum_{k=1}^{n} I_{M_k=\mathsf{S}}.$$

LEMMA 4.1. *For $1 \leqslant k \leqslant n$ and any complex numbers z, w we have*

$$\mathsf{E}_k\big(z^{R_k} w^{U_k}\big) = \frac{w(z-1)}{k}\, \mathsf{E}_{k-1}\big(z^{R_{k-1}} w^{U_{k-1}}\big) + \frac{w}{k}\, \mathsf{E}_{k-1}\big(z^{R_{k-1}} (w+1)^{U_{k-1}}\big).$$

PROOF. By conditioning on $\mathcal{F}_{k-1}$ we get

$$\mathsf{E}_k(z^{R_k} w^{U_k}) = \mathsf{E}_k\left(z^{R_{k-1}}\, \mathsf{E}_k(z^{I_{M_k=\mathsf{S}}} w^{U_k} \mid \mathcal{F}_{k-1})\right). \tag{4.1}$$

Note that $M_k = \mathsf{S}$ if and only if $U_k = 1 + U_{k-1}$ and, conditionally on $\mathcal{F}_{k-1}$, this happens with probability $1/2^{U_{k-1}}$. Hence, using also Lemma 3.3 with $z = 1$ (or the argument in (2.3)),

$$\begin{aligned}
\mathsf{E}_k(z^{I_{M_k=\mathsf{S}}} w^{U_k} \mid \mathcal{F}_{k-1}) &= \mathsf{E}_k\big((z^{I_{M_k=\mathsf{S}}} - 1) w^{U_k} \mid \mathcal{F}_{k-1}\big) + \mathsf{E}_k(w^{U_k} \mid \mathcal{F}_{k-1}) \\
&= \mathsf{E}_k\big((z-1) I_{M_k=\mathsf{S}} w^{U_k} \mid \mathcal{F}_{k-1}\big) + w \left(\frac{w+1}{2}\right)^{U_{k-1}} \\
&= (z-1)\frac{w^{1+U_{k-1}}}{2^{U_{k-1}}} + w \left(\frac{w+1}{2}\right)^{U_{k-1}} \\
&= w(z-1)\left(\frac{w}{2}\right)^{U_{k-1}} + w \left(\frac{w+1}{2}\right)^{U_{k-1}}.
\end{aligned}$$

Putting this into (4.1) and applying (2.1) proves the lemma. □

We can now compute the probability generating function of R_n.

THEOREM 4.2. *We have*

$$g_{R_n}(z) = \frac{1}{n!} \sum_{r=1}^{n} \left\langle {n \atop r-1} \right\rangle z^r,$$

where $\left\langle {n \atop k} \right\rangle$ are the Eulerian numbers counting the number of permutations $\sigma = (\sigma_1, \dots, \sigma_n)$ of $[n]$ with k descents (recall that a descent is a pair (σ_j, σ_{j+1}) such that $\sigma_j > \sigma_{j+1}$). Thus R_n has the same distribution as 1 plus the number of descents in a random permutation of $[n]$.

PROOF. Consider $\mathsf{E}_n\, z^{R_n}$. Applying Lemma 4.1 k times and collecting the terms involving the expectation of the same expression, we see that it is of the form

$$\mathsf{E}_n\, z^{R_n} = \frac{1}{n \cdots (n-k+1)} \sum_{m=0}^{k} c_{k,m}(z)\, \mathsf{E}_{n-k}\big(z^{R_{n-k}} (m+1)^{U_{n-k}}\big) \tag{4.2}$$

(for $0 \leqslant k \leqslant n$), with certain coefficient functions $c_{k,m} = c_{k,m}(z)$. Apply Lemma 4.1 again to each of the expectations in the sum to get

$$\begin{aligned}
\mathsf{E}_{n-k}\big(z^{R_{n-k}} (m+1)^{U_{n-k}}\big) = {} & \frac{(m+1)(z-1)}{n-k}\, \mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}} (m+1)^{U_{n-k-1}}\big) \\
& + \frac{m+1}{n-k}\, \mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}} (m+2)^{U_{n-k-1}}\big).
\end{aligned}$$

Putting that back in, we see that $(n\cdots(n-k))g_{R_n}(z)$ is equal to

$$\sum_{m=0}^{k} c_{k,m}(m+1)(z-1)\,\mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}}(m+1)^{U_{n-k-1}}\big) + \sum_{m=0}^{k} c_{k,m}(m+1)\,\mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}}(m+2)^{U_{n-k-1}}\big).$$

By rearranging the terms this is

$$\begin{aligned}
&c_{k,0}(z-1)\,\mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}}\big)\\
&\qquad+\sum_{m=1}^{k} c_{k,m}(m+1)(z-1)\,\mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}}(m+1)^{U_{n-k-1}}\big)\\
&\qquad+\sum_{m=0}^{k-1} c_{k,m}(m+1)\,\mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}}(m+2)^{U_{n-k-1}}\big)\\
&\qquad+c_{k,k}(k+1)\,\mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}}(k+2)^{U_{n-k-1}}\big)\\
&=c_{k,0}(z-1)\,\mathsf{E}_{n-k-1}\,z^{R_{n-k-1}}\\
&\qquad+\sum_{m=1}^{k}\{c_{k,m}(m+1)(z-1)+mc_{k,m-1}\}\,\mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}}(m+1)^{U_{n-k-1}}\big)\\
&\qquad+c_{k,k}(k+1)\,\mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}}(k+2)^{U_{n-k-1}}\big)\\
&=\sum_{m=0}^{k+1} c_{k+1,m}\,\mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}}(m+1)^{U_{n-k-1}}\big),
\end{aligned}$$

where the coefficients $c_{k,m}$ satisfy the following recurrence: $c_{k,m}=0$ unless $0\leqslant m\leqslant k$, $c_{0,0}=1$, and for $0\leqslant m\leqslant k$,

$$c_{k,m}=m\,c_{k-1,m-1}+(m+1)(z-1)c_{k-1,m}.$$

It follows by induction that $c_{k,m}(z)$ is a constant times $(z-1)^{k-m}$, so

$$c_{k,m}=a_{k,m}(z-1)^{k-m},$$

where $a_{0,0}=1$ and

$$a_{k,m}=\begin{cases} ma_{k-1,m-1}+(m+1)a_{k-1,m}, & \text{if } 0\leqslant m\leqslant k,\ k\geqslant 1;\\ 0, & \text{otherwise.}\end{cases} \tag{4.3}$$

If we now let $k=n-1$ in (4.2) and use $\mathsf{E}_1\big(z^{R_1}(m+1)^{U_1}\big)=z(m+1)$, then we get

$$g_{R_n}(z)=\frac{z}{n!}\sum_{m=0}^{n-1}(m+1)c_{n-1,m}=\frac{z}{n!}\sum_{m=0}^{n-1}(m+1)a_{n-1,m}(z-1)^{n-1-m}. \tag{4.4}$$

The recurrence (4.3) is solved by

$$a_{k,m}=m!\left\{{k+1\atop m+1}\right\},$$

where $\left\{ {k \atop m} \right\}$ is the number of partitions of a k-element set into m non-empty subsets (recall the basic recurrence $\left\{ {k+1 \atop m+1} \right\} = (m+1)\left\{ {k \atop m+1} \right\} + \left\{ {k \atop m} \right\}$). Hence, (4.4) becomes

$$\begin{aligned} g_{R_n}(z) &= \frac{z}{n!} \sum_{m=0}^{n-1} (m+1)! \left\{ {n \atop m+1} \right\} (z-1)^{n-(m+1)} \\ &= \frac{z}{n!} \sum_{m=0}^{n} m! \left\{ {n \atop m} \right\} (z-1)^{n-m}, \end{aligned}$$

where in the last step we shifted the index by one and used $\left\{ {n \atop 0} \right\} = 0$ for $n \geqslant 1$. This completes the proof, since for a complex z (see [**K**, Section 5.1.3]),

$$\sum_{r=1}^{n} \left\langle {n \atop r-1} \right\rangle \frac{z^r}{n!} = \frac{z}{n!} \sum_{m=0}^{n} m! \left\{ {n \atop m} \right\} (z-1)^{n-m}. \qquad \square$$

Once the coefficients have been identified as Eulerian numbers we can use their known properties (see, e.g., [**DB**, Chapter 10, pp. 150–154]) to obtain

COROLLARY 4.3. *The number of rows R_n in a random permutation tableau of length n satisfies*

$$\frac{R_n - (n+1)/2}{\sqrt{(n+1)/12}} \xrightarrow{\mathrm{d}} N(0,1). \tag{4.5}$$

In fact, a local limit theorem holds as well (see, e.g., [**CKSS**] or [**E**]).

Moreover, for the number of columns in a random permutation tableau, by definition, we have $C_n = n - R_n$. Further, let D_n denote the number of descents in a random permutation; thus $R_n \stackrel{\mathrm{d}}{=} D_n + 1$ by Theorem 4.2. Hence $C_n = n - R_n \stackrel{\mathrm{d}}{=} n - 1 - D_n = A_n$, the number of ascents in a random permutation. By symmetry, $A_n \stackrel{\mathrm{d}}{=} D_n \stackrel{\mathrm{d}}{=} R_n - 1$, and thus we obtain the following symmetry property:

COROLLARY 4.4. *The number C_n of columns in a random permutation tableau satisfies $C_n \stackrel{\mathrm{d}}{=} R_n - 1$. In particular,*

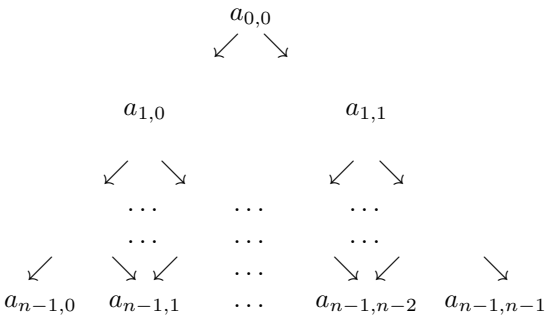
$$\frac{C_n - (n-1)/2}{\sqrt{(n+1)/12}} \xrightarrow{\mathrm{d}} N(0,1).$$

The fact that $R_n - 1 \stackrel{\mathrm{d}}{=} C_n \stackrel{\mathrm{d}}{=} D_n \stackrel{\mathrm{d}}{=} A_n$ follows also, by the bijections described in [**CoN**].

REMARK 4.5. The coefficients in (4.4) can be more explicitly written as

$$a_{n-1,n-r-1} = (n-r-1)! \sum_{1 \leqslant j_1 < \cdots < j_r \leqslant n-1} j_1 (j_2 - 1) \cdots (j_r - (r-1)). \tag{4.6}$$

This can be seen by putting them in a Pascal-type triangle



and observing that by (4.3), a move down (either SW or SE) from a coefficient $a_{k,m}$ has weight $m+1$. The value of a given coefficient at the bottom is obtained by summing, over all possible paths leading to it from the root $a_{0,0}$, the products of weights corresponding to the moves along the path. Any path contributing to $a_{n-1,n-1-r}$, $0 \leqslant r \leqslant n-1$ has exactly r SW moves, and if they are from levels $j_1, \dots, j_r$ (with level 0 at the top and $n-1$ at the bottom), then the path has weight

$$(n-1-r)!\,(j_1+1)j_2(j_3-1)\cdots(j_r-(r-2)),$$

and (4.6) follows by renumbering the terms.

Expression (4.6) gives a direct way to get the moments; from (4.4) the probability generating function of $R_n - 1$ is

$$g_{R_n-1}(z) = \frac{1}{n!}\sum_{r=0}^{n-1}(n-r)a_{n-1,n-r-1}(z-1)^r,$$

and thus the rth factorial moment of $R_n - 1$ is

$$\begin{aligned}\mathsf{E}_n\big((R_n-1)(R_n-2)\cdots(R_n-r)\big) &= \frac{\mathrm{d}^r}{\mathrm{d}z^r}g_{R_n-1}(z)\Big|_{z=1} = \frac{r!(n-r)}{n!}a_{n-1,n-r-1}\\ &= \frac{r!\,(n-r)!}{n!}\sum_{1\leqslant j_1<\cdots<j_r\leqslant n-1} j_1(j_2-1)\cdots(j_r-(r-1)).\end{aligned}$$

In particular,

$$\mathsf{E}_n\, R_n = 1 + \mathsf{E}_n(R_n-1) = 1 + \frac{1}{n}\sum_{j=1}^{n-1} j = 1 + \frac{n-1}{2} = \frac{n+1}{2},$$

and

$$\begin{aligned}\mathsf{var}(R_n) = \mathsf{var}(R_n-1) &= \mathsf{E}_n((R_n-1)(R_n-2)) + \mathsf{E}_n(R_n-1) - (\mathsf{E}_n(R_n-1))^2\\ &= \frac{2}{(n-1)n}\sum_{1\leqslant j_1<j_2\leqslant n-1} j_1(j_2-1) + \frac{n-1}{2} - \frac{(n-1)^2}{4}\\ &= \frac{(n-2)(3n-5)}{12} + \frac{n-1}{2} - \frac{(n-1)^2}{4} = \frac{n+1}{12},\end{aligned}$$

which explains the normalization in (4.5).

5. Probability generating function of the number of superfluous 1s

Let S_n be the number of superfluous 1s in a random tableau of length n. In [**CoH**] calculations based on (2.1) and (2.2) were used to show that

$$\mathsf{E}_n\, S_n = \frac{(n-1)(n-2)}{12}. \tag{5.1}$$

In this section we use the same approach to derive the expression for the probability generating function of S_n, see Proposition 5.3 or equation (5.5) below. Although the form of the probability generating function looks rather unwieldy, it can be used to compute (in practice low order) moments of S_n in a relatively straightforward, albeit a bit tedious way. We shall illustrate it by deriving the exact expression for the variance of S_n (see Proposition 5.5) below. However, there does not seem to be an easy way to derive the central limit theorem for S_n directly from the probability generating function. For this reason, in the next section we shall follow a different approach to establish the CLT for S_n.

We write in this section

$$S_n = \sum_{k=1}^{n} V_k,$$

where V_k is the increase in the number of superfluous 1s when the length of a tableau is increased from $k-1$ to k. As in the preceding section, let G_k be the row number of the topmost 1 (provided that the kth move is W and counting only unrestricted rows), and recall that it is set to be $U_{k-1}+1$ if the kth move is S. Then by the same argument as in Lemma 3.2,

$$\mathsf{P}_k(G_k = j \mid \mathcal{F}_{k-1}) = \frac{1}{2^j}, \quad j = 1, \dots, U_{k-1},$$

and

$$\mathsf{P}_k(G_k = U_{k-1}+1 \mid \mathcal{F}_{k-1}) = \frac{1}{2^{U_{k-1}}}.$$

Moreover, the joint conditional distribution of (U_k, V_k) given $\mathcal{F}_{k-1}$ is

$$\mathcal{L}((U_k, V_k) \mid \mathcal{F}_{k-1}) = (G_k + \mathrm{Bin}(U_{k-1} - G_k), \mathrm{Bin}(U_{k-1} - G_k)), \tag{5.2}$$

in the sense that if we further condition on $G_k = m$, then the distribution equals that of $(m + X, X)$ with $X \sim \mathrm{Bin}(U_{k-1} - m)$ (thus the two occurences of the expression $\mathrm{Bin}(U_{k-1} - G_k)$ in (5.2) signify the same random variable); further we interpret $\mathrm{Bin}(-1)$ as 0.

Lemma 5.1. *We have, for all complex z and w, and $k \geqslant 1$:*

$$\mathsf{E}_k(z^{V_k} w^{U_k} \mid \mathcal{F}_{k-1}) = w\left\{\frac{1}{1+w(z-1)}\left(\frac{zw+1}{2}\right)^{U_{k-1}} + \left(1 - \frac{1}{1+w(z-1)}\right)\left(\frac{w}{2}\right)^{U_{k-1}}\right\}$$

and

$$\mathsf{E}_k(z^{S_k} w^{U_k}) = \frac{w}{k}\left\{\frac{1}{1+w(z-1)}\,\mathsf{E}_{k-1}\left(z^{S_{k-1}}(zw+1)^{U_{k-1}}\right) + \left(1 - \frac{1}{1+w(z-1)}\right)\mathsf{E}_{k-1}\left(z^{S_{k-1}} w^{U_{k-1}}\right)\right\}.$$

PROOF. Using (5.2) we have

$$\begin{aligned}
\mathsf{E}_k(z^{V_k}w^{U_k} \mid \mathcal{F}_{k-1}) &= \mathsf{E}_k(z^{\mathrm{Bin}(U_{k-1}-G_k)}w^{G_k+\mathrm{Bin}(U_{k-1}-G_k)} \mid \mathcal{F}_{k-1})\\
&= \sum_{m=1}^{U_{k-1}} \mathsf{E}_k\big((zw)^{\mathrm{Bin}(U_{k-1}-m)}w^m I(G_k=m) \mid \mathcal{F}_{k-1}\big)\\
&\qquad + w^{1+U_{k-1}}\,\mathsf{P}_k(G_k = U_{k-1}+1 \mid \mathcal{F}_{k-1})\\
&= \sum_{m=1}^{U_{k-1}} \Big(\frac{w}{2}\Big)^m \mathsf{E}_k((zw)^{\mathrm{Bin}(U_{k-1}-m)} \mid \mathcal{F}_{k-1}) + w\Big(\frac{w}{2}\Big)^{U_{k-1}}\\
&= \sum_{m=1}^{U_{k-1}} \Big(\frac{w}{2}\Big)^m \Big(\frac{zw+1}{2}\Big)^{U_{k-1}-m} + w\Big(\frac{w}{2}\Big)^{U_{k-1}}\\
&= \sum_{\ell=0}^{U_{k-1}-1} \Big(\frac{zw+1}{2}\Big)^{\ell} \Big(\frac{w}{2}\Big)^{U_{k-1}-\ell} + w\Big(\frac{w}{2}\Big)^{U_{k-1}}\\
&= \Big(\frac{w}{2}\Big)^{U_{k-1}} \left\{ \frac{1}{(1+zw)/w-1}\left(\Big(\frac{zw+1}{w}\Big)^{U_{k-1}} - 1\right) + w \right\}\\
&= w\left\{ \frac{1}{1+w(z-1)}\Big(\frac{zw+1}{2}\Big)^{U_{k-1}} + \left(1 - \frac{1}{1+w(z-1)}\right)\Big(\frac{w}{2}\Big)^{U_{k-1}} \right\},
\end{aligned}$$

which is the first formula.

To prove the second, write

$$\mathsf{E}_k\big(z^{S_k}w^{U_k}\big) = \mathsf{E}_k\,\mathsf{E}_k(z^{S_k}w^{U_k} \mid \mathcal{F}_{k-1}) = \mathsf{E}_k\big(z^{S_{k-1}}\,\mathsf{E}_k(z^{V_k}w^{U_k} \mid \mathcal{F}_{k-1})\big),$$

and use the first part and the usual reduction by (2.1) from P_k to P_{k-1}. □

For $\ell \geqslant 0$, set $b_\ell = b_\ell(z) = \sum_{j=0}^{\ell} z^j$ so that we have

$$zb_\ell + 1 = b_{\ell+1} \quad \text{and} \quad 1 + (z-1)b_\ell = z^{\ell+1}. \tag{5.3}$$

If we substitute b_ℓ for w in Lemma 5.1 and use (5.3), we obtain the following.

LEMMA 5.2. *For $0 \leqslant k \leqslant n-1$ and $\ell \geqslant 0$,*

$$\begin{aligned}
\mathsf{E}_{n-k}\big(z^{S_{n-k}}b_\ell^{U_{n-k}}\big) = \frac{b_\ell}{n-k}\Big(&z^{-\ell-1}\,\mathsf{E}_{n-k-1}\big(z^{S_{n-k-1}}b_{\ell+1}^{U_{n-k-1}}\big)\\
&+ (1-z^{-\ell-1})\,\mathsf{E}_{n-k-1}\big(z^{S_{n-k-1}}b_\ell^{U_{n-k-1}}\big)\Big).
\end{aligned}$$

PROPOSITION 5.3. *We have*

$$g_{S_n}(z) = \frac{1}{n!}\sum_{\ell=0}^{n-1} c_{n-1,\ell}(z)b_\ell(z),$$

where the coefficients $c_{m,\ell} = c_{m,\ell}(z)$ satisfy $c_{0,0} = 1$ and for $m \geqslant 0$ the recurrence

$$c_{m+1,\ell} = c_{m,\ell}(1-z^{-\ell-1})b_\ell(z) + c_{m,\ell-1}z^{-\ell}b_{\ell-1}(z).$$

In particular, $c_{m,\ell} = 0$ unless $0 \leqslant \ell \leqslant m$.

PROOF. Omitting the dependence of b_ℓ's on z for notational convenience, we show first that for $0 \leqslant k \leqslant n-1$ we have

$$\mathsf{E}_n\, z^{S_n} = \mathsf{E}_n\, z^{S_n} b_0^{U_n} = \frac{1}{n\cdots(n-k+1)} \sum_{m=0}^{k} c_{k,m}\, \mathsf{E}_{n-k}\, z^{S_{n-k}} b_m^{U_{n-k}}. \tag{5.4}$$

Indeed, assume inductively (5.4) for some $k \geqslant 0$ ($k=0$ being trivial). Apply the previous lemma to each of the terms $\mathsf{E}_{n-k}\, z^{S_{n-k}} b_m^{U_{n-k}}$ to get

$$\mathsf{E}_n\, z^{S_n} = \frac{1}{n\cdots(n-k)} \sum_{m=0}^{k} c_{k,m} b_m \Big\{ z^{-m-1}\, \mathsf{E}_{n-k-1}\, z^{S_{n-k-1}} b_{m+1}^{U_{n-k-1}} + (1-z^{-m-1})\, \mathsf{E}_{n-k-1}\, z^{S_{n-k-1}} b_m^{U_{n-k-1}} \Big\}.$$

Separating the sums, rearranging the terms, and collecting the coefficients in front of $\mathsf{E}_{n-k-1}\, z^{S_{n-k-1}} b_m^{U_{n-k-1}}$, we obtain that $\mathsf{E}_n\, z^{S_n}$ is equal to

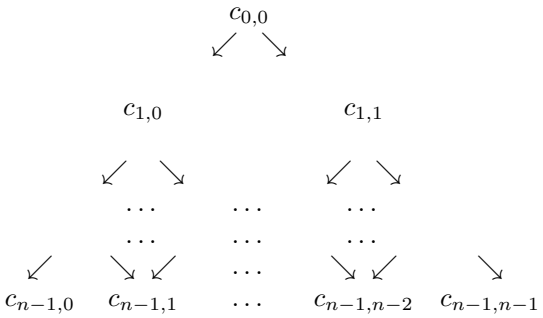
$$\frac{1}{n\cdots(n-k)} \Big\{ \sum_{m=0}^{k+1} \big(c_{k,m-1} b_{m-1} z^{-m} + c_{k,m} b_m (1-z^{-m-1}) \big)\, \mathsf{E}_{n-k-1}\, z^{S_{n-k-1}} b_m^{U_{n-k-1}} \Big\}.$$

This completes the inductive proof of (5.4) with the coefficients $c_{k,m}$ given by the specified recurrence. Choosing $k = n-1$ in (5.4), and using the observation that

$$\mathsf{E}_1\, z^{S_1} b_m^{U_1} = b_m$$

(because $S_1 = 0$ and $U_1 = 1$), completes the proof. □

REMARK 5.4. Just as in the proof of Theorem 4.2, the coefficients $c_{k,m}$ can be put in a Pascal-type triangle



This time a SE move from a coefficient $c_{m,\ell}$ has weight $z^{-\ell-1} b_\ell(z)$, and a SW move has weight $(1-z^{-\ell-1}) b_\ell(z)$. As was the case with the number of rows, the value of a given coefficient at the bottom is obtained by summing, over all possible paths leading to it from the root $c_{0,0}$, the products of weights corresponding to the moves along the path. This can be used to obtain an explicit expressions for the coefficients $(c_{n-1,m})$. Any path from $c_{0,0}$ to $c_{n-1,n-1-r}$, $0 \leqslant r \leqslant n-1$, has exactly r SW moves and $n-1-r$ SE moves. The total weight of SE moves (they are from $c_{\cdot,0}, c_{\cdot,1}, \dots, c_{\cdot,n-2-r}$) is

$$\prod_{j=0}^{n-2-r} z^{-j-1} b_j(z) = z^{-\binom{n-r}{2}} \prod_{j=0}^{n-r-2} b_j(z).$$

The SW moves may be from $c_{\cdot,\ell_1},\dots,c_{\cdot,\ell_r}$, for some $0\leqslant\ell_1\leqslant\ldots\leqslant\ell_r\leqslant n-1-r$. A SW move from $c_{\cdot,\ell_j}$ has weight $(1-z^{-\ell_j-1})b_{\ell_j}(z)$. Therefore,

$$c_{n-1,n-1-r}(z)=z^{-\binom{n-r}{2}}\prod_{k=0}^{n-r-2}b_k(z)\sum_{0\leqslant\ell_1\leqslant\ldots\leqslant\ell_r\leqslant n-1-r}\prod_{j=1}^{r}(1-z^{-\ell_j-1})b_{\ell_j}(z).$$

This gives an explicit expression for the probability generating function of S_n, namely

$$g_{S_n}(z)=\frac{1}{n!}\sum_{r=0}^{n-1}z^{-\binom{n-r}{2}}\left(\prod_{k=0}^{n-r-1}b_k(z)\right)A_{n,r}(z), \tag{5.5}$$

where

$$A_{n,r}(z)=\sum_{0\leqslant\ell_1\leqslant\ldots\leqslant\ell_r\leqslant n-1-r}\prod_{j=1}^{r}(1-z^{-\ell_j-1})b_{\ell_j}(z). \tag{5.6}$$

(Note that all negative powers cancel in (5.5) since $g_{S_n}(z)$ is a polynomial in z.)

Although the above expression looks quite complicated and it is not clear to us at the moment how to deduce the asymptotic normality of S_n from it, some information can be extracted from it. We shall illustrate this by deriving an exact expression for the variance of S_n.

PROPOSITION 5.5. *For $n\geqslant 2$, the variance of S_n satisfies*

$$\mathsf{var}(S_n)=\frac{(n-2)(2n^2+11n-1)}{360}. \tag{5.7}$$

PROOF. We compute the second factorial moment of S_n:

$$\mathsf{E}_n\, S_n(S_n-1)=\frac{1}{n!}\,\frac{\mathrm{d}^2}{\mathrm{d}z^2}\Big(\sum_{k=0}^{n-1}c_{n-1,k}b_k\Big)\Big|_{z=1}. \tag{5.8}$$

Notice that every path contributing to $c_{n-1,r}$, for $0\leqslant r<n-3$ has at least three SW moves, and so its weight will have at least three factors of the form $1-z^{-j}$. Hence if it is differentiated twice and evaluated at $z=1$, it will vanish. It follows that the sum on the right hand side of (5.8) reduces to the last three terms. For $k=n-3$ we have

$$\frac{1}{n!}c_{n-1,n-3}b_{n-3}=\frac{z^{-\binom{n-2}{2}}}{n!}\left(\prod_{k=0}^{n-4}b_k\right)b_{n-3}A_{n,2}(z)=G(z)A_{n,2}(z),$$

where we have set

$$G(z):=\frac{1}{n!}z^{-\binom{n-2}{2}}\prod_{k=0}^{n-3}b_k(z).$$

So, the second derivative of $G(z)A_{n,2}(z)$ is

$$G''(z)A_{n,2}(z)+2G'(z)A'_{n,2}(z)+G(z)A''_{n,2}(z).$$

Since

$$A_{n,2}(z)=\sum_{1\leqslant\ell\leqslant m\leqslant n-2}(1-z^{-\ell})(1-z^{-m})b_{\ell-1}(z)b_{m-1}(z),$$

$A_{n,2}(1) = A'_{n,2}(1) = 0$, so we only need $G(1)A''_{n,2}(1)$. Now,

$$G(1) = \frac{1}{n!}\prod_{k=0}^{n-3}(k+1) = \frac{1}{n(n-1)}. \tag{5.9}$$

To compute $A''_{n,2}(1)$, writing

$$(1-z^{-\ell})(1-z^{-m}) = h_1(z), \quad b_{\ell-1}(z)b_{m-1}(z) = h_2(z),$$

we see that

$$h_1''(1)h_2(1) + 2h_1'(1)h_2'(1) + h_1(1)h_2''(1) = h_1''(1)h_2(1) = 2\ell^2 m^2.$$

Therefore,

$$\begin{aligned} A''_{n,2}(1) &= 2\sum_{m=1}^{n-2} m^2 \sum_{\ell=1}^{m} \ell^2 = \frac{1}{3}\sum_{m=1}^{n-2} m^3(m+1)(2m+1) \\ &= \frac{1}{180}n(n-1)(n-2)(5n-11)(2n-1)(2n-3), \end{aligned}$$

and combining this with (5.9) we obtain

$$\frac{1}{n!}\frac{\mathrm{d}^2}{\mathrm{d}z^2}\left(c_{n-1,n-3}b_{n-3}\right)\Big|_{z=1} = \frac{(n-2)(5n-11)(2n-1)(2n-3)}{180}.$$

We next handle

$$\begin{aligned} \frac{1}{n!}c_{n-1,n-1}b_{n-1} &= \frac{z^{-\binom{n}{2}}}{n!}\left(\prod_{k=0}^{n-2} b_k\right) b_{n-1} \\ &= z^{-\binom{n}{2}}\prod_{j=1}^{n-1}\frac{b_{j-1}}{j}. \end{aligned}$$

The last expression is the probability generating function of a sum of independent random variables $W_0, \dots, W_n$, where $W_0 \equiv -n(n-1)/2$ and for $1 \leqslant j \leqslant n$, W_j is a discrete uniform random variable on $\{0, \dots, j-1\}$. So, with $Y_n = \sum_{j=0}^{n} W_j$ we have

$$\frac{1}{n!}\frac{\mathrm{d}^2}{\mathrm{d}z^2}c_{n-1,n-1}b_{n-1}\Big|_{z=1} = \mathsf{E}\, Y_n(Y_n-1) = \sum_{j=1}^{n}\mathsf{var}(W_j) + (\mathsf{E}\, Y_n)^2 - \mathsf{E}\, Y_n.$$

Now,

$$\mathsf{E}\, W_j = \frac{j-1}{2}, \qquad 1 \leqslant j \leqslant n,$$

so that

$$\mathsf{E}\, Y_n = \frac{1}{2}\sum_{j=1}^{n}(j-1) - \frac{n(n-1)}{2} = -\frac{n(n-1)}{4}.$$

Furthermore, for $1 \leqslant j \leqslant n$

$$\mathsf{var}(W_j) = \frac{1}{j}\sum_{k=0}^{j-1} k^2 - \left(\frac{j-1}{2}\right)^2 = \frac{(j-1)(j+1)}{12}.$$

Therefore,

$$\sum_{j=1}^{n}\mathsf{var}(W_j) = \frac{n(n-1)(2n+5)}{72}.$$

Finally, putting the above together we get

$$\frac{1}{n!}\frac{\mathrm{d}^2}{\mathrm{d}z^2}c_{n-1,n-1}b_{n-1}\Big|_{z=1} = \frac{n(n-1)(2n+5)}{72} + \frac{n^2(n-1)^2}{16} + \frac{n(n-1)}{4}$$
$$= \frac{n(n-1)(9n^2-5n+46)}{144}.$$

It remains to handle $\frac{1}{n!}c_{n-1,n-2}b_{n-2}$. We write

$$c_{n-1,n-2}b_{n-2} = f(z)A_{n,1}(z),$$

where

$$f(z) = z^{-\binom{n-1}{2}}\prod_{k=0}^{n-2} b_k(z)$$

and, according to (5.6),

$$A_{n,1}(z) = \sum_{\ell=0}^{n-2}(1-z^{-\ell-1})b_\ell(z).$$

Then the second derivative of $c_{n-1,n-2}b_{n-2}$ is

$$f''(z)A_{n,1}(z) + 2f'(z)A'_{n,1}(z) + f(z)A''_{n,1}(z).$$

At $z=1$ the first product vanishes. For the remaining two, first notice that

$$\frac{1}{(n-1)!}f(z) = z^{-\binom{n-1}{2}}\prod_{j=1}^{n-1}\frac{b_{j-1}}{j}$$

is a generating function of a legitimate distribution function. Therefore, $f(1) = (n-1)!$ and $f'(1)/(n-1)!$ is the expected value of a random variable represented by $f(z)/(n-1)!$. Since this expected value is

$$-\binom{n-1}{2} + \sum_{j=1}^{n-1}\frac{1}{j}\sum_{k=0}^{j-1}k = -\frac{(n-1)(n-2)}{4},$$

we obtain

$$f'(1) = -\frac{(n-1)(n-2)(n-1)!}{4}.$$

It remains to compute the first two derivatives of $A_{n,1}(z)$ at $z=1$.

$$A'_{n,1}(z) = \sum_{\ell=1}^{n-1}\left(b'_{\ell-1}(1-z^{-\ell}) + b_{\ell-1}\ell z^{-\ell-1}\right),$$

so that

$$A'_{n,1}(1) = \sum_{\ell=1}^{n-1}\ell^2 = \frac{n(n-1)(2n-1)}{6}. \tag{5.10}$$

Also

$$A''_{n,1}(z) = \sum_{\ell=1}^{n-1}\left(b''_{\ell-1}(1-z^{-\ell}) + 2b'_{\ell-1}\ell z^{-\ell-1} - b_{\ell-1}\ell(\ell+1)z^{-\ell-2}\right).$$

Since $b'_{\ell-1}(1) = (\ell-1)\ell/2$ we get

$$A''_{n,1}(1) = \sum_{\ell=1}^{n-1}\left((\ell-1)\ell^2 - \ell^2(\ell+1)\right) = -\frac{n(n-1)(2n-1)}{3}.$$

Hence,

$$\begin{aligned}\frac{1}{n!}\frac{\mathrm{d}^2}{\mathrm{d}z^2}c_{n-1,n-2}b_{n-2}\Big|_{z=1} &= \frac{1}{n!}\big(2f'(1)A'_{n,1}(1)+f(1)A''_{n,1}(1)\big)\\ &= -\frac{(n-1)^2(n-2)(2n-1)}{12} - \frac{(n-1)(2n-1)}{3}\\ &= -\frac{(n-1)(2n-1)(n^2-3n+6)}{12}.\end{aligned}$$

Combining all of these calculations gives

$$\begin{aligned}\mathsf{E}_n\, S_n(S_n-1) &= \frac{(n-2)(5n-11)(2n-1)(2n-3)}{180}\\ &\quad - \frac{(n-1)(2n-1)(n^2-3n+6)}{12}\\ &\quad + \frac{n(n-1)(9n^2-5n+46)}{144}\\ &= \frac{(n-2)(n-3)(5n^2-n-16)}{720}.\end{aligned}$$

By the same argument

$$\begin{aligned}\mathsf{E}_n\, S_n &= \frac{1}{n!}\frac{\mathrm{d}}{\mathrm{d}z}\Big(c_{n-1,n-1}b_{n-1}+c_{n-1,n-2}b_{n-2}\Big)\Big|_{z=1}\\ &= \mathsf{E}\, Y_n + \frac{1}{n!}\left(f'(1)A_{n,1}(1)+f(1)A'_{n,1}(1)\right).\end{aligned}$$

Using $A_{n,1}(1)=0$, $f(1)=(n-1)!$, (5.10), and the value of $\mathsf{E}\, Y_n$, we get

$$\mathsf{E}_n\, S_n = -\frac{n(n-1)}{4}+\frac{(n-1)(2n-1)}{6} = \frac{(n-1)(n-2)}{12},$$

which conforms to (5.1). Finally,

$$\begin{aligned}\mathsf{var}(S_n) &= \mathsf{E}\, S_n(S_n-1) - (\mathsf{E}\, S_n)^2 + \mathsf{E}\, S_n\\ &= \frac{(n-2)(n-3)(5n^2-n-16)}{720} - \left(\frac{(n-1)(n-2)}{12}\right)^2 + \frac{(n-1)(n-2)}{12}\\ &= \frac{(n-2)(2n^2+11n-1)}{360},\end{aligned}$$

which proves Proposition 5.5. □

6. Asymptotic normality of S_n

In this section we provide a self-contained proof of

THEOREM 6.1. *As $n\to\infty$ we have*

$$\frac{S_n - n^2/12}{\sqrt{n^3/180}} \xrightarrow{\mathrm{d}} N(0,1). \tag{6.1}$$

PROOF. As we mentioned, the form of the probability generating function of S_n obtained in the previous section does not seem to be convenient to yield the central limit theorem. For this reason, we shall rely on a bijective result of Steingrímsson and Williams [**SW**]. According to their result, the number of superfluous 1s in a permutation tableau of length n is equidistributed with the number of occurrences of the generalized pattern $31{-}2$ in a random permutation of $[n]$. (An occurrence

of the generalized pattern $31{-}2$ in a permutation σ is a pair $1 < i < j$ such that $\sigma_{i-1} > \sigma_j > \sigma_i$.) To analyze that quantity, it will be convenient to think of a random permutation as being generated from a sample of i.i.d. random variables $X_1, \dots, X_n$, each uniform on $[0,1]$. (The permutation is obtained by reading off the ranks of $X_1, \dots, X_n$.) If for $2 \leqslant i < j \leqslant n$ we let $I_{i,j} := I_{X_{i-1} > X_j > X_i}$, then

$$S_n = \sum_{2 \leqslant i < j \leqslant n} I_{i,j}.$$

Notice that from this representation we immediately recover (5.1), since

$$\mathsf{E}\, S_n = \mathsf{E} \sum_{2 \leqslant i < j \leqslant n} I_{i,j} = \binom{n-1}{2} \mathsf{P}(X_1 > X_3 > X_2) = \frac{1}{6}\binom{n-1}{2}.$$

Similarly, we can easily obtain the asymptotic value of the variance: we write

$$\mathsf{var}(S_n) = \sum_{\substack{i_1 < j_1 \\ i_2 < j_2}} \mathsf{cov}(I_{i_1,j_1}, I_{i_2,j_2}),$$

and note that if $\{i_1 - 1, i_1, j_1\} \cap \{i_2 - 1, i_2, j_2\} = \varnothing$ then I_{i_1,j_1} and I_{i_2,j_2} are independent, and so their covariance vanishes. In the complementary case, the main contribution comes from the cases that contribute $\Theta(n^3)$ terms to the sum. We obtain

$$\begin{aligned}\mathsf{var}(S_n) \sim \frac{n^3}{3}\Big(&\mathsf{cov}(I_{2,3}, I_{2,4}) + \mathsf{cov}(I_{2,5}, I_{4,5}) + \mathsf{cov}(I_{2,4}, I_{3,5}) \\ &+ \mathsf{cov}(I_{2,5}, I_{3,4}) + \mathsf{cov}(I_{2,3}, I_{4,5}) + \mathsf{cov}(I_{2,4}, I_{4,5})\Big),\end{aligned}$$

as all other cases contribute $O(n^2)$ terms to the sum. We calculate:

$$\mathsf{E}\, I_{2,3} \cap I_{2,4} = \mathsf{P}(X_1 > X_3 > X_2, X_1 > X_4 > X_2) = 2\,\mathsf{P}(X_1 > X_3 > X_4 > X_2) = \frac{2}{4!},$$

so that

$$\mathsf{cov}(I_{2,3}, I_{2,4}) = \frac{1}{12} - \left(\frac{1}{6}\right)^2.$$

Similarly,

$$\begin{aligned}
\mathsf{E}\, I_{2,5} \cap I_{4,5} &= \mathsf{P}(X_1 > X_5 > X_2, X_3 > X_5 > X_4) \\
&= 4\,\mathsf{P}(X_1 > X_3 > X_5 > X_2 > X_4) \quad = \frac{1}{30}, \\
\mathsf{E}\, I_{2,4} \cap I_{3,5} &= \mathsf{P}(X_1 > X_4 > X_2, X_2 > X_5 > X_3) = \frac{1}{120}, \\
\mathsf{E}\, I_{2,5} \cap I_{3,4} &= \mathsf{P}(X_1 > X_5 > X_2, X_2 > X_4 > X_3) = \frac{1}{120}, \\
\mathsf{E}\, I_{2,3} \cap I_{4,5} &= \mathsf{P}(X_1 > X_3 > X_2, X_3 > X_5 > X_4) \\
&= 3\,\mathsf{P}(X_1 > X_3 > X_2 > X_5 > X_4) \quad = \frac{1}{40}, \\
\mathsf{E}\, I_{2,4} \cap I_{4,5} &= \mathsf{P}(X_1 > X_4 > X_2, X_3 > X_5 > X_4) = \frac{1}{40}.
\end{aligned}$$

Hence,

$$\mathsf{var}(S_n) \sim \frac{n^3}{3} \cdot \Big(\frac{1}{12} + \frac{1}{30} + \frac{1}{120} + \frac{1}{120} + \frac{1}{40} + \frac{1}{40} - \frac{6}{36}\Big) = \frac{n^3}{180}. \tag{6.2}$$

REMARK 6.2. The exact value of the variance could be obtained by computing the other terms.

Finally, to establish (6.1) we shall rely on results presented in [**J**] and [**JŁR**, Section 6.1]; see [**Bó**] for a closely related theorem proved by the same method. See also [**E**] for a related simple proof of the asymptotic normality of the number of descents, cf. Corollary 4.3 above. We let $A = A_n := \{\alpha = (i,j) : \ 2 \leqslant i < j \leqslant n\}$. Then $S_n = \sum_{\alpha\in A} I_\alpha$. Recall that a dependency graph L for $\{I_\alpha : \alpha \in A\}$ is any graph whose vertex set is A and which has the property that if V_1, V_2 are two disjoint subsets of A such that L has no edges with one endpoint in V_1 and the other in V_2, then the families $\{I_\alpha : \ \alpha \in V_1\}$ and $\{I_\alpha : \ \alpha \in V_2\}$ are mutually independent. For our purposes it is enough to consider L defined by the following rule: we put an edge between $\alpha_1 = (i_1, j_1)$ and $\alpha_2 = (i_2, j_2)$ iff $\{i_1 - 1, i_1, j_1\} \cap \{i_2 - 1, i_2, j_2\} \neq \varnothing$. If $\alpha_1, \dots, \alpha_r \in A$ then the closed neighborhood of $\{\alpha_1, \dots, \alpha_r\}$ in L is defined by

$$\bar{N}_L(\alpha_1, \dots, \alpha_r) = \bigcup_{i=1}^{r} \{\beta \in A : \ \beta = \alpha_i \text{ or } \alpha_i\beta \in E(L)\},$$

where $E(L)$ denotes the edge set of L. Note that in our case for every fixed $r \geqslant 2$ $\left|\bar{N}_L(\alpha_1, \dots, \alpha_{r-1})\right| = O_r(n)$, where $O_r(\ \cdot\ )$ means that the constant may depend on r. Hence, trivially

$$\sum_{\alpha\in\bar{N}_L(\alpha_1,\dots,\alpha_{r-1})} \mathsf{E}\big(I_\alpha \big| I_{\alpha_1}, \dots, I_{\alpha_{r-1}}\big) = O_r(n).$$

Since, cf. (5.1) for the exact value,

$$\sum_{\alpha\in A} \mathsf{E}\, I_\alpha \leqslant |A| = O(n^2),$$

by [**JŁR**, Lemma 6.17] we conclude that the rth cumulant of S_n (defined by $\kappa_r(S_n) = \mathrm{i}^{-k}\frac{\mathrm{d}^k}{\mathrm{d}t^k}\log\phi_{S_n}(0)$, where $\phi_X(t) = g_X(e^{\mathrm{i}t})$ is the characteristic function) satisfies

$$|\kappa_r(S_n)| = O_r(n^2 \cdot n^{r-1}) = O_r(n^{r+1}).$$

Hence, for $r \geqslant 3$,

$$\left|\kappa_r\left(\frac{S_n - n^2/12}{\sqrt{n^3/180}}\right)\right| = O_r(n^{r+1}n^{-\frac{3}{2}r}) = o(1)$$

as $n \to \infty$. Since $\kappa_1(X) = \mathsf{E}\, X$ and $\kappa_2(X) = \mathsf{var}(X)$ we have

$$\kappa_1\left(\frac{S_n - n^2/12}{\sqrt{n^3/180}}\right) \longrightarrow 0, \qquad \kappa_2\left(\frac{S_n - n^2/12}{\sqrt{n^3/180}}\right) \longrightarrow 1,$$

and the theorem follows by the cumulant convergence theorem (see, e.g., [**JŁR**, Corollary 6.15]). □

REMARK 6.3. Our results can be used to draw conclusions about some other parameters. For example, let Y_n be the number of 1s in the random permutation tableau T_n. Although we have not computed an explicit formula for the distribution of Y_n, we can easily obtain its asymptotic distribution. Namely, as $n \to \infty$ we have

$$\frac{Y_n - n^2/12}{\sqrt{n^3/180}} \xrightarrow{\mathrm{d}} N(0,1).$$

This follows immediately from Theorem 6.1 upon noting that

$$Y_n = S_n + C_n = S_n + O(n).$$

It seems straightforward to prove a similar central limit theorem for Z_n, the number of 0s in the random permutation tableau T_n, using further bijective results of [**SW**]; the main difference is that we need to consider several generalized patterns simultaneously, and the joint distribution of their numbers of occurrences. We leave this to the reader.

References

[BaS] E. Babson and E. Steingrímsson, *Generalized permutation patterns and a classification of the Mahonian statistics*, Sém. Lothar. Combin. **44** (2000), article B44b, 18 pp.

[Bó] M. Bóna, The copies of any permutation pattern are asymptotically normal, preprint, 2007. Available as arXiv:math.CO/0712.2792.

[BrCPRE] R. Brak, S. Corteel, R. Parviainen, A. Rechnitzer, and J. Essam, *A combinatorial derivation of the PASEP stationary state*, Electron. J. Combin. **13** (2006), no. 1, research paper 108, 23 pp.

[BrE] R. Brak and J. Essam, *Asymmetric exclusion model and weighted lattice paths*, J. Phys. A **37** (2004), no. 14, 4183–4217.

[Bu] A. Burstein, *On some properties of permutation tableaux*, Ann. Comb. **11** (2007), no. 3–4, 355–368.

[CKSS] L. Carlitz, D. S. Kurtz, R. Scoville, and O. P. Stackelberg, *Asymptotic properties of Eulerian numbers*, Z. Wahrscheinlichkeitstheorie und verw. Gebiete **23** (1972), 47–54.

[Cl] A. Claesson, *Generalized pattern avoidance*, European J. Combin. **22** (2001), no. 7, 961–971.

[ClM] A. Claesson and T. Mansour, *Enumerating permutations avoiding a pair of Babson–Steingrímsson patterns*, Ars Combin. **77** (2005), 17–31.

[Co] S. Corteel, *Crossings and alignments of permutations*, Adv. in Appl. Math. **38** (2007), no. 2, 149–163.

[CoH] S. Corteel and P. Hitczenko, *Expected values of statistics on permutation tableaux*, pp. 325–339 in *2007 Conference on Analysis of Algorithms, AofA 07*, Assoc. Discrete Mathematics and Theoretical Computer Science (DMTCS), Nancy, France, 2007.

[CoN] S. Corteel and P. Nadeau, *Bijections for permutation tableaux*, European J. Combin. **30** (2009), no. 1, 295–310.

[CoW] S. Corteel and L. K. Williams, *Tableaux combinatorics for the asymmetric exclusion process*, Adv. in Appl. Math. **39** (2007), no. 3, 293–310.

[CoW1] S. Corteel and L. K. Williams, *A Markov chain on permutations which projects to the PASEP*, Int. Math. Res. Not. IMRN (2007), no. 17, article ID rnm055, 27 pp.

[DB] F. N. David and D. E. Barton, *Combinatorial Chance*, Hafner Publishing Co., New York, 1962.

[DDM] B. Derrida, E. Domany, and D. Mukamel, *An exact solution of a one-dimensional asymmetric exclusion model with open boundaries*, J. Statist. Phys. **69** (1992), no. 3–4, 667–687.

[DEHP] B. Derrida, M. Evans, V. Hakim, V. Pasquier, *Exact solution of a 1D asymmetric exclusion model using a matrix formulation*, J. Phys. A **26** (1993), no. 7, 1493–1517.

[DS] E. Duchi and G. Schaeffer, *A combinatorial approach to jumping particles*, J. Combin. Theory Ser. A **110** (2005), no. 1, 1–29.

[E] C.-G. Esseen, *On the application of the theory of probability to two combinatorial problems involving permutations*, pp. 137–147 in *Proceedings of the Seventh Conference on Probability Theory (Braşov, 1982)*, VNU Sci. Press, Utrecht, 1985.

[F] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. I, Wiley, New York, 3rd edition, 1968.

[J] S. Janson, *Normal convergence by higher semi-invariants with applications to sums of dependent random variables and random graphs*, Ann. Probab. **16** (1988), no. 1, 305–312.

[JŁR] S. Janson, T. Łuczak, and A. Rucinski, *Random Graphs*, Wiley, New York, 2000.
[K] D. N. Knuth, *The Art of Computer Programming*, vol. 3, Addison–Wesley, Reading, MA, 1996.
[P] A. Postnikov, *Total positivity, Grassmannians, and networks*, preprint, 2006. Available as arXiv:math.CO/0609764.
[R] J. Riordan, *An Introduction to Combinatorial Analysis*, Wiley, New York, 1958.
[S] T. Sasamoto, *One-dimensional partially asymmetric simple exclusion process with open boundaries: Orthogonal polynomials approach*, J. Phys. A **32** (1999), no. 41, 7109–7131.
[SW] E. Steingrímsson and L. K. Williams, *Permutation tableaux and permutation patterns*, J. Combin. Theory Ser. A **114** (2007), no. 2, 211–234.
[USW] M. Uchiyama, T. Sasamoto, and M. Wadati, *Asymmetric simple exclusion process with open boundaries and Askey–Wilson polynomials*, J. Phys. A **37** (2004), no. 18, 4985–5002.
[W] L. K. Williams, *Enumeration of totally positive Grassmann cells*, Adv. in Math. **190** (2005), no. 2, 319–342.

Department of Mathematics, Drexel University, Philadelphia, Pennsylvania 19104, USA

E-mail address: phitczenko@math.drexel.edu
URL: http://www.math.drexel.edu/~phitczen

Department of Mathematics, Uppsala University, Box 480, SE-751 06 Uppsala, Sweden

E-mail address: svante.janson@math.uu.se
URL: http://www.math.uu.se/staff/pages/?uname=svante

Contemporary Mathematics
Volume **520**, 2010

Rotor walks and Markov chains

Alexander E. Holroyd and James Propp

Dedicated to Oded Schramm, 1961–2008

Abstract. The rotor walk is a derandomized version of the random walk on a graph. On successive visits to any given vertex, the walker is routed to each of the neighboring vertices in some fixed cyclic order, rather than to a random sequence of neighbors. The concept generalizes naturally to countable Markov chains. Subject to general conditions, we prove that many natural quantities associated with the rotor walk (including normalized hitting frequencies, hitting times and occupation frequencies) concentrate around their expected values for the random walk. Furthermore, the concentration is stronger than that associated with repeated runs of the random walk; the discrepancy is at most C/n after n runs (for an explicit constant C), rather than $c/\sqrt{n}$.

1. Introduction

Let $X_0, X_1, \ldots$ be a Markov chain on a countable set V with transition probabilities $p\colon V \times V \to [0,1]$ (see, e.g., [**20**] for background). We call the elements of V *vertices*. We write $\mathbb{P}_u$ for the law of the Markov chain started at vertex u (so $\mathbb{P}_u$-a.s. we have $X_0 = u$).

The *rotor-router walk* or *rotor walk* is a deterministic cellular automaton associated with the Markov chain, defined as follows. Assume that all transition probabilities $p(u,v)$ are rational (later we shall address relaxation of this assumption) and that for each u there are only finitely many v such that $p(u,v) > 0$. To each vertex u we associate a positive integer $d(u)$ and a finite sequence of (not necessarily distinct) vertices $u^{(1)}, \ldots, u^{(d(u))}$, called the *successors* of u, in such a way that

$$p(u,v) = \frac{\#\{i : u^{(i)} = v\}}{d(u)} \qquad \text{for all} \quad u, v \in V. \tag{1.1}$$

2000 *Mathematics Subject Classification.* Primary 82C20; Secondary 20K01, 05C25.

Key words and phrases. Rotor router, quasirandom, derandomized random walk, Eulerian walkers, Markov chain, discrepancy.

AEH was funded in part by Microsoft Research and an NSERC Grant.

JP was funded in part by NSF grant no. DMS-0644877.

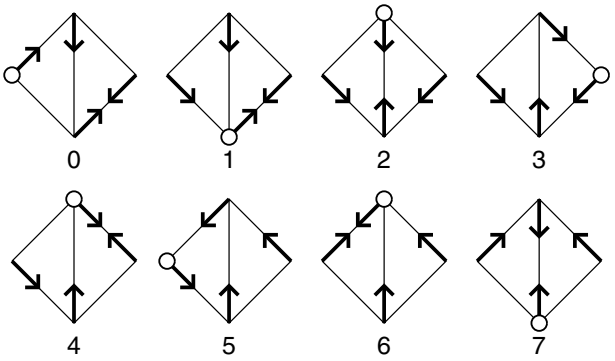


FIGURE 1. Steps $0, \dots, 7$ of a rotor walk associated with the simple random walk on a graph with 4 vertices. The thin lines represent the edges of the graph, the circle is the particle location, and the thick arrows are the rotors. The rotor mechanism in this case is such that each rotor successively points to the vertex's neighbors in anticlockwise order.

(This is clearly possible under the given assumptions; $d(u)$ may be taken to be the lowest common denominator of the transition probabilities from u.) The set V together with the quantities $d(u)$ and the assignments of successors will sometimes be called the *rotor mechanism*.

A *rotor configuration* is a map r that assigns to each vertex v an integer $r(v) \in \{1, \dots, d(v)\}$. (We think of an arrow or *rotor* located at each vertex, with the rotor at v pointing to vertex $v^{(r(v))}$). We let a rotor configuration evolve in time, in conjunction with the position of a particle moving from vertex to vertex: the rotor at the current location v of the particle is incremented, and the particle then moves in the new rotor direction. More formally, given a rotor mechanism, an initial particle location $x_0 \in V$ and an initial rotor configuration r_0, the *rotor walk* is a sequence of vertices $x_0, x_1, \dots \in V$ (called *particle locations*) together with rotor configurations $r_0, r_1, \dots$, constructed inductively as follows. Given x_t and r_t at time t we set:

(i) $r_{t+1}(v) := \begin{cases} (r_t(v)+1) \bmod d(v), & v = x_t; \\ r_t(v), & v \neq x_t \end{cases}$
(i.e., increment the rotor at the current particle location); and

(ii) $x_{t+1} := (x_t)^{(r_{t+1}(x_t))}$
(i.e., move the particle in the new rotor direction).

See Figure 1 for a simple illustration of the mechanism.

Given a rotor walk, write

$$n_t(v) := \#\big\{s \in [0, t-1] : x_s = v\big\}$$

for the number of times the particle visits vertex v before (but not including) time t.

We next state general results, Theorems 1–4, relating basic Markov chain objects to their rotor walk analogues (under suitable conditions). We then state a more refined result (Theorem 5) for the important special case of simple random walk on $\mathbb{Z}^2$, followed by extensions to infinite times (Theorem 8) and irrational

transition probabilities (Theorem 12). We postpone discussion of history and background to the end of the introduction, and proofs to the later sections.

1.1. Hitting probabilities. Let $T_v := \min\{t \geqslant 0 : X_t = v\}$ be the first hitting time of vertex v by the Markov chain (where $\min \varnothing := \infty$). Fix two distinct vertices b, c and consider the hitting probability

$$h(v) = h_{b,c}(v) := \mathbb{P}_v(T_b < T_c). \tag{1.2}$$

Note that $h(b) = 1$ and $h(c) = 0$. In order to connect hitting probabilities with rotor walks, fix a starting vertex $a \notin \{b, c\}$, and modify the transition probabilities from b and c so that $p(b, a) = p(c, a) = 1$. (Thus, after hitting b or c, the particle is returned to a.) Note that this modification does not change the function h. Modify the rotor mechanism accordingly by setting $d(b) = d(c) = 1$ and $b^{(1)} = c^{(1)} = a$. Let $x_0, x_1, \ldots$ be a rotor walk associated with the modified chain. The following is our most basic result.

THEOREM 1 (Hitting probabilities). *Under the above assumptions, suppose that the quantity*

$$K_1 := 1 + \frac{1}{2} \sum_{\substack{u \in V \setminus \{b,c\}, \\ v \in V}} d(u)p(u, v)\,|h(u) - h(v)|$$

is finite. Then for any rotor walk and all t,

$$\left| h(a) - \frac{n_t(b)}{n_t(b) + n_t(c)} \right| \leqslant \frac{K_1}{n_t(b) + n_t(c)}.$$

Theorem 1 implies that the proportion of times that the rotor walk hits b as opposed to c converges to the Markov chain hitting probability $h(a)$, provided the rotor walk hits $\{b, c\}$ infinitely often (we shall consider cases where this does not hold in the later discussion on transfinite rotor walks). Furthermore, after n visits to $\{b, c\}$, the discrepancy in this convergence is at most K/n for a fixed constant K. In contrast, for the proportion of visits by the Markov chain itself, the discrepancy is asymptotically a random multiple of $1/\sqrt{n}$ (by the central limit theorem).

The condition $K_1 < \infty$ holds in particular whenever V is finite, as well as in many cases when it is infinite; for examples see [**22**].

In the case when the Markov chain (before modification) is a simple random walk on an undirected graph $G = (V, E)$ (thus $p(u, v)$ equals $1/d(u)$ if (u, v) is an edge, and 0 otherwise, with $d(u)$ being the degree of u), we obtain the particularly simple bound $K_1 \leqslant 1 + \sum_{(u,v)\in E} |h(u) - h(v)|$.

Theorem 1 can be easily adapted to give similar results for the probability of *returning* to b before hitting c when started at $a = b$, and for the probability of hitting one *set* of vertices before another. This can be done either by adapting the proof or by adding appropriate extra vertices and then appealing to Theorem 1. For brevity we omit such variations.

We next discuss extensions of Theorem 1 in the following directions: hitting times and stationary distributions, an example where $K_1 = \infty$, cases where the particle can escape to infinity, and irrational transition probabilities.

1.2. Hitting times. Fix a vertex b and let

$$k(v) = k_b(v) := \mathbb{E}_v T_b \tag{1.3}$$

be its expected hitting time. Fix also an initial vertex $a \neq b$ and modify the transition probabilities from b so that $p(b,a) = 1$. (Then $k(a)$ is also the expected return time from b to b in the reduced chain in which the vertices a and b are conflated.) Let $x_0, x_1, \ldots$ be a rotor walk associated with the modified chain.

THEOREM 2 (Hitting times). *Under the above assumptions, suppose that V is finite, and let*

$$K_2 := \max_{v\in V} k(v) + \frac{1}{2} \sum_{\substack{u\in V\setminus\{b\},\\ v\in V}} d(u)p(u,v)\,|k(u) - k(v) - 1|\,.$$

Then for any rotor walk and all t,

$$\left|(k(a)+1) - \frac{t}{n_t(b)}\right| \leqslant \frac{K_2}{n_t(b)}.$$

Thus the average time for the rotor walk to get from a to b concentrates around the expected hitting time. The "+1" term corresponds to the time step to move from b to a.

Note that, in contrast with Theorem 1, in the above result we require V to be finite. Leaving aside some degenerate cases, such a bound cannot hold when V is infinite. Indeed, if V is infinite and the Markov chain is irreducible, then $|(k(a)+1)n_t(b) - t|$ is unbounded in t, since the rotor walk has arbitrarily long excursions between successive visits to b; hence the conclusion of Theorem 2 cannot hold (for any constant K_2) in this case. In contrast, in the next result we again allow V to be infinite.

1.3. Stationary vectors. Suppose that the Markov chain is irreducible and recurrent, and let $\pi\colon V \to (0,\infty)$ be a stationary vector (so that $\pi p = \pi$ as a matrix product). Let $x_0, x_1, \ldots$ be an associated rotor walk. Fix two vertices $b \neq c$ and let $h = h_{b,c}$ be as in (1.2) above. Also let $T_u^+ := \min\{t \geqslant 1 : X_t = u\}$ denote the first return time to u, and define the escape probability $e_{u,v} := \mathbb{P}_u(T_v < T_u^+)$.

THEOREM 3 (Occupation frequencies). *For any irreducible, recurrent Markov chain, with the above notation, suppose that the quantity*

$$K_3 := 1 + \frac{1}{2}\Big(d(b) + d(c) + \sum_{u,v\in V} d(u)p(u,v)\,|h(u) - h(v)|\Big)$$

is finite. Then for all t,

$$\left|\frac{n_t(b)}{\pi(b)} - \frac{n_t(c)}{\pi(c)}\right| \leqslant \frac{K_3}{\pi(b)e_{b,c}}.$$

Thus, the ratio of times spent at different vertices by the rotor walk concentrates around the ratio of corresponding components of the stationary vector.

Now suppose that the Markov chain is irreducible and positive recurrent, and let π be the stationary *distribution* (so that $\sum_{v\in V}\pi(v) = 1$). Fix a vertex b and let $k = k_b$ be as in (1.3). The following result states that the proportion of time spent by the rotor walk at b concentrates around $\pi(b)$.

THEOREM 4 (Stationary distribution). *For an irreducible positive recurrent Markov chain with V finite, with the above notation, let*

$$K_4 := \max_{v\in V} k(v) + \frac{1}{2}\left(\frac{d(b)}{\pi(b)} + \sum_{u,v\in V} d(u)p(u,v)\,|k(u)-k(v)-1|\right).$$

Then for all t,

$$\left|\pi(b) - \frac{n_t(b)}{t}\right| \leqslant \frac{K_4\pi(b)}{t}.$$

1.4. Logarithmic discrepancy for walks on $\mathbb{Z}^2$. While Theorem 1 requires the quantity K_1 to be finite, experiments suggest that similar conclusions hold in many cases where it is infinite. We next treat one interesting example in which such a conclusion provably holds, but with an additional logarithmic factor in the bound on the discrepancy. (Additional such examples will appear in [**22**].)

Consider simple symmetric random walk on the square lattice $\mathbb{Z}^2$. That is, let $V = \mathbb{Z}^2$, and let $p(u,v) := 1/4$ for all $u, v \in V$ with $\|u-v\|_1 = 1$ and $p(u,v) := 0$ otherwise. Let each rotor rotate anticlockwise; that is for each $u \in V$, we set $d(u) := 4$ and

$$u^{(i)} := u + (\cos\tfrac{i\pi}{2}, \sin\tfrac{i\pi}{2}), \qquad i = 1,\dots,4. \tag{1.4}$$

Consider the particular initial rotor configuration r given by

$$r\big((x,y)\big) := \lfloor \tfrac{1}{2} + \tfrac{2}{\pi}\arg(x-\tfrac{1}{2}, y-\tfrac{1}{2})\rfloor \bmod 4 \tag{1.5}$$

(where $\arg(x,y)$ denotes the angle $\theta \in [0,2\pi)$ such that $(x,y) = r(\cos\theta, \sin\theta)$ with $r > 0$). See Figure 2.

Fix vertices a, b, c of $\mathbb{Z}^2$ with $b \neq c$ and modify p by setting $p(b,a) = p(c,a) = 1$. If $a = b$ then also split this vertex into two vertices a and b, let b inherit all the incoming transition probabilities of the original random walk, and let a inherit the outgoing probabilities; similarly if $a = c$. Also modify the rotor mechanism and the rotor configuration r accordingly.

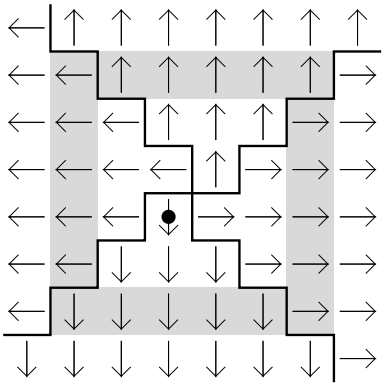


FIGURE 2. The initial rotor configuration in Theorem 5. The dot shows the location of $(0,0)$. The third layer is shaded (see the later proofs).

THEOREM 5 (Hitting probabilities for walk on $\mathbb{Z}^2$). *Let a, b, c be vertices of $\mathbb{Z}^2$ with $b \neq c$, and consider the rotor walk associated with the random walk, rotor mechanism and initial rotor configuration described above, started at vertex a. Then for any t, with $h(a) = h_{b,c}(a)$ and $n = n_t(b) + n_t(c)$,*

$$\left| h(a) - \frac{n_t(b)}{n} \right| \leqslant \frac{C \ln n}{n}.$$

Furthermore, $t \leqslant C' n^3$. Here C, C' are finite constants depending on a, b, c.

In contrast to the above result for the rotor walk, for the Markov chain itself, after n visits to $\{b, c\}$ the proportion of visits to b differs from its limit $h(a)$ by $K/\sqrt{n}$ in expected absolute value (by the central limit theorem), while the median number of time steps needed to achieve n visits is at least $(K')^n$ where $K > 0$ and $K' > 1$ are constants depending on a, b, c. (The latter fact is an easy consequence of the standard fact [**23**] that the expected number of visits to the origin of $\mathbb{Z}^2$ after t steps of random walk is $O(\ln t)$ as $t \to \infty$.)

Simulations suggest that a much tighter bound on the discrepancy should actually hold in the situation of Theorem 5, and in fact the results seem consistent with a bound of the form const$/n$. The rotor configurations at large times are very interesting; see Plate 1. (Also compare with Plate 2.) Further discussion of these issues will appear in [**22**].

1.5. Transfinite walks. As mentioned above, Theorem 1 implies convergence of $n_t(b)/(n_t(b) + n_t(c))$ to $h(a)$ only if $n_t(b) + n_t(c) \to \infty$ as $t \to \infty$; we now investigate when this holds, and what can be done if it does not. We say that a rotor walk is *recurrent* if it visits every vertex infinitely often, and *transient* if it visits every vertex only finitely often.

LEMMA 6 (Recurrence and transience). *Any rotor walk associated with an irreducible Markov chain is either recurrent or transient.*

Note in particular that if V is finite and p is irreducible then every rotor walk is recurrent.

Fix an initial rotor configuration r_0 and an initial vertex $x_0 = a$. Suppose that the rotor walk $x_0, x_1, \ldots$ is transient. Then we can define a rotor configuration r_ω by $r_\omega(v) := \lim_{t\to\infty} r_t(v)$ (the limit exists since the sequence $r_t(v)$ is eventually constant). Now *restart* the particle at a by setting $x_\omega := a$, and define a rotor walk $x_\omega, x_{\omega+1}, x_{\omega+2}, \ldots$ according to the usual rules. If this is again transient we can set $r_{2\omega} := \lim_{t\to\infty} r_{\omega+t}$ and restart at $x_{2\omega} := a$, and so on. Continue in this way up to the first m for which the walk $x_{m\omega}, x_{m\omega+1}, \ldots$ is recurrent, or indefinitely if it is transient for all m. Call this sequence of walks a *transfinite rotor walk* started at a.

A *transfinite time* is a quantity of the form $\tau = \omega^2$, or $\tau = m\omega + t$ where m, t are non-negative integers. There is a natural order on transfinite times given by $m\omega + t < m'\omega + t'$ if and only if either $m < m'$ or both $m = m'$ and $t < t'$, while $m\omega + t < \omega^2$ for all m and t. For a transfinite walk and a transfinite time τ we write $n_\tau(v) = \#\{\alpha < \tau : x_\alpha = v\}$ for the number of visits to v before time τ. We sometimes say that the walk *goes to infinity* just before each of the times $\omega, 2\omega, \ldots$ at which it is defined.

LEMMA 7 (Transfinite recurrence and transience). *For an irreducible Markov chain and a transfinite rotor walk started at a, for any transfinite time τ, either $n_\tau(v)$ is finite for all v or $n_\tau(v)$ is infinite for all v. Also there exists $M \in \{1, 2, \ldots, \omega\}$ such that $n_{M\omega}(v)$ is infinite for all v and the rotor walk is defined at all $\tau < M\omega$.*

Note that while it is not obvious how to use a finite computer running in finite time to compute transfinite rotor walks in general, it is at least possible in certain settings, such as a random walk on the integers with a periodic initial rotor configuration.

THEOREM 8 (Transfinite walks and hitting probabilities). *Under the assumptions of Theorem 1, suppose further that p is irreducible, and that*

$$\limsup_{v \in V} h(v) = 0.$$

Then for any transfinite time $\tau = m\omega + t$ at which all vertices have been visited only finitely often,

$$\left| h(a) - \frac{n_\tau(b)}{n_\tau(b) + n_\tau(c) + m} \right| \leqslant \frac{K_1}{n_\tau(b) + n_\tau(c) + m}.$$

Thus the proportion of times the particle hits b as opposed to hitting c or going to infinity concentrates around $h(a)$. Furthermore, Lemma 7 ensures that $n_\tau(b) + n_\tau(c) + m \to \infty$ as $\tau \to M\omega$, so that the proportion converges to $h(a)$. The proof of Theorem 8 may be easily adapted to cover the probability of hitting a single vertex b, as opposed to escaping to infinity.

Next, for a vertex b, write $g(v) = g_b(v) := \mathbb{E}_v \sum_{t=0}^{\infty} \mathbf{1}[X_t = b]$ for the expected total number of visits to b. Note that this is finite for an irreducible, transient Markov chain.

THEOREM 9 (Transfinite walks and number of visits). *Consider an irreducible, transient Markov chain and fix vertices a, b. Suppose that*

$$\limsup_{v \in V} g(v) = 0.$$

Suppose moreover that the quantity

$$K_5 := \sup_{v \in V} g(v) + \frac{1}{2}\left(d(b) + \sum_{u,v \in V} d(u)p(u,v)\,|g(u) - g(v)| \right)$$

is finite. Then for any transfinite walk started at a, and for any transfinite time $\tau = m\omega + t$ at which all vertices have been visited only finitely often,

$$\left| g(a) - \frac{n_\tau(b)}{m} \right| \leqslant \frac{K_5}{m}.$$

It is natural to ask how recurrence and transience of rotor walks are related to recurrence and transience of the associated Markov chain. The following variant of an unpublished result of Oded Schramm provides an answer in one direction: in a certain asymptotic sense, the rotor walk is no more transient than the Markov chain. For a transfinite rotor walk started at vertex a, let I_n be the number of times the walk goes to infinity before the nth return to a. (I.e., $I_n := \max\{m \geqslant 0 : n_{m\omega}(a) < n\}$. This is well defined by Lemma 7; recall that the walk is restarted at a after each escape to infinity.)

THEOREM 10 (Transience density; Oded Schramm). *Consider an irreducible Markov chain, and an associated transfinite rotor walk started at vertex a. With I_n defined as above we have*

$$\limsup_{n\to\infty} \frac{I_n}{n} \leqslant \mathbb{P}_a(T_a^+ = \infty).$$

In particular we note that for a recurrent Markov chain the right side in Theorem 10 is zero, so the sequence of escapes to infinity has density zero in the sequence of returns to a. On the other hand, for a recurrent Markov chain it is possible for a rotor walk to go to infinity, for example in the case of simple symmetric random walk on $\mathbb{Z}$, with all rotors initially pointing in the same direction.

Moreover, for simple random walk on $\mathbb{Z}^2$ with all rotors initially in the same direction, the rotor walk goes to infinity infinitely many times. (To check this, suppose the rotors rotate anticlockwise and initially point East. Whenever the particle's horizontal coordinate achieves a new maximum, it is immediately sent directly Northwards to infinity. This happens infinitely often by Lemma 7.) See Plate 2 for a simulation of this remarkable process, and see [**22**] for further discussion.

On the other hand, it should be noted that the rotor walk on $\mathbb{Z}^2$ is recurrent for the initial configuration in Theorem 5 (see Figure 2). It is also possible for the rotor walk to be recurrent for a transient Markov chain, for example in the case of simple random walk on an infinite binary tree, with all rotors arranged so as to next send the particle towards the root. Landau and Levine [**14**] studied the rotor walk on regular trees in great detail, in particular identifying exactly which sequences $(I_n)_{n\geqslant 0}$ are possible on the binary tree. Further work on rotor walks on trees will appear in [**1**].

1.6. Stack walks. To generalize rotor walks to Markov chains with irrational transition probabilities, we must allow the particle to be routed to a non-periodic sequence of vertices on its successive visits to a given vertex.

Given a set V, a *stack mechanism* is an assignment of an infinite sequence of successors $u^{(1)}, u^{(2)}, \dots$ to each vertex $u \in V$. The *stack walk* started at x_0 is a sequence of vertices $x_0, x_1, \dots$ defined inductively by

$$x_{t+1} := x_t^{(n_t(x_t)+1)},$$

where

$$n_t(v) := \#\{s \in [0, t-1] : x_s = v\}.$$

(Note that, in the case of rational transition probabilities considered previously, the rotor walk can be regarded as a special case of a stack walk, with the periodic stacks given by $u^{(kd(u)+j)} = u^{(j)}$ for $1 \leqslant j \leqslant d(u)$ and $k \geqslant 0$.)

We illustrate the use of stacks with Theorem 12 below on hitting probabilities. The following will enable us to choose a suitable stack mechanism.

PROPOSITION 11 (Low-discrepancy sequence). *Let $p_1, \dots, p_n \in (0, 1]$ satisfy $\sum_i p_i = 1$. There exists a sequence $z_1, z_2, \dots \in \{1, \dots, n\}$ such that for all i and t,*

$$\left| p_i t - \#\{s \leqslant t : z_s = i\} \right| \leqslant 1. \tag{1.6}$$

Let p be a Markov transition kernel on V, and suppose that for each vertex u there are only finitely many vertices v such that $p(u, v) > 0$. We may then choose

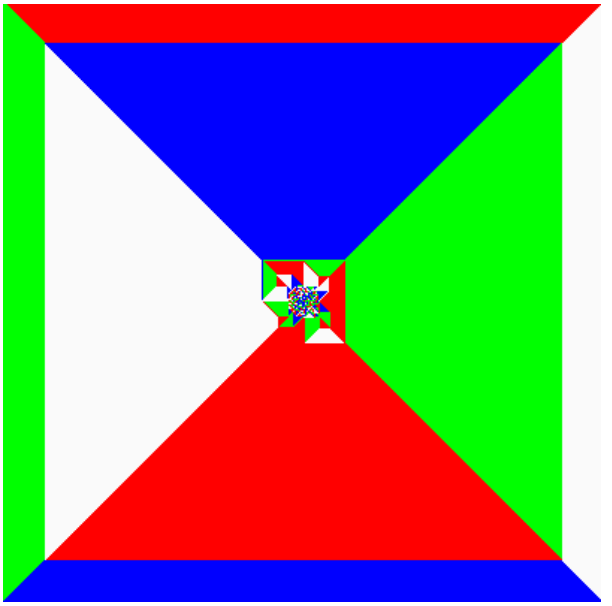


PLATE 1. The rotor configuration after 500 visits to b, starting in the configuration of Figure 2 with $a = c = (0, 0)$ and $b = (1, 1)$. The rotor directions are: East=white, North=red, West=green, South=blue.

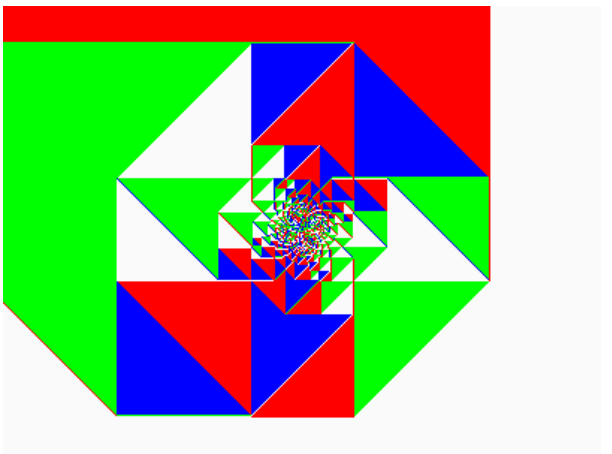


PLATE 2. The rotor configuration after 500 restarts from $a = (0, 0)$, for the transfinite rotor walk on $\mathbb{Z}^2$ with all rotors initially pointing East. The rotor directions are: East=white, North=red, West=green, South=blue. The red region extends infinitely far to the North.

a stack mechanism according to Proposition 11. More precisely, for each vertex u, enumerate the vertices v such that $p(u,v) > 0$ as $v_1, \dots, v_n$, and set $p_i = p(u, v_i)$. Then let $u^{(j)} := v_{z_j}$ where z is the sequence given by Proposition 11. Now let a, b, c be distinct vertices and assume that $p(b,a) = p(c,a) = 1$ and $b^{(i)} = c^{(i)} = a$ for all i. Write $h = h_{b,c}$.

THEOREM 12 (Stack walks). *Under the above assumptions, suppose that*

$$K_6 := 1 + \sum_{\substack{u \in V \setminus \{b,c\}, v \in V: \\ p(u,v) > 0}} |h(u) - h(v)|$$

is finite. For the stack mechanism described above, and any t,

$$\left| h(a) - \frac{n_t(b)}{n_t(b) + n_t(c)} \right| \leqslant \frac{K_6}{n_t(b) + n_t(c)}.$$

Proposition 11 can in fact be extended to the case of infinite probability vectors [**2**], and Theorem 12 carries over straightforwardly to this case. However, the result appears to have few applications in this broader context, since $\sum_v |h(u) - h(v)|$ is typically infinite when u has infinitely many successors.

1.7. Further remarks.

1.7.1. *History.* The rotor-router model was introduced by Priezzhev, Dhar, Dhar and Krishnamurthy [**21**] (under the name "Eulerian walkers model"), in connection with self-organized criticality. A special case was rediscovered in [**9**], in the analysis of some combinatorial games. The present article reports the first work on the close connection between rotor walks and Markov chains, originating in discussions between the two authors at a meeting in 2003. (Such a connection was however anticipated in the "whirling tours" theorem of [**9**], which shows that for random walk on a tree, the expected hitting time from one vertex to another can be computed by means of a special case of rotor walk; see also [**22**].) A special case of results presented here was reported in [**13**], and earlier drafts of the current work provided partial inspiration for some of the recent progress in [**3, 4, 5, 6, 8, 12, 17, 18, 19**], which we discuss below.

The idea of stack walks has its roots in Wilson's approach to random walks via random stacks; see [**24**].

1.7.2. *Time-dependent bounds.* We have chosen to focus on upper bounds of the form K_i/n, where K_i is a fixed constant not depending on time t. If this latter requirement is relaxed, our proofs may be adapted to give bounds that are stronger in some specific cases (at the expense of less clean formulations). Specifically, in each of Theorems 1–4 and 12, the claimed bound still holds if the relevant constant K_i is replaced with a modified quantity $K_i(t)$ obtained from K_i by:

(i) multiplying the initial additive term "1" or "$\max k(v)$" or "$\sup g(v)$" by the indicator $\mathbf{1}[x_t \neq x_0]$ (so that the term vanishes when the particle returns to its starting point); and

(ii) multiplying the summand in the sum $\sum_{u,v}$ by $\mathbf{1}[r_t(u) \neq r_0(u)]$ (so that in particular, terms corresponding to vertices u that have not been visited by time t vanish).

The same holds for Theorems 8 and 9 in the transfinite case, but replacing t with τ.

The above claims follow by straightforward modifications to our proofs. Indeed, our proof of Theorem 5 employs a special case of this argument. These and other refinements will be discussed more fully in the forthcoming article [**22**].

1.7.3. *Abelian property.* The rotor-router model has a number of interesting properties that will not be used directly in most of our proofs but which are nonetheless relevant. In particular, it enjoys an "Abelian property" which allows rotor walks to be parallelized. Specifically, consider a Markov chain on a finite set V with one or more *sinks*, i.e., vertices s with $p(s,s) = 1$, and suppose that from every vertex, some sink is accessible (so that the Markov chain eventually enters a sink almost surely). Then we may run several rotor walks simultaneously as follows. Start with an initial rotor configuration, and some non-negative number of particles at each vertex. At each step, choose a particle and route it according to the usual rotor mechanism; i.e., increment the rotor at its current vertex and move the particle in the new rotor direction. Continue until all particles are at sinks. It turns out that the resulting configuration of particles and rotors is independent of the order in which we chose to route the particles. This is the Abelian property; see, e.g., [**12**, Lemma 3.9] for a proof (and generalizations).

In the situation of Theorem 1, for example, assume that V is finite and the Markov chain is irreducible, and then modify it to make vertices b and c sinks. Start n particles at vertex a and perform simultaneous rotor walks. The Abelian property implies that the number of particles eventually at b is the same as the number $n_t(b)$ of times that b is visited when $n_t(b) + n_t(c) = n$ in the original set-up of Theorem 1, and the bound of Theorem 1 therefore applies.

A similar Abelian property holds for the "chip-firing" model introduced by Engel [**10, 11**] (later re-invented by Dhar [**7**] under the name "abelian sandpile model" as another model for self-organized criticality). The two models have other close connections, and in particular there is a natural group action involving sandpile configurations acting on rotor configurations. More details may be found in [**12**] and references therein. Engel's work was motivated by an analogy between Markov chains and chip-firing (indeed, he viewed chip-firing as an "abacus" for Markov chain calculations).

1.7.4. *Periodicity.* In the case when V is finite, we note the following very simple argument which gives bounds similar to Theorems 1–4 but with (typically) much worse constants. Since there are only finitely many rotor configurations, the sequence of vertices $((x_t, r_t))_{t \geqslant 0}$ is eventually periodic (with explicit upper bounds on the period and the time taken to become periodic which are exponentially large in the number of vertices). Therefore the proportion of time $n_t(v)/t$ spent at vertex v converges as $t \to \infty$ to some quantity $\mu(v)$, say, with a discrepancy bounded by const/t. Furthermore, as a consequence of the rotor mechanism, we have $\mu(u) = \sum_{v \in V} p(u,v)\mu(v)$ for all vertices u (because after many visits to u, the particle will have been routed to each successor approximately equal numbers of times). Thus μ is a stationary distribution for the Markov chain. This implies the bound in Theorem 4, except with a different (and typically much larger) constant in place of $K_4\pi(b)$. Similar arguments yield analogues of Theorems 1–3, but only in the case where V is finite.

1.7.5. *Related work.* As remarked earlier, rotor walks on trees were studied in detail by Landau and Levine [**14**]. Further results on rotor walks on trees will

appear in a forthcoming work of Angel and Holroyd [**1**], and further refinements and discussions of the results presented here will appear in Propp [**22**].

Cooper and Spencer [**6**] studied the following closely related problem. For the rotor walk associated with simple symmetric random walk on $\mathbb{Z}^d$, start with n particles at the origin, or more generally distributed in any fashion on vertices $(i_1, \ldots, i_d)$ with $i_1 + \cdots + i_d$ even, and apply one step of the rotor walk to each particle; repeat this t times. (It should be noted that the Abelian property does *not* apply here—the result is not the same as applying t rotor steps to each particle in an arbitrary order; see [**12**].) It is proved in [**6**] (see Figure 8) that the number of particles at a given vertex differs from the expected number of particles for n random walks by at most a constant (depending only on d). Furthermore, more precise estimates are proved in dimension $d = 1$ in [**5**] and in dimension $d = 2$ in [**8**].

The following rotor-based model for internal diffusion-limited aggregation (with acronym IDLA) was proposed by the second author, James Propp, and studied by Levine and Peres in [**16, 17, 18, 19**]. Starting with a rotor configuration on $\mathbb{Z}^d$, perform a sequence of rotor walks starting at the origin, stopping each walk as soon as it reaches a vertex not occupied by a previously stopped particle. It is proved in [**18**] that, as the number of particles n increases, the shape of the set of occupied vertices converges to a d-dimensional Euclidean ball; generalizations and more accurate bounds are proved in [**17, 19**].

2. Proofs of basic results

Theorems 1–4 will all follow as special cases of Proposition 13 below, and the remaining results will also follow by adapting the same proof. For any Markov transition kernel p and any function $f\colon V \to \mathbb{R}$ we define the *Laplacian* $\Delta f\colon V \to \mathbb{R}$ by

$$\Delta f(u) := \sum_{v \in V} p(u,v) f(v) - f(u). \tag{2.1}$$

Proposition 13 (Key bound). *For any rotor walk $x_0, x_1, \ldots$ associated with p, any function f and any t we have*

$$\left| \sum_{s=0}^{t-1} \Delta f(x_s) \right| \leqslant |f(x_t) - f(x_0)| + \frac{1}{2} \sum_{u,v \in V} d(u) p(u,v) \, |f(u) - f(v) + \Delta f(u)| \,.$$

The proofs of Theorems 1–4 will proceed by applying Proposition 13 to a suitable f. The proof of Proposition 13 will use the following simple fact.

Lemma 14. *If $\sum_{i=1}^n a_i = 0$ then $\left|\sum_{i=1}^j a_i - \sum_{i=1}^k a_i\right| \leqslant \frac{1}{2} \sum_{i=1}^n |a_i|$ for all $j, k \in [1, n]$.*

Proof. We prove the stronger statement that $|\sum_{i \in S} a_i| \leqslant \frac{1}{2} \sum_{i=1}^n |a_i|$ for any subset S of $\{1, \ldots, n\}$: assuming without loss of generality that $\sum_{i \in S} a_i$ is positive, it is at most $\sum_{a_i : a_i > 0} a_i = \frac{1}{2}(\sum_{i : a_i > 0} a_i - \sum_{i : a_i < 0} a_i) = \frac{1}{2} \sum_{i=1}^n |a_i|$. □

Proof of Proposition 13. Recall that r_0 denotes the initial rotor configuration. For a vertex x and a rotor configuration r, consider the quantity

$$\Phi(x, r) := f(x) + \sum_{u \in V} \big[\phi(u, r(u)) - \phi(u, r_0(u))\big],$$

where

$$\phi(u,j) := \sum_{i=1}^{j} \big[f(u) - f(u^{(i)}) + \Delta f(u) \big].$$

Note that $\Phi(x, r_t)$ is finite if r_t is any rotor configuration encountered by the rotor walk, since the only non-zero terms in the sum over u are those corresponding to vertices that the walk has visited (this is the reason for including the term "$-\phi(u, r_0(u))$" in the above definition). Note also that the definition of the Laplacian (2.1) and the rotor property (1.1) imply for all $u \in V$ that

$$\phi(u, d(u)) = 0. \tag{2.2}$$

Let us compute the change in Φ produced by a step of the rotor walk from (x_t, r_t) to (x_{t+1}, r_{t+1}). The only term in the sum over u that changes is the one corresponding to $u = x_t$, and thus

$$\begin{aligned}
\Phi(x_{t+1}, r_{t+1}) - \Phi(x_t, r_t) &= f(x_{t+1}) - f(x_t) + [\phi(x_t, r_{t+1}(x_t)) - \phi(x_t, r_t(x_t))] \\
&= f(x_{t+1}) - f(x_t) + [f(x_t) - f(x_t^{(r_{t+1}(x_t))}) + \Delta f(x_t)] \\
&= \Delta f(x_t),
\end{aligned}$$

where we have used (2.2) in the case when $r_{t+1}(x_t) = 1$. Therefore $\Phi(x_t, r_t) - \Phi(x_0, r_0) = \sum_{s=0}^{t-1} \Delta f(x_s)$. Also $\Phi(x_0, r_0) = f(x_0)$, so we obtain

$$\sum_{s=0}^{t-1} \Delta f(x_s) = f(x_t) - f(x_0) + \sum_{u \in V} \big[\phi(u, r_t(u)) - \phi(u, r_0(u)) \big]. \tag{2.3}$$

In order to bound the last sum in (2.3), we use (2.2) together with Lemma 14 and the definition of ϕ to deduce

$$\begin{aligned}
\big| \phi(u, r_t(u)) - \phi(u, r_0(u)) \big| &\leqslant \frac{1}{2} \sum_{i=1}^{d(u)} \big| f(u) - f(u^{(i)}) + \Delta f(u) \big| \\
&= \frac{1}{2} \sum_{v \in V} d(u) p(u,v) \big| f(u) - f(v) + \Delta f(u) \big| \qquad (2.4)
\end{aligned}$$

(since $d(u)p(u,v)$ is the number of i such that $u^{(i)} = v$). We conclude by applying the triangle inequality to (2.3). □

Proof of Theorem 1. We shall apply Proposition 13 with $f = h_{b,c}$. Note that $h(b) = 1$ and $h(c) = 0$, while conditioning on the first step of the Markov chain gives $h(u) = \sum_{v \in V} p(u,v) h(v)$ for $u \neq b, c$. Hence, using $p(b,a) = p(c,a) = 1$,

$$\Delta h(u) = \begin{cases} 0, & u \neq b, c; \\ h(a) - 1, & u = b; \\ h(a), & u = c, \end{cases}$$

and thus $\sum_{s=0}^{t-1} \Delta h(x_s) = h(a)[n_t(b) + n_t(c)] - n_t(b)$.

Turning to the other terms in Proposition 13, note that $|h(x_t) - h(x_0)| \leqslant 1$, and $h(u) - h(v) + \Delta h(u) = 0$ when $u \in \{b, c\}$ and $v = a$. Substituting into Proposition 13 gives $|h(a)[n_t(b) + n_t(c)] - n_t(b)| \leqslant K_1$ as required. □

PROOF OF THEOREM 2. We shall apply Proposition 13 with $f = k_b$. In this case

$$\Delta k(u) = \begin{cases} -1, & u \neq b; \\ k(a), & u = b, \end{cases}$$

and thus $\sum_{s=0}^{t-1} \Delta k(x_s) = (n_t(b))(k(a)) + (t - n_t(b))(-1) = (k(a) + 1)n_t(b) - t$. Substituting into Proposition 13 and using $|k(x_t) - k(x_0)| \leqslant \max_{v \in V} k(v)$ and $k(b) - k(a) + \Delta k(b) = 0$ completes the proof. □

To prove Theorem 3 we note some elementary facts about Markov chains.

LEMMA 15. *Let b, c be two distinct vertices of an irreducible recurrent Markov chain, and let π be a stationary vector. Then $\pi(b)e_{b,c} = \pi(c)e_{c,b}$. Also the hitting probabilities $h = h_{b,c}$ satisfy $\Delta h(b) = -e_{b,c}$ and $\Delta h(c) = e_{c,b}$.*

PROOF. Let N denote the number of visits to c before the first return to b when started from b. It is a standard fact (see, e.g., [**20**, Theorem 1.7.6]) that $\mathbb{E}N = \pi(c)/\pi(b)$. On the other hand $\mathbb{P}(N = n) = e_{b,c}(1 - e_{c,b})^{n-1}e_{c,b}$ for $n \geqslant 1$, so $\mathbb{E}N = e_{b,c}/e_{c,b}$, and the first claim follows. For the remaining claims we compute Δh by conditioning on the first step: $\Delta h(b) = (1 - e_{b,c}) - h(b) = -e_{b,c}$ and $\Delta h(c) = e_{c,b} - h(c) = e_{c,b}$. □

PROOF OF THEOREM 3. We shall again apply Proposition 13 with $f = h = h_{b,c}$ (now without the restriction $p(b, a) = p(c, a) = 1$). Lemma 15 gives

$$\Delta h(u) = \begin{cases} 0, & u \neq b, c; \\ -e_{b,c}, & u = b; \\ e_{c,b}, & u = c, \end{cases}$$

and so $\sum_{s=0}^{t-1} \Delta h(x_s) = -n_t(b)e_{b,c} + n_t(c)e_{c,b}$. In order to bound the terms in the last sum in Proposition 13 in the cases $u = b, c$ note that $|\Delta h(u)| \leqslant 1$ in these cases, and so, for $u = b, c$,

$$\sum_{v \in V} p(u, v) \left|h(u) - h(v) + \Delta h(u)\right| \leqslant 1 + \sum_{v \in V} p(u, v) \left|h(u) - h(v)\right|.$$

Hence Proposition 13 gives

$$\left|n_t(b)e_{b,c} - n_t(c)e_{c,b}\right| \leqslant K_3.$$

Now divide through by $\pi(b)e_{b,c}$ (which equals $\pi(c)e_{c,b}$ by Lemma 15). □

PROOF OF THEOREM 4. We shall apply Proposition 13 with $f = k = k_b$. Note that $k(b) = 0$, while $\mathbb{E}_b T_b^+ = 1 + \sum_{v \in V} p(b, v)k(v)$. Also we have $\mathbb{E}_b T_b^+ = 1/\pi(b)$ (see [**20**]), hence

$$\Delta k(u) = \begin{cases} -1, & u \neq b; \\ 1/\pi(b) - 1, & u = b. \end{cases}$$

We bound the term for $u = b$ in Proposition 13 thus:

$$\sum_{v \in V} p(b, v) \left|k(b) - k(v) + \Delta k(b)\right| \leqslant 1/\pi(b) + \sum_{v \in V} p(b, v) \left|k(b) - k(v) - 1\right|.$$

We obtain

$$\left|\frac{n_t(b)}{\pi(b)} - t\right| \leqslant K_4,$$

and multiply by $\pi(b)/t$ to conclude. □

3. Proofs for walks on $\mathbb{Z}^2$

Our proof of Theorem 5 is based on the two lemmas below. For $k \geqslant 1$ we define the kth *box* $B(k) := (-k,k]^2 \cap \mathbb{Z}^2$ and the kth *layer* $\partial B(k) := B(k) \setminus B(k-1)$. See Figure 2.

LEMMA 16. *Fix two distinct vertices b, c of $\mathbb{Z}^2$, and let $h = h_{b,c}$ be the hitting probability for the simple random walk on the square lattice. There exists $C = C(b,c) \in (0,\infty)$ such that for all positive integers k,*

$$\sum_{\substack{u,v \in B(k): \\ \|u-v\|_1 = 1}} |h(u) - h(v)| \leqslant C \ln k.$$

PROOF. Fix b, c and write $C_1, C_2, \ldots$ for constants depending on b, c. We claim first that for all $v \in \mathbb{Z}^2$,

$$h(v) = C_1 + e_{b,c}[a(v-c) - a(v-b)], \tag{3.1}$$

where $a\colon \mathbb{Z}^2 \to \mathbb{R}$ is the *potential kernel* of $\mathbb{Z}^2$. (The function a may be expressed as $a(v) := \lim_{n\to\infty} \sum_{t=0}^{n} [\mathbb{P}(X_t = 0) - \mathbb{P}(X_t = v)]$, where (X_t) is the simple random walk on $\mathbb{Z}^2$; for more information see, e.g., [**23**, Ch. 3] or [**15**, Sect. 1.6].) To check (3.1), we note the following facts about a. Firstly,

$$a(v) = A + \tfrac{2}{\pi} \ln |v| + O(|v|^{-2}) \quad \text{as } |v| \to \infty, \tag{3.2}$$

where $|v| := \|v\|_2$ and A is an absolute constant (see [**15**, p. 39]). Since $\frac{\mathrm{d}}{\mathrm{d}x}(A + \frac{2}{\pi} \ln x) = \frac{2}{\pi} x^{-1}$ we deduce

$$|a(v-c) - a(v-b)| \leqslant C_2 |v|^{-1}.$$

Secondly, writing Δ for the Laplacian of the random walk on $\mathbb{Z}^2$, i.e., $\Delta f(u) := \frac{1}{4} \sum_{v:\|u-v\|_1=1} f(v) - f(u)$, we have

$$\Delta a(v) = \mathbf{1}[v = 0].$$

Hence, using Lemma 15 and the fact that $\pi \equiv 1$ is a stationary vector for the random walk, the function $v \mapsto h(v) - e_{b,c}[a(v-c) - a(v-b)]$ is bounded and harmonic, therefore constant, establishing (3.1).

We now claim that for all u, v with $\|u - v\|_1 = 1$,

$$|h(u) - h(v)| \leqslant C_3 |v|^{-2}. \tag{3.3}$$

Once this is established we obtain

$$\sum_{\substack{u,v \in B(k): \\ \|u-v\|_1 = 1}} |h(u) - h(v)| \leqslant \sum_{j=1}^{k} C_4 j (C_3 j^{-2}) \leqslant C \ln k.$$

as required.

Finally, turning to the proof of (3.3), combining (3.1) and (3.2) gives

$$h(u) - h(v) = C_5 \big(\ln|u-c| - \ln|u-b| - \ln|v-c| + \ln|v-b| \big) + O(|v|^{-2}).$$

In order to bound the above expression, fix $u - v$ to be one of the 4 possible integer unit vectors, and write $v - c = z$ and $c - b = \alpha$ and $u - v = \beta$. For convenience

identify the vector (x, y) with the complex number $x + \mathrm{i}y$ and let $|\cdot|$ denote the modulus. We have

$$\begin{aligned}
&\ln|u-c| - \ln|u-b| - \ln|v-c| + \ln|v-b| \\
&\quad = \ln\left|\frac{(z+\alpha)(z+\beta)}{(z+\alpha+\beta)z}\right| = \ln\left|1 + \frac{\alpha}{z} + \frac{\beta}{z} - \frac{\alpha}{z} - \frac{\beta}{z} + O(|z|^{-2})\right| \\
&\quad = O(|z|^{-2}), \qquad \text{as } z \to \infty.
\end{aligned}$$

□

Fix $a, b, c \in \mathbb{Z}^2$, and consider the rotor walk $x_0, x_1, \ldots$ started at a with rotor mechanism (1.4) and rotor configuration (1.5) modified so that $p(b, a) = p(c, a) = 1$ as discussed in the paragraph preceding the statement of Theorem 5. We say that the walk *enters a new layer* at time t if for some k we have $x_0, \ldots, x_{t-1} \in B(k)$ but $x_t \notin B(k)$.

LEMMA 17. *Under the above assumptions, between any two times at which the rotor walk enters a new layer, it must visit vertex a at least once. Also, between any two consecutive visits to vertex a, no vertex is visited more than 4 times.*

PROOF. We start by proving the first assertion. The reader may find it helpful to consult Figure 2 throughout. Suppose for a contradiction that $a \in B(k-1)$, and that the rotor walk enters both the layers $\partial B(k)$ and $\partial B(k+1)$ for the first time without visiting a in between. Let s be the time of the last visit to a prior to entering $\partial B(k)$, and let t be the first time at which $\partial B(k+1)$ is entered.

We claim that some vertex v emitted the particle at least 5 times during $[s, t]$. To prove this, note first that $x_{t-1} \in \partial B(k)$, and consider the following two cases. If x_{t-1} is not one of the four "corner vertices" of $\partial B(k)$, then immediately after the particle moves from x_{t-1} to $x_t \in \partial B(k+1)$, the rotor at x_{t-1} is pointing in the same direction as in the initial rotor configuration r. Since this rotor did not move before time s, vertex x_{t-1} must have emitted the particle at least 4 times during $[s, t]$. Therefore, x_{t-1} must have received the particle at least 4 times from among its 4 neighbors in $[s, t]$; but it has not received the particle from x_t, therefore by the pigeonhole principle it received it at least twice from some other neighbor v. And $v \notin \{b, c\}$ since $x_{t-1} \neq a$. By considering the rotor at v, we see that this implies that v emitted the particle at least 5 times during $[s, t]$. On the other hand, if x_{t-1} is a corner vertex of $\partial B(k)$, then on comparing with the initial rotor configuration r we see that x_{t-1} has emitted (and hence received) the particle 3 or 4 times, but two of its neighbors lie in $\partial B(k+1)$, so it did not receive the particle from them, and the same argument now applies. Thus we have proved the above claim.

Now let u be the first vertex to emit the particle 5 times during $[s, t]$. Then $u \notin \{a, b, c\}$, otherwise we should have a contradiction to our assumption that a is visited only once. But now repeating the argument above, u must have received the particle 5 times, so it must have received it at least twice from some neighbor, not in $\{a, b, c\}$; so this neighbor must have emitted the particle 5 times by some earlier time in $[s, t]$, a contradiction. Thus the first assertion is established.

The second assertion follows by an almost identical argument: if some vertex is visited at least 5 times between visits to a, then considering the first vertex to be so visited leads to a contradiction. □

PROOF OF THEOREM 5. We write $C_1, C_2, \ldots$ for constants which may depend on a, b, c. We use the proof of Proposition 13 in the case $f = h$. As in the proof of

Theorem 1, equation (2.3) becomes

$$h(a)n - n_t(b) = h(x_t) - h(a) + \sum_{u \in V \setminus \{b,c\}} \big[\phi(u, r_t(u)) - \phi(u, r_0(u))\big],$$

where $n = n_t := n_t(b) + n_t(c)$. However, the term $\phi(u, r_t(u)) - \phi(u, r_0(u))$ is non-zero only for those vertices which have been visited by time t. Now the first assertion of Lemma 17 implies that at most one new layer is entered for each visit to a, and thus for each visit to $\{b, c\}$. Hence for some C_1, all the vertices visited by time t lie in $B(n + C_1)$ (where the constant C_1 depends on the layer of the initial vertex a).

Now proceeding as in the proof of Proposition 13 and using Lemma 16,

$$\begin{aligned}\left| \sum_{u \in B(n+C_1)\setminus\{b,c\}} \big[\phi(u, r_t(u)) - \phi(u, r_0(u))\big] \right| &\leqslant \tfrac{1}{2} \sum_{\substack{u,v \in B(n+C_1+1): \\ \|u-v\|_1 = 1}} |h(u) - h(v)| \\ &\leqslant C \ln n.\end{aligned}$$

Combining this with the above facts gives

$$\big|h(a)n - n_t(b)\big| \leqslant 1 + C \ln n,$$

as required.

Finally to prove the bound $t \leqslant C'n^3$, we note by the second assertion of Lemma 17 that after n visits to vertex a, each of the at most $C_2 n^2$ vertices in $B(n + C_1)$ has been visited at most $4n$ times, so the total number of time steps is at most $4C_2 n^3$. □

4. Proofs for transfinite walks

Proof of Lemma 6. By irreducibility it is enough to show that if u is visited infinitely often and $p(u, v) > 0$ then v is visited infinitely often. But this is immediate since $v = u^{(i)}$ for some i, so the rotor at u will be incremented to point to v infinitely often. □

Proof of Lemma 7. As in the preceding proof, if u is visited infinitely often and $p(u, v) > 0$ then v is visited infinitely often, proving the first assertion. For the second assertion, let M be one greater than the first m for which the walk $x_{m\omega}, x_{m\omega+1}, \dots$ is recurrent, or $M = \omega$ if all are transient. Then a is visited infinitely often before time $M\omega$, and we apply the first assertion. □

Proof of Theorem 8. We consider the quantity Φ defined in the proof of Proposition 13, with $f = h = h_{b,c}$ (as in the proof of Theorem 1). Suppose $x_0, x_1, \dots$ is a transient rotor walk. We claim that

$$\Phi(x_\omega, r_\omega) - \lim_{t \to \infty} \Phi(x_t, r_t) = h(a). \tag{4.1}$$

The claim is proved as follows. The assumption of the theorem and the fact that the walk is transient imply that $\lim_{t\to\infty} h(x_t) = 0$. We clearly have $\lim_{t\to\infty} \phi(u, r_t(u)) = \phi(u, r_\omega(u))$ for each u, and by (2.4) and the definition of K_1 in Theorem 1 we have for all u and t that $|\phi(u, r_t(u)) - \phi(u, r_0(u))| \leqslant F(u)$ where $\sum_{u \in V} F(u) \leqslant 2(K_1 - 1)$. Hence by the dominated convergence theorem,

$$\lim_{t \to \infty} \Phi(x_t, r_t) = 0 + \sum_{u \in V} \big[\phi(u, r_\omega(u)) - \phi(u, r_0(u))\big] = \Phi(x_\omega, r_\omega) - h(a).$$

We have proved claim (4.1); thus whenever we "restart from infinity to a", the quantity Φ increases by $h(a)$. Combining this with the argument from the proof of Theorem 1, we get

$$\big[n_\tau(b) + n_\tau(c) + m\big]h(a) - n_\tau(b) = \Phi(x_\tau, r_\tau) - \Phi(x_0, r_0)$$

for $\tau = m\omega + t$, and the right side is bounded in absolute value by K_1 exactly as in the proof of Theorem 1. □

Proof of Theorem 9. We consider the quantity Φ defined in the proof of Proposition 13, with $f = g = g_b$. Note that

$$\Delta g(u) = \begin{cases} 0, & u \neq b; \\ -1, & u = b. \end{cases}$$

Mimicking the proof of Theorem 8, we obtain

$$g(a)m - n_\tau(b) = \Phi(x_\tau, r_\tau) - \Phi(x_0, r_0),$$

and we bound the right side as in the previous proofs, noting that when $u = b$ we have $|g(u) - g(v) + \Delta g(u)| \leqslant |g(u) - g(v)| + 1$. □

Our proof of Theorem 10 is based on an argument of Oded Schramm which is unpublished (although we present the details in a somewhat different way). We shall need some preparation. It will be convenient to work with $R_n := n - I_n$, i.e., the number of times the transfinite rotor walk returns to a *without* going to infinity up to the time of the nth return to a. We also introduce some modified Markov chains and rotor mechanisms as follows.

Firstly, replace the vertex a with two vertices a_0 and a_1. Let $\widehat{V} = (V \setminus \{a\}) \cup \{a_0, a_1\}$ denote this modified vertex set. Introduce a modified transition kernel $\widehat{p}$ by letting a_0 inherit all the outgoing transition probabilities from a, and letting a_1 inherit all the incoming transition probabilities to a (i.e., let $\widehat{p}(a_0, v) = p(a, v)$ and $\widehat{p}(v, a_1) = p(v, a)$ for all $v \in V \setminus \{a\}$); also let $\widehat{p}(a_1, a_0) = 1$ and $\widehat{p}(a_0, a_1) = 0$, and let $\widehat{p}$ otherwise agree with p.

Secondly, for a positive integer d, let $B(d)$ denote the set of vertices that can be reached in at most d steps of the original Markov chain starting from a, and let $\partial B(d) := B(d) \setminus B(d-1)$. Let $\widehat{p}^d$ be $\widehat{p}$ modified so that $\widehat{p}^d(b, a_0) = 1$ for all $b \in \partial B(d)$. (Thus, on reaching distance d from a, the particle is immediately returned to a_0).

Fix a rotor mechanism and initial rotor configuration for the original Markov chain, and modify them accordingly to obtain a rotor walk associated with $\widehat{p}^d$, started at a_0. Let R_n^d be the number of times this rotor walk hits a_1 before the nth return to a_0 (i.e., before the $(n+1)$st visit to a_0). Also note that R_n is the number of times the transfinite rotor walk associated with $\widehat{p}$ and started at a_0 hits a_1 before the nth return to a_0.

Lemma 18. *For a fixed initial rotor configuration, and any non-negative integer n, we have $R_n^d \to R_n$ as $d \to \infty$ (i.e., $R_n^d = R_n$ for d sufficiently large).*

Proof. For $v \in \widehat{V}$, let $N_n^d(v)$ (respectively $N_n(v)$) be the number of visits to vertex v before the nth return to a_0 for the (transfinite) rotor walk associated with $\widehat{p}^d$ (respectively $\widehat{p}$). We claim that

$$N_n^d \to N_n \qquad \text{as } d \to \infty, \tag{4.2}$$

where the convergence is in the product topology on $\mathbb{N}^{\widehat{V}}$; in other words, for any finite set $F \subset \widehat{V}$, if d is sufficiently large then $N_n^d(v) = N_n(v)$ for all $v \in F$. The required result follows immediately from this, because $R_n^d = N_n^d(a_1)$ and $R_n = N_n(a_1)$.

We prove (4.2) by induction on n. It holds trivially for $n = 0$ because N_0^d and N_0 equal zero everywhere. Assume it holds for $n-1$. This implies in particular that the rotor configuration at the time of the $(n-1)$st return to a_0 similarly converges as $d \to \infty$ to the corresponding rotor configuration in the transfinite case. Now consider the portion of the transfinite rotor walk corresponding to $\widehat{p}$, starting just after the $(n-1)$st return to a_0, up until the nth return to a_0. Consider the following two possibilities. If this walk is recurrent (so that it returns to a_0 via a_1) then it visits only finitely many vertices, so if d is sufficiently large that N_{n-1}^d and N_{n-1} agree on all the vertices it visits, then N_n^d and N_n also agree on the same set of vertices, establishing (4.2) in this case. On the other hand, suppose the aforementioned walk is transient (so that it goes to infinity before being restarted at a_0). Given a finite set $F \subset \widehat{V}$, let d be such that that when this walk leaves F for the last time, it has never been outside $B(d)$. Now let d' be such that $N_{n-1}^{d'}$ and N_{n-1} agree on $B(d)$. Then N_n^d and N_n will agree on F. So (4.2) holds in this case also, and the induction is complete. □

LEMMA 19. *For all positive integers n and d we have $R_n^{d+1} \geqslant R_n^d$.*

PROOF. This will follow by a special case of the Abelian property for rotor walks on finite graphs with a sink (see, e.g., [**12**, Lemma 3.9]). First we slightly modify the mechanism yet again. Consider the rotor mechanism and initial rotor configuration corresponding to $\widehat{p}^{d+1}$. Remove all the vertices in $V \setminus B(d+1)$ (these cannot be visited by the rotor walk started at a_0 anyway). Introduce an additional *absorbing* vertex s (called the sink), and modify the transition probabilities so that on hitting a_1 or $\partial B(d+1)$, particles are sent immediately to s instead of to a_0. Modify the rotor mechanism accordingly, but do not otherwise modify the initial rotor configuration.

We now consider the following multi-particle rotor walk (see, e.g., [**12**] or the discussion in the introduction for more information). Start with n particles at a_0, and perform a sequence of rotor steps. That is, at each step, choose any non-sink vertex which has a positive number of particles (if such exists), and *fire* the vertex; i.e., increment its rotor, and move one particle in the new rotor direction. Continue in this way until all particles are at the sink. [**12**, Lemma 3.9] states that the total number of times any given vertex fires during this procedure is independent of our choices of which vertex to fire.

In particular, consider the firing order in which we first move one particle repeatedly (so that it performs an ordinary rotor walk) until it reaches s, then move the second particle in the same way, and so on. Thus the number of times a_1 fires is R_n^{d+1}. Alternatively, we may move one particle until the first time it reaches $\partial B(d) \cup \{s\}$, then "freeze" it, and move the second particle until it reaches $\partial B(d) \cup \{s\}$, and so on. At this stage, the number of times a_1 has fired is R_n^d. Now we can continue firing until the frozen particles reach s. Comparing the two procedures shows $R_n^{d+1} \geqslant R_n^d$. □

COROLLARY 20. *For all positive integers n and d we have $R_n \geqslant R_n^d$.*

PROOF. Immediate from Lemmas 18 and 19. □

PROOF OF THEOREM 10. Since $R_n = n - I_n$, the required result is clearly equivalent to $\liminf_{n\to\infty} R_n/n \geqslant \mathbb{P}_a(T_a^+ < \infty)$. Fix any $\epsilon > 0$. Then there exists d such that $\mathbb{P}_a(T_a^+ < T_{\partial B(d)}) \geqslant \mathbb{P}_a(T_a^+ < \infty) - \epsilon$. Now consider the modified rotor walk corresponding to $\widehat{p}^d$ as defined above. Since the set of vertices that can be reached from a_0 is finite (so in effect the vertex set is finite), Theorem 1 implies that $R_n^d/n \to \mathbb{P}_a(T_a^+ < T_{\partial B(d)})$ as $n \to \infty$. Putting these facts together with Corollary 20 we obtain

$$\liminf_{n\to\infty} \frac{R_n}{n} \geqslant \lim_{n\to\infty} \frac{R_n^d}{n} = \mathbb{P}_a(T_a^+ < T_{\partial B(d)}) \geqslant \mathbb{P}_a(T_a^+ < \infty) - \epsilon. \qquad \square$$

5. Proofs for stack walks

In this section we shall prove Proposition 11, and use it together with Proposition 21 below to prove Theorem 12. Given a Markov chain and a stack mechanism, we define the discrepancy functions

$$D_n(u,v) := \#\{i \leqslant n : u^{(i)} = v\} - np(u,v).$$

PROPOSITION 21. *For any Markov chain, any stack walk, any function f and any t,*

$$\sum_{s=0}^{t-1} \Delta f(x_s) = f(x_t) - f(x_0) + \sum_{u,v\in V} D_{n_t(u)}(u,v)\big[f(u) - f(v) + \Delta f(u)\big].$$

PROOF. Consider the function

$$\Psi(t) := f(x_t) + \sum_{u\in V} \psi(u, n_t(u)),$$

where

$$\psi(u,n) := \sum_{i=1}^{n} \big[f(u) - f(u^{(i)}) + \Delta f(u)\big].$$

As in the proof of Proposition 13 we have $\sum_{s=0}^{t-1} \Delta f(x_s) = \Psi(t) - \Psi(0)$.

From the definition of D we have

$$\psi(u,n) = \sum_{v\in V} \big[D_n(u,v) + np(u,v)\big]\big[f(u) - f(v) + \Delta f(u)\big].$$

But by the definition of the Laplacian, $\sum_{v\in V} p(u,v)[f(u) - f(v) + \Delta f(u)] = 0$; therefore

$$\psi(u,n) = \sum_{v\in V} D_n(u,v)\big[f(u) - f(v) + \Delta f(u)\big],$$

and the result follows on substituting. □

PROOF OF PROPOSITION 11. First note that it suffices to prove the case in which $p_1, \dots, p_n$ are all rational. The irrational case then follows by a limiting argument. Specifically, let $p_1^k, \dots, p_n^k$ be rational with $p_i^k \to p_i$ as $k \to \infty$, and let $z_1^k, z_2^k, \dots$ be a sequence satisfying (1.6) for the p_i^k's; then by a compactness argument (since $\{1, \dots, n\}$ is finite) there is a subsequence (k_j) and a sequence $z_1, z_2, \dots$ such that $z_t^{k_j} \to z_t$ for each t, and (z_t) then satisfies (1.6) for the p_i's.

Now suppose that $p_1, \dots, p_n$ are rational, and let d be their least common denominator. Consider the finite bipartite graph G with vertex classes $L := \{1, \dots, d\}$ and $R := \bigcup_{i=1}^n R_i$ where $R_i := \{(i, m) : m \in \{1, \dots, p_i d\}\}$, and with an edge from $t \in L$ to $(i, m) \in R$ if and only if

$$\left\lceil \frac{m-1}{p_i} \right\rceil \leqslant t \leqslant \left\lceil \frac{m}{p_i} \right\rceil. \tag{5.1}$$

We shall prove that G has a perfect matching between L and R. Note first that $\#L = d = \sum_i p_i d = \#R$. We claim that any set $T \subseteq L$ has at least $p_i \#T$ neighbors in R_i; from this it follows that it has at least $\#T$ neighbors in R, and the existence of a perfect matching then follows from Hall's marriage theorem. To prove the claim, fix $i \in \{1, \dots, n\}$ and note that (5.1) is equivalent to $p_i t - p_i < m \leqslant p_i t + 1$. Therefore in the case when T is an interval $[s, t]$, it is adjacent to all those pairs $(i, m) \in R_i$ for which m is an integer in $(p_i s - p_i, p_i t + 1] \cap [1, d]$; this includes all integers in $(p_i s - p_i, p_i t + 1)$ (since $p_i s - p_i \geqslant 0$ and $p_i t + 1 \leqslant d + 1$). The latter interval has length $p_i(t - s + 1) + 1$, therefore it contains at least $p_i(t - s + 1) = p_i \#T$ integers as required. Now consider the case $T = [s, t] \cup [u, v]$ (where $u > t + 1$). If the two intervals have disjoint neighborhoods in R_i, the claim follows by applying the single-interval case to each, and summing. On the other hand if the neighborhoods of the two intervals intersect, we see from (5.1) that the neighborhood in R_i of T is the same as the neighborhood in R_i of the larger set $[s, v]$, so the claim again follows from the single-interval case. Finally, the case when T is a union of three or more intervals is handled by applying the same reasoning to each adjacent pair, proving the claim and hence the existence of a matching.

Fix a perfect matching of G, and for $t = 1, \dots, d$, let $z_t := i$ where R_i contains the partner of t. It follows from (5.1) that if (i, m) and $(i, m + 1)$ have respective partners t and t' then $t < t'$. Therefore t and (i, m) are partners if and only if z_t is the mth occurrence of i in the sequence z; from (5.1) this mth occurrence appears between positions $\lceil \frac{m-1}{p_i} \rceil$ and $\lceil \frac{m}{p_i} \rceil$. Thus,

$$\lfloor p_i t \rfloor \leqslant \#\{s \leqslant t : z_s = i\} \leqslant \lfloor p_i t \rfloor + 1,$$

and it follows that (1.6) holds for all $t \leqslant d$. Note also that the left side of (1.6) is zero for $t = d$, therefore continuing the sequence z so as to be periodic with period d completes the proof. □

(For an alternative proof of Proposition 11 that applies also to infinite probability vectors, see [**2**].)

Proof of Theorem 12. Choosing the stack mechanism according to Proposition 11 ensures that $|D_n(u, v)| \leqslant 1$ for all u, v and n. Now apply Proposition 21 to $f = h$ to obtain

$$\begin{aligned} &|h(a)[n_t(b) + n_t(c)] - n_t(b)| \\ &\qquad \leqslant |h(x_t) - h(x_0)| + \sum_{\substack{u \in V \setminus \{b, c\}, \\ v \in V}} \left| D_{n_t(u)}(u, v) \right| \cdot |h(u) - h(v)|, \end{aligned}$$

and conclude by noting that $D_{n_t(u)}(u, v) = 0$ unless $p(u, v) > 0$. □

Open questions

As the burgeoning literature on Eulerian walk and rotor-routing attests, there are numerous interesting open problems. Here we focus on a few that are related to rotor walk on Euclidean lattices.

(i) Can the bound $C \log n/n$ in Theorem 5 for the discrepancy in hitting probabilities for simple random walk on $\mathbb{Z}^2$ be improved to C/n? Do similar results hold in $\mathbb{Z}^d$ and for other initial rotor configurations?

(ii) For simple random walk on $\mathbb{Z}^2$, let all rotors initially point East, and consider the transfinite rotor walk restarted at the origin after each escape to infinity. What is the asymptotic behavior of I_n, the number of escapes to infinity before the nth return to the origin? Theorem 10 implies that $I_n/n \to 0$ as $n \to \infty$, but simulations suggest that the convergence is rather slow.

(iii) For simple random walk on $\mathbb{Z}^d$ with $d \geqslant 3$, does there exist an initial rotor configuration for which the rotor walk is recurrent?

Acknowledgements. We thank Omer Angel, David desJardins, Lionel Levine, Russell Lyons, Karola Mészáros, Yuval Peres, Oded Schramm and David Wilson for valuable discussions.

References

[1] O. Angel and A. E. Holroyd, *Rotor walks on general trees*, in preparation.

[2] O. Angel, A. E. Holroyd, J. Martin, and J. Propp, *Discrete low-discrepancy sequences*, preprint, 2009. Available as arXiv:0910.1077.

[3] J. Cooper, B. Doerr, T. Friedrich, and J. Spencer, *Deterministic random walks on regular trees*, pp. 766–772 in *Proceedings of the Nineteenth Annual ACM–SIAM Symposium on Discrete Algorithms (SODA 2008)*, Association for Computing Machinery (ACM), New York, 2008.

[4] J. Cooper, B. Doerr, J. Spencer, and G. Tardos, *Deterministic random walks*, pp. 185–197 in *Proceedings of the Eighth Workshop on Algorithm Engineering and Experiments and the Third Workshop on Analytic Algorithmics and Combinatorics* (R. Raman, R. Sedgewick, and M. F. Stallman, eds.), Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2006.

[5] J. Cooper, B. Doerr, J. Spencer, and G. Tardos, *Deterministic random walks on the integers*, European J. Combin. **28** (2007), no. 8, 2072–2090.

[6] J. N. Cooper and J. Spencer, *Simulating a random walk with constant error*, Combin. Probab. Comput. **15** (2006), no. 6, 815–822.

[7] D. Dhar, *Self-organized critical state of sandpile automaton models*, Phys. Rev. Lett. **64** (1990), no. 14, 1613–1616.

[8] B. Doerr and T. Friedrich, *Deterministic random walks on the two-dimensional grid*, Combin. Probab. Comput. **18** (2009), no. 1–2, 123–144.

[9] I. Dumitriu, P. Tetali, and P. Winkler, *On playing golf with two balls*, SIAM J. Discrete Math. **16** (2003), no. 4, 604–615.

[10] A. Engel, *The probabilistic abacus*, Ed. Stud. Math. **6** (1975), no. 1, 1–22.

[11] A. Engel, *Why does the probabilistic abacus work?*, Ed. Stud. Math. **7** (1976), no. 1–2, 59–69.

[12] A. E. Holroyd, L. Levine, K. Mészáros, Y. Peres, J. Propp, and D. B. Wilson, *Chip-firing and rotor-routing on directed graphs*, pp. 331–364 in *In and Out of Equilibrium 2* (V. Sidoravicius and M. E. Vares, eds.), no. 60 in Progress in Probability, Birkhäuser, Boston/Basel, 2008.

[13] M. Kleber, *Goldbug variations*, Math. Intelligencer **27** (2005), no. 1, 55–63.

[14] I. Landau and L. Levine, *The rotor-router model on regular trees*, J. Combin. Theory Ser. A **116** (2009), no. 2, 421–433.

[15] G. F. Lawler, *Intersections of Random Walks*, Birkhäuser, Boston/Basel, 1991.

[16] L. Levine, *The rotor-router model*, preprint, 2004. Available as arXiv:math/0409407.

[17] L. Levine and Y. Peres, *Scaling limits for internal aggregation models with multiple sources*, J. d'Analyse Math., to appear. Available as arXiv:0712.3378.
[18] L. Levine and Y. Peres, *Spherical asymptotics for the rotor-router model in* $\mathbb{Z}^d$, Indiana Univ. Math. J. **57** (2008), no. 1, 431–449.
[19] L. Levine and Y. Peres, *Strong spherical asymptotics for rotor-router aggregation and the divisible sandpile*, Potential Analysis **30** (2009), no. 1, 1–27.
[20] J. R. Norris. *Markov Chains*, no. 2 in Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge Univ. Press, Cambridge, UK, 1998. Reprint of 1997 original.
[21] V. B. Priezzhev, D. Dhar, A. Dhar, and S. Krishnamurthy, *Eulerian walkers as a model of self-organized criticality*, Phys. Rev. Lett. **77** (1996), no. 25, 5079–5082.
[22] J. Propp, *Rotor walk and Markov chains: Refinements and examples*, in preparation.
[23] F. Spitzer, *Principles of Random Walk*, no. 34 in Graduate Texts in Mathematics, Springer-Verlag, New York, second edition, 1976.
[24] D. B. Wilson, *Generating random spanning trees more quickly than the cover time*, pp. 296–303 in *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing (Philadelphia, 1996)*, Association for Computing Machinery (ACM), New York, 1996.

Microsoft Research, 1 Microsoft Way, Redmond, WA 98052, USA; and Dept. of Mathematics, University of British Columbia, 1984 Mathematics Rd., Vancouver, BC V6T 1Z2, Canada
E-mail address: holroyd@math.ubc.ca
URL: http://www.math.ubc.ca/~holroyd

Dept. of Mathematical Sciences, University of Massachusetts Lowell, 1 University Ave., Olney Hall 428, Lowell, MA 01854, USA
E-mail address: jpropp@cs.uml.edu
URL: http://faculty.uml.edu/jpropp

Contemporary Mathematics
Volume **520**, 2010

Approximate enumeration of self-avoiding walks

E. J. Janse van Rensburg

Abstract. Algorithms for the approximate enumeration of lattice self-avoiding walks are reviewed. Innovations in the approximate counting of such walks started with the invention of PERM (the pruned enhanced Rosenbluth method) in 1997. The recent generalization of the underlying Rosenbluth method (RM) to the GARM (generalized atmospheric RM), and to the GAS (generalized atmospheric sampling) algorithm, opens up exciting new possibilities for the approximate enumeration of walks. The implementation and use of these algorithms are described, as are associated results, including numerical data that show that high-quality estimates of the number of walks can be obtained.

1. Introduction

A difficult classical problem in polymer physics is the determination of c_n, the number of self-avoiding walks of n steps from the origin in the hypercubic lattice $\mathbb{Z}^d$. The sequence $\{c_n\}_{n\geqslant 0}$ provides information on polymer entropy and scaling [**6, 8, 9, 10, 35**]. An example of a self-avoiding walk is given in Figure 1.

An n-step walk in $\mathbb{Z}^d$ is defined as a sequence $\langle v_0, e_1, v_1, e_2, \ldots, e_n, v_n\rangle$ of vertices $v_0, v_1, v_2, \ldots, v_n$ and edges $e_1, e_2, \ldots, e_n$, such that the endpoints of the edge e_j are vertices v_{j-1} and v_j. The initial vertex v_0 is at the origin of the lattice. The walk is said to be self-avoiding if the vertices $v_0, v_1, v_2, v_3, \ldots, v_n$ are distinct. The length of the walk is its number of edges.

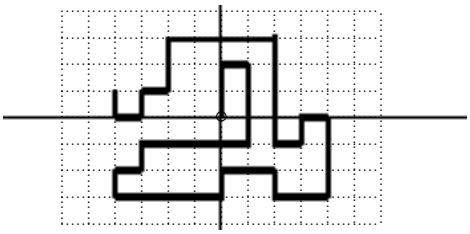


Figure 1. A self-avoiding walk in the square lattice $\mathbb{Z}^2$. The walk has initial (zeroth) vertex at the origin, and makes unit steps to nearest-neighbor vertices.

2000 *Mathematics Subject Classification.* Primary 82B41; Secondary 05A15, 65C05, 82B80.
Key words and phrases. Self-avoiding walk, statistical mechanics.
The author acknowledges support in the form of a Discovery Grant from NSERC (Canada).

The number c_n is generally difficult to determine, even for modest values of n. In two dimensions $c_0 = 1$, $c_1 = 4$, $c_2 = 12$, $c_3 = 36$, $c_4 = 100$; and c_n is known to $n = 71$ [**26**]. In three dimensions $c_0 = 1$, $c_1 = 6$, $c_2 = 30$, $c_3 = 150$; and c_n is known to $n = 30$ [**4**]. These sequences are `A0001411`, `A001412` in the OEIS [**45**]. In $\mathbb{Z}^d$,

$$d^n \leqslant c_n \leqslant 2d(2d-1)^{n-1}, \tag{1.1}$$

so that for all d, c_n grows exponentially with n. Enumerating walks exactly poses formidable numerical problems, and only sophisticated ideas and programming, and improvements in computing power following Moore's law, made possible the determinations of c_n in references [**26**] and [**4**].

A few more facts are known about c_n. For example, $c_{n+1} \geqslant c_n$, i.e., the sequence $\{c_n\}_{n\geqslant 0}$ is increasing [**40**]. By cutting a self-avoiding walk of length $n+m$ at its vertex v_n, two subwalks of lengths n and m are obtained. The number of choices for a walk of length $n+m$ is c_{n+m}, but the number of resulting pairs of subwalks is at most $c_n c_m$. Thus, c_n is a submultiplicative function on $\mathbb{N}$: $c_{n+m} \leqslant c_n c_m$. Together with the bounds in equation (1.1), this implies that the limit

$$\mu = \lim_{n\to\infty} (c_n)^{1/n} = \inf_{n\geqslant 0} (c_n)^{1/n} \tag{1.2}$$

exists, and that $d \leqslant \mu \leqslant 2d-1$ [**13, 15, 16**]. Kesten's pattern theorem [**29, 30**] shows that

$$\lim_{n\to\infty} \frac{c_{n+2}}{c_n} = \mu^2, \tag{1.3}$$

but it is not known whether $\lim_{n\to\infty}(c_{n+1}/c_n)$ exists.

An n-step lattice polygon in $\mathbb{Z}^d$ is a sequence $\langle v_0, e_1, v_1, e_2, \ldots, e_n, v_n\rangle$ of vertices $v_0, v_1, v_2, \ldots, v_n$ and edges $e_1, e_2, \ldots, e_n$, such that the endpoints of the edge e_j are vertices v_{j-1} and v_j, the vertices $v_0, v_1, v_2, \ldots, v_{n-1}$ are distinct, and $v_0 = v_n$. Equivalently, an n-step polygon may be defined as an $(n-1)$-step self-avoiding walk with final vertex v_{n-1} adjacent to the initial vertex v_0 at the origin. Polygons defined in this way have root vertex $v_0 = v_n$ at the origin, but the convention is to count polygons modulo translation in $\mathbb{Z}^d$. This is equivalent to removing the root.

Set p_n equal to the number of n-step polygons in $\mathbb{Z}^d$ modulo translation. For example, in the square lattice $\mathbb{Z}^2$ one has $p_4 = 1$, $p_6 = 2$, $p_8 = 7$, $p_{10} = 28$; and p_n is known to $n = 110$ [**25, 28**]. This sequence is `A002931` in the OEIS. It is a theorem [**14**] that the limit

$$\mu = \lim_{n\to\infty} (p_{2n})^{1/2n} = \sup_{n>0} (p_{2n})^{1/2n} \tag{1.4}$$

exists and is equal to the connective constant μ of self-avoiding walks, defined in equation (1.2). Note that in bipartite lattices such as $\mathbb{Z}^d$, the limit and supremum in equation (1.4) are taken through even integers, since $p_n = 0$ for odd values of n in such lattices. It is also known that

$$\lim_{n\to\infty} \frac{p_{n+2}}{p_n} = \mu^2, \tag{1.5}$$

a result which is due to Kesten [**29, 30**]; see reference [**35**] for a simpler proof.

1.1. Scaling of c_n. Equations (1.2) and (1.3) indicate that the leading behavior of c_n is exponential in n. The most natural correction to this growth is

a multiplicative power law. It is accordingly conjectured (see for example reference [**6**]) that to leading order c_n grows as

$$c_n \sim A\, n^{\gamma-1}\mu^n, \qquad n \to \infty. \tag{1.6}$$

Here $a_n \sim b_n$ signifies that $\lim_{n\to\infty}(a_n/b_n) = 1$, and γ is the critical exponent. If this asymptotic behavior of c_n holds, then it follows from equation (1.2) that $\gamma \geqslant 1$. The connective constant μ and critical exponent γ have been estimated to many digits (see below), but it has never been *proved* that equation (1.6) holds, except in dimensions $d \geqslant 5$ (where Hara and Slade [**17, 18**] have shown that it holds with $\gamma = 1$). The best rigorous upper bounds on c_n in dimensions $d \leqslant 4$ are all of the form $e^{Cn^q}\mu^n$ with $0 < q < 1$; see references [**29**] and [**35**].

Simple random walks in dimensions $d \geqslant 4$ are known to have few long range self-intersections (see [**32**, Theorem 3.3.2]). This suggests that the self-avoidance constraint may only affect a self-avoiding walk 'locally,' and in the scaling limit will not have a global effect on the structure of the walk. The result of Hara and Slade is that if $d \geqslant 5$ this is the case, equation (1.6) holding with $\gamma = 1$ because $\gamma = 1$ if there is no self-avoidance. If $d = 4$ it is believed that $\gamma = 1$ [**6**], but that the asymptotic behavior in equation (1.6) is modified by a logarithmic factor: $c_n \sim A\,(\log n)^{1/4}\,\mu^n$.

Dimension $d = 4$ is the 'upper critical dimension' for the self-avoiding walk since the behavior of this model changes radically for $d < 4$. If $d = 2$, it is believed that $\gamma = 43/32$ exactly; this is obtained from non-rigorous conformal field theory and Coulomb gas arguments [**5**]. In $d = 3$ dimensions γ is thought to be irrational, and is not expected to be the solution of any simple equation.

The *susceptibility* of the self-avoiding walk, $\chi(t)$, is the generating function of the sequence $\{c_n\}_{n\geqslant 0}$, defined by

$$\chi(t) = \sum_{n=0}^{\infty} c_n t^n, \tag{1.7}$$

which by equation (1.2) is convergent if $|t| < 1/\mu$ and diverges as $t \nearrow t_c = 1/\mu$. It follows from equation (1.6) that

$$\chi(t) \sim A'\,(t_c - t)^{-\gamma}, \qquad t \nearrow t_c, \tag{1.8}$$

with $A' = \Gamma(\gamma)/\mu^\gamma$. This explains the term 'critical exponent' for γ.

The estimation of μ and γ is a major motivation for computing the numbers c_n. Monte Carlo methods have also been used to estimate μ and γ. These methods are not nearly so accurate as methods based on exact enumeration, but they do provide independent checks on the latter.

1.2. Estimating μ and γ. The connective constant μ can be estimated from the exact values of c_n. In low dimensions such series data give the best estimates of μ; for the square lattice $\mathbb{Z}^2$ the result

$$\mu = 2.63815856 \pm 0.00000003, \tag{1.9}$$

as estimated in reference [**26**]. See references [**12, 27**] for additional results.

Monte Carlo estimates of μ and γ can be obtained from 'grand canonical' sampling of self-avoiding walks. In this approach, the probability of a generated walk having length n is $(c_n t^n)/\chi(t)$, where t is a parameter. From a sample of independently generated walks, μ and γ can be estimated. This was done in particular by

Berretti and Sokal [**1**], giving

$$\mu = 2.63820 \pm 0.00034 \quad \text{and} \quad \gamma = 1.352 \pm 0.031 \tag{1.10}$$

for $\mathbb{Z}^2$. See also reference [**38**], and a generalization of their algorithm in reference [**39**]. In reference [**38**] it is reported that

$$\mu = 2.638164 \pm 0.000014. \tag{1.11}$$

In addition, an analysis of series data on polygons [**28**] gives μ to very high precision:

$$\mu = 2.63815853034 \pm 0.00000000010. \tag{1.12}$$

Reference [**25**] can be consulted for a slight improvement of this estimate.

Less precise estimates of the connective constant μ are available for the cubic lattice $\mathbb{Z}^3$; see reference [**4**], where

$$\mu = 4.684043 \pm 0.000012 \tag{1.13}$$

is estimated from series data on walks obtained by the lace expansion. A more accurate estimate is due to Hsu and Grassberger [**20**], namely

$$\mu = 4.684038 \pm 0.000007, \tag{1.14}$$

and these two estimates are consistent within confidence intervals. A Monte Carlo estimate obtained from the 'canonical' sampling of polygons, and their atmospheric statistics [**23**], is

$$\mu = 4.68398 \pm 0.00016. \tag{1.15}$$

Reference [**19**] reports an (unpublished) estimate $\mu = 4.683907 \pm 0.000022$ due to A. J. Guttmann.

In two and three dimensions, $\gamma > 1$. As noted, in two dimensions conformal field theory [**5**] predicts the value $\gamma = 43/32 = 1.34375$, while in three dimensions the ϵ-expansion in a field-theoretic renormalization group approach [**33**] has been used to predict that $1.157 \lesssim \gamma \lesssim 1.160$. The best estimates of γ based on actual walks rely on exact enumeration data when $d = 2$ and Monte Carlo simulations when $d = 3$:

$$\gamma = \begin{cases} 1.343745 \pm 0.000015, & \text{if } d = 2, \text{ [26]}; \\ 1.1575 \pm 0.0006, & \text{if } d = 3, \text{ [2]}. \end{cases} \tag{1.16}$$

See reference [**4**] for more on results for $d = 3$. The estimate $\gamma = 1.1573 \pm 0.0002$ for $d = 3$ has been obtained in reference [**21**] by using the PERM algorithm on the Domb–Joyce model.

In this article we review algorithms and general ideas underlying the *approximate* enumeration of self-avoiding walks. The invention of the Rosenbluth method in 1955 [**44**] provided the first means of approximating c_n, i.e., estimating it rather than computing it exactly. This technique was extended to include pruning and enrichment, yielding PERM [**11**], and later the underlying Rosenbluth method was generalized to GARM [**43**]. The further generalization of GARM to GAS [**24**] is recent. We shall consider each of these algorithms briefly, to demonstrate approximate enumeration. Many of the underlying ideas are quite general and are applicable to other problems, such as the counting of self-avoiding polygons [**43**] and other lattice objects, and the counting of elements of finite groups [**7**].

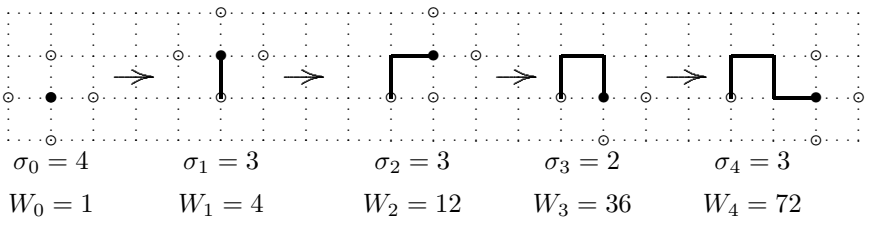


FIGURE 2. The Rosenbluth method for generating a walk. Starting with the trivial walk at the origin on the left, one repeatedly adds unoccupied nearest-neighbor vertices (denoted by ∘'s). As vertices are added to the walk, its weight is updated.

2. Approximate counting by Rosenbluth sampling

Rosenbluth sampling generates a sample of self-avoiding walks s_n of length n (or less), and corresponding weights W_n, such that the sample mean of the weights, $\langle W_n \rangle$, is an estimate of c_n. The walks in the sample are grown independently.

An individual walk s_n is grown as follows. Let s_0 be the trivial walk of length zero starting at the origin, and put $W_0 = 1$. Suppose that the walk has already grown to length j, i.e., that s_j and W_j are known. Determine the number of open lattice sites that are nearest neighbors of the endpoint of s_j; let σ_j be this number. With uniform probability, choose one of these sites and append it to s_j to obtain s_{j+1}, and update the weight by $W_{j+1} = \sigma_j W_j$. Repeat this process to grow a walk of any desired length, such as an n-step walk s_n with an associated weight W_n. The generation of a walk and its weight by this algorithm is illustrated in Figure 2.

There is the possibility of having $\sigma_{j_0} = 0$ for some j_0, i.e., of an endpoint of the walk having no adjacent unoccupied lattice sites. In this case the walk is *trapped*, and can no longer grow. The algorithm assigns zero weight W_j to any such walk for all $j > j_0$, i.e., effectively discards it; but it is still kept as a member of any sample of generated walks.

The probability that after n steps, this random growth produces any specified self-avoiding walk s_n of length n is

$$\Pr(s_n) = \prod_{j=0}^{n-1} \left[\frac{1}{\sigma_j(s_n)} \right] = \frac{1}{W_n(s_n)}. \tag{2.1}$$

The expectation of the weight W_n of a walk which is randomly grown as described, with target length n, is

$$\mathrm{E}\, W_n = \sum_{s_n \in S_n} \Pr(s_n) W_n(s_n) = \sum_{s_n \in S_n} 1 = c_n, \tag{2.2}$$

where S_n is the collection of all self-avoiding walks of length n. Therefore, the mean weight $\langle W_n \rangle$ of a sample of N walks $s_n^{(m)}$, $m = 1, \ldots, N$, that are grown independently, each with target length n, is an estimator for c_n. By the law of large numbers, $\langle W_n \rangle$ will satisfy

$$\langle W_n \rangle = \frac{1}{N} \sum_{m=1}^{N} W_n(s_n^{(m)}) \longrightarrow \mathrm{E}\, W_n = c_n, \qquad N \to \infty. \tag{2.3}$$

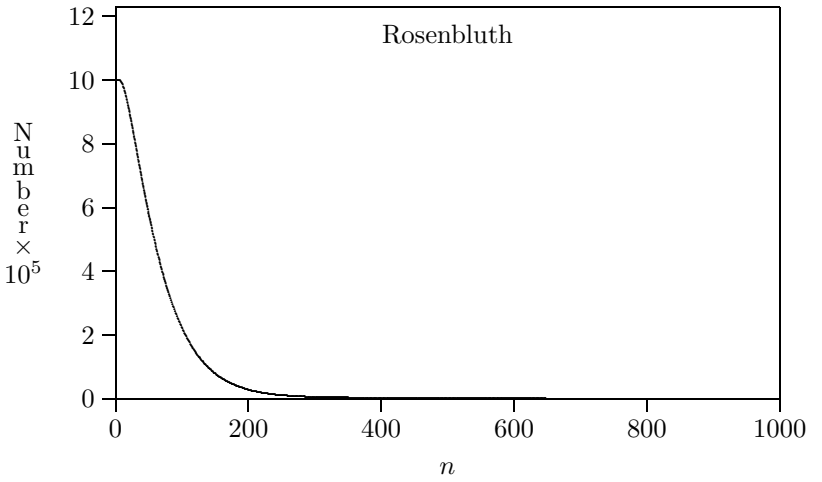


FIGURE 3. Attrition of started walks in the Rosenbluth method, for the square lattice $\mathbb{Z}^2$. Of one million started walks, only 20% survive to 100 steps and almost all are discarded by 200 steps.

To emphasize: in computing the sample mean $\langle W_n \rangle$, the 'discarded' walks if any, each of which has zero weight, must be taken into account.

The performance of this unbiased estimator for c_n is affected by the choice of N, the number of started walks, and the common target length n of the walks. For small values of n, attrition by trapping is not serious, but for larger n it poses a problem. In $\mathbb{Z}^2$ the attrition is mild for walks of lengths less than about 50 steps, but becomes serious for walks of lengths 100 steps or longer. Figure 3 illustrates the phenomenon. Attrition causes the variance of the estimator, relative to c_n^2, to increase with n. A more subtle problem is that the weights of non-discarded walks disperse over many orders of magnitude as the walk length n increases, which also increases the variance of the estimator. Eventually, the non-discarded portion of the sample becomes both small and dominated by a few walks with very large weights, and the estimate of c_n becomes unreliable.

Estimates of c_n for $\mathbb{Z}^2$ by the Rosenbluth method are given in Table 1. For one million started walks ($N = 1000000$), the mean weight was computed at each value of n. For small values of n the results are good, but deteriorate with increasing n as more walks are lost to attrition, and as the weights disperse. Over the range of values of n displayed in this table, the Rosenbluth estimates are fairly accurate, in part because, as Figure 3 shows, more than 50% of the started walks survive to length $n = 70$. The quality of the estimates decreases drastically when $n \gtrsim 100$, as attrition of the walks, and the dispersion of weights, take their toll. The reliable estimation of c_n for larger n would require an impractically large number of started walks N.

3. The pruned enriched Rosenbluth method (flatPERM)

3.1. From Rosenbluth to PERM. The two basic flaws in the Rosenbluth method are (1) the attrition due to trapping, and (2) the increasing dispersion of weights in the surviving walks. They can be addressed simultaneously by adding

pruning and *enrichment*, yielding PERM. (See reference [**11**]; cf. reference [**46**].) These additions turn the Rosenbluth algorithm into a dynamical simulation which is in a certain sense asymptotically stationary; but PERM differs from the Berretti–Sokal algorithm [**1**] in that walk lengths always increase, with pruning and enrichment being executed when weights go outside certain preset bounds.

In PERM, the size of the sample of self-avoiding walks is not fixed at its initial value N, the number of started walks, but evolves non-deterministically with n. The PERM algorithm is as follows. Recall that the weight of a walk s of length n generated by the Rosenbluth method is

$$W_n(s) = \prod_{j=0}^{n-1} \sigma_j(s). \tag{3.1}$$

n	c_n	Rosenbluth $\langle W_n \rangle$	flatPERM $\langle W_n \rangle$
0	1	1	1
1	4	4	4
2	12	12	12
3	36	36.0023	35.9999
4	100	100.005	99.9670
5	284	283.969	284.024
6	780	780.187	780.085
7	2172	2173.69	2172.49
8	5916	5918.86	5921.01
9	16268	16275.1	16279.4
10	44100	44098.6	44131.9
11	120292	120344	120370
12	324932	325039	325293
13	881500	881624	882273
14	2374444	2.37473×10^6	2.37759×10^6
15	6416596	6.41677×10^6	6.42589×10^6
16	17245332	1.72457×10^7	1.72694×10^7
17	46466676	4.64554×10^7	4.65682×10^7
18	124658732	1.24646×10^8	1.24896×10^8
19	335116620	3.35045×10^8	3.35598×10^8
20	897697164	8.97597×10^8	8.98656×10^8
25	123481354908	1.23661×10^{11}	1.23552×10^{11}
30	16741957935348	1.67595×10^{13}	1.67569×10^{13}
35	2252534077759844	2.25640×10^{15}	2.25608×10^{15}
40	300798249248474268	3.01374×10^{17}	3.01558×10^{17}
45	39992704986620915140	4.00832×10^{19}	4.01102×10^{19}
50	529279466872483720664	5.30379×10^{21}	5.31268×10^{21}
55	6985017002775819546740	6.99913×10^{23}	7.01539×10^{23}
60	918958360250562146340477	9.24390×10^{25}	9.21359×10^{25}
65	120662711363467257265478106	1.21868×10^{28}	1.20728×10^{28}
70	158078467825057188201748024363	1.58563×10^{30}	1.58013×10^{30}
71	419089302090393505461912000591	4.20159×10^{30}	4.18935×10^{30}

TABLE 1. Approximate enumeration in 2-D (Rosenbluth and flatPERM).

At each length n, introduce an upper cutoff or threshold T_n, and if $W_n(s) > T_n$ for any walk s in the sample, enrich the sample by a adding a copy of s to it, at the same time reducing the weight $W_n(s)$ by a factor of two. There will then be two copies of s, but each will have weight $W_n(s)/2$; and the sample size will be incremented by one. In PERM the 'sample mean' $\langle W_n \rangle$ is redefined to equal the sum of the weights of the walks in the sample, divided by the *original* size of the sample, N, rather than by its current value. With this redefinition, the enrichment will not affect the sample mean $\langle W_n \rangle$, which will continue to be an unbiased estimator for c_n. Because the two copies of s will grow randomly along different trajectories, the enrichment, including the replacement of a large weight by two smaller ones, will have the effect of reducing the dispersion of the weights.

The problem of walks with small weight is dealt with by introducing a lower cutoff or threshold t_n at each length n, at which pruning will take place. If a walk s has grown to length n and is of weight $W_n(s) < t_n$, then 'discard' it as if it were trapped (i.e., stop growing it and change its weight to zero, though keeping it in the sample) with probability $1 - 1/q$, where q is a parameter of the algorithm. If the walk is not discarded (with probability $1/q$), then increase its weight by a factor of q. (Values $q \approx 2$ are typical.) The zero weights of discarded (i.e., pruned) walks will be kept in the formula for the sample mean $\langle W_n \rangle$, as usual.

The thresholds t_n and T_n should be adjusted during the random evolution of the sample. The initial values are usually $t_n = 0$ and T_n some large constant, but they should be chosen to increase rapidly with n, due to the expected rapid growth of the weights. It is possible to set $T_n/t_n = K$ eventually, for $K > 1$ some constant.

The defining feature of PERM is the scheme for adjusting t_n and T_n with n, which affects every growing walk in the sample. It is performed in a way that depends on the *entire sample*, thereby coupling the evolutions of the walks to one another. Suppose that from N started walks, a sample of walks of length n has been generated, and that the mean sample weight, defined as above, equals $\langle W_n \rangle$. The thresholds t_n and T_n are adjusted according to

$$t_n = c\langle W_n \rangle \qquad \text{and} \qquad T_n = C\langle W_n \rangle, \tag{3.2}$$

for some $c < 1$ and $C > 1$ satisfying $C/c = K$. With $K \approx 5$, and appropriate choices of c, C, the dispersion of the weights of the walks in the sample will typically be reduced to about one order of magnitude. Also, the enrichment will increase the number of completed walks, thereby reducing the underlying Rosenbluth attrition.

As described, PERM is a significant advance in the approximate enumeration of self-avoiding walks; see reference [**11**] for more details. PERM has also been used in generating samples of lattice animals and trees [**22**].

3.2. Flat-histogram PERM (flatPERM). The PERM algorithm can be improved by replacing the pruning and enrichment based on thresholds, however chosen, by *continual* pruning and enrichment. In this alternative algorithm, the weight of each growing walk of length n in the sample is kept as close as possible to the current estimate of c_n, to minimize the variance of the estimator $\langle W_n \rangle$.

This is accomplished as follows. Suppose that at length n, some number N_n of walks are growing. Let them be denoted $s^{(m)}$, $m = 1, \ldots, N_n$. For each m, compute the ratio

$$r_m = \frac{W_n^{(m)}}{\langle W_n \rangle}, \tag{3.3}$$

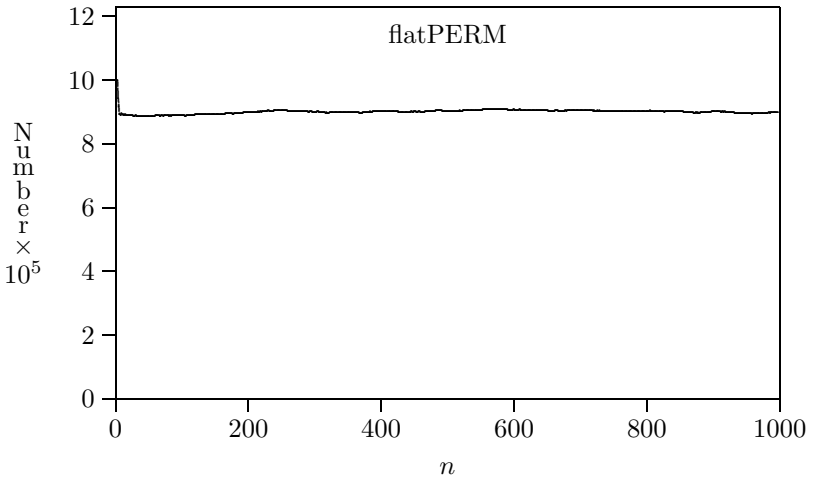


FIGURE 4. Attrition of started walks in flatPERM. One million walks were started, and the number of non-discarded walks at each length n is plotted. After some initial attrition, a nearly constant number of walks is obtained. By comparison with the Rosenbluth data in Figure 3, flatPERM is highly effective in self-tuning.

where $W_n^{(m)}$ is the weight of $s^{(m)}$, and $W_n^{(m)}$ is included in the calculation of the sample mean $\langle W_n \rangle$. If $r_m > 1$ then the walk $s^{(m)}$ is a candidate for enrichment. If r_m is also larger than the number of possible ways of extending the walk (i.e., $r_m > \sigma_n(s^{(m)})$) then put $c = \min\{\lfloor r_m \rfloor, \sigma_n(s^{(m)})\}$, and enrich $s^{(m)}$ in the sample by making c copies of it, and by reducing its weight by a factor of c.

On the other hand, if $r_m < 1$, then the walk $s^{(m)}$ has weight smaller than the sample mean, so with probability $1 - r_m$, prune it (i.e., 'discard' it by setting its weight to zero, though keeping it in the sample). If it is not pruned, then increase its weight by a factor $1/r_m$.

Note that pruning and enrichment are performed on any walk *after* it is included in the computation of the sample mean $\langle W_n \rangle$, i.e., the estimate of c_n. The estimate $\langle W_n \rangle$ for c_n may initially be very wrong, but with the preceding scheme for pruning and enrichment, it will typically improve with increasing n. The algorithm is, in a sense, self-tuning.

This algorithm is called flatPERM because it produces a sample of length-n walks, the size of which is (eventually) roughly 'flat' as a function of n. On average, for every walk pruned, one is enriched; and moreover, the variance of the walk weights is tightly controlled. These two features are improvements over Rosenbluth sampling, where attrition makes the generation of long walks very difficult. In Figure 4 the attrition of walks in flatPERM is shown for one million started walks in $\mathbb{Z}^2$. In this run the attrition is less than 10% even for $n = 1000$, and the number of walks in the sample is more or less flat over the entire range of n. This gives a large sample of long walks, compared to the Rosenbluth method.

The flatPERM algorithm requires more CPU time for a given number of started walks, compared to the Rosenbluth method, because more walks are completed. The flatPERM data presented in Figure 4 required roughly twenty times the CPU time of the Rosenbluth data in Figure 3.

As mentioned, when n is small a flatPERM estimate of c_n by the sample mean $\langle W_n \rangle$ will usually be poor. This leads to a decrease in the initial effectiveness of pruning and enrichment, which is what causes the initial attrition seen in Figure 4. To deal with this difficulty, one should increase N, the number of started walks. Observe that a sample of walks can be enlarged incrementally: when N is increased to $N+1$, each walk in the former sample can be used in the latter. Hence it is reasonable to restrict the walk length n until N has grown sufficiently large. One particular scheme limits n to αN, where N is the number of walks being generated. As the algorithm runs, N is incremented until n reaches its target value, at which point the restriction is removed. The constant α is typically chosen to be between 1 and 10.

3.3. Microcanonical sampling with flatPERM. The flatPERM algorithm was originally applied to the 'microcanonical' sampling of self-avoiding walks [**41**], to produce estimates of $c_{n,e}$. This is the number of walks of length n and 'energy' e. The energy of a walk depends on the model being considered: it may be the number of nearest-neighbor contacts between vertices of the walk, the number of vertices of the walk adsorbed in a line or plane, etc.

In microcanonical flatPERM, with every edge that is added to a growing walk, the weight $W_{n,e}$ of the walk is updated as usual, and its energy is also updated. The mean weight of the walks in the sample that have energy e gives an estimate of $c_{n,e}$. Pruning and enrichment take place as usual, at each energy. The number of walks in the sample, for each length n and energy e, eventually becomes roughly independent of (n, e), over a wide range. That is, it becomes 'flat' in both directions.

This approach has also been applied in reference [**31**], to a model of pulled adsorbing walks.

4. Generalized atmospheres and GARM

GARM (see reference [**43**]), the generalized atmospheric Rosenbluth method, is a generalization of the original Rosenbluth method (RM), which employs more general ways of growing self-avoiding walks. In the RM, a walk is grown by adding an edge to its endpoint. In GARM, it is grown by adding an edge in its *positive atmosphere* (or in some variants, in its neutral atmosphere). Just as the RM can be enhanced to yield flatPERM, so can GARM be enhanced to yield flatGARM.

4.1. Endpoint and generalized atmospheres. There are various ways in which 'atmospheres' of a self-avoiding walk can be defined. The first explicit definitions can be found in reference [**42**], but they have been used in generating (samples of) self-avoiding walks since the Rosenbluth method was introduced in 1955 [**44**]. More can be found in references [**23, 24, 43**].

Endpoint Atmospheres: The set of lattice edges which can be added to the final vertex of a self-avoiding walk to extend its length by 1 is its *positive endpoint atmosphere*. The size of this set will be denoted by a^e_+. In Figure 5, the positive endpoint atmosphere of the walk ending in the arrowhead consists of the pair of bold edges. The last edge of the walk (indicated by the arrowhead) can be deleted to create a self-avoiding walk of length reduced by one. The set of possible deletions is the *negative endpoint atmosphere* of the walk. Its size will be denoted by a^e_-. The trivial walk of length zero has $a^e_- = 0$, but every other walk has $a^e_- = 1$.

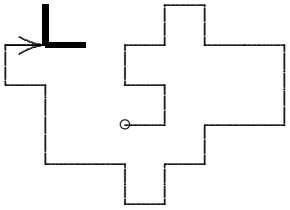


FIGURE 5. A self-avoiding walk in $\mathbb{Z}^2$, with its positive endpoint atmospheric edges indicated in bold. The last step in the walk is its negative endpoint atmospheric edge. In this case, $a_+^e = 2$ and $a_-^e = 1$.

A *neutral endpoint atmosphere* can also be defined, and in Figure 6 the definition is indicated. The bold edge is the final step in the walk, and it can be changed to either of the dashed edges. The latter make up the neutral endpoint atmosphere of the walk. In general, its size will be denoted by a_0^e.

Generalized Atmospheres: Somewhat different definitions of atmospheres (the positive and negative *generalized* atmospheres) are given in Figures 7 and 8. The size of the former is denoted by a_+^g. (For instance, the trivial walk in $\mathbb{Z}^2$ has $a_+^g = 4$.) The size of the latter is denoted by a_-^g. For any nontrivial walk the final edge is a negative generalized atmospheric edge, hence $a_-^g \geqslant 1$. Neutral generalized atmospheres, of size a_0^g, can be defined in any of several ways [**24**].

Alternative definitions of the three types of atmosphere may be used; see for example references [**23, 43**].

4.2. Atmospheric moves in self-avoiding walks. Positive, neutral and negative atmospheres specify possible *atmospheric moves* that add edges, rearrange edges, or subtract edges from a walk. These include but are not limited to endpoint atmospheric moves.

Several Monte Carlo algorithms for generating random samples of self-avoiding walks are based on atmospheric moves, including the Rosenbluth and PERM algorithms of Sections 2 and 3, and also the Berretti–Sokal algorithm [**1**]. The pivot

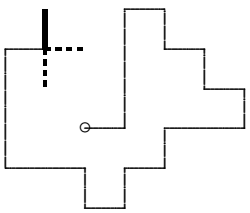


FIGURE 6. A self-avoiding walk in $\mathbb{Z}^2$, with its final step indicated in bold. The final step may be changed to either of the two dashed edges, so they make up the neutral endpoint atmosphere. In this case, $a_0^e = 2$.

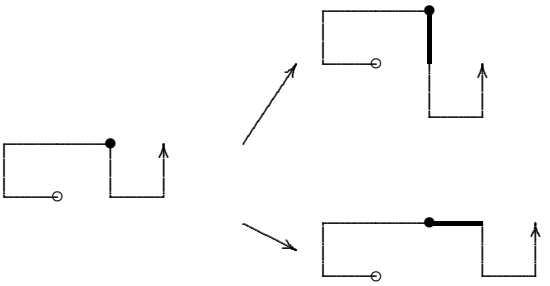


FIGURE 7. A self-avoiding walk in $\mathbb{Z}^2$ and a small part of its positive generalized atmosphere. There are two ways in which an edge can be inserted in the walk at the vertex $\bullet$ on the left. The outcomes are shown on the right, with the inserted edge indicated in bold. The edges that can be inserted in this way make up the positive generalized atmosphere, of size a_+^g. In this case, $a_+^g = 15$.

algorithm (see references [**34, 37**]) is based on neutral atmospheric moves defined using pivots.

Observe that any positive atmospheric move defined above (whether endpoint or generalized) can be reversed to yield a negative atmospheric move. Similarly, any negative atmospheric move can be reversed to yield a positive one. Furthermore, given an arbitrary self-avoiding walk, one can reduce it to the trivial 'state' (i.e., walk) by a sequence of negative atmospheric moves. That is, the trivial state communicates with any given state s, in an infinite *state space* of self-avoiding walks, along a sequence of positive atmospheric moves.

Define $S(n)$ to be the set of self-avoiding walks of length n, started at the origin, so that the state space is $S = \cup_{n=0}^{\infty} S(n)$. Then the collection of positive atmospheric moves defines a multivalued map $f\colon S(n) \to S(n+1)$, as shown schematically in

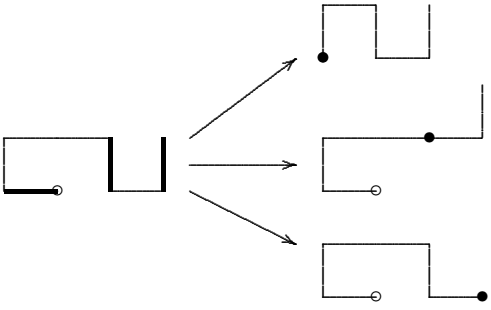


FIGURE 8. A self-avoiding walk in $\mathbb{Z}^2$ and its negative generalized atmosphere. By deleting and contracting one of the three bold edges in the walk on the left, the walks on the right are obtained. The edges that can be deleted and contracted make up the negative generalized atmosphere, of size a_-^g. In this case, $a_-^g = 3$.

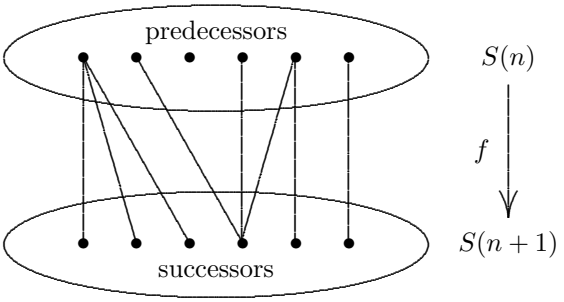


FIGURE 9. Positive generalized moves, whether endpoint or generalized, define a multivalued map $f\colon S(n) \to S(n+1)$ that takes walks of length n to walks of length $n+1$. The map f may be represented as a graph, as shown. Some states in $S(n)$ have no successors; they are trapped walks.

Figure 9. Iterating this map, i.e., repeatedly performing positive generalized atmospheric moves, gives a method of growing a walk, which is illustrated in Figure 10. Using only positive *endpoint* moves gives the Rosenbluth method. Using positive *generalized* moves, such as those shown in Figure 7, gives the GARM algorithm, which is explained further in the following subsection.

The edges in Figure 9 are called *linkages* between $S(n)$ and $S(n+1)$. Consider a walk $s \in S(n)$ together with its associated atmospheres of sizes $a_+(s)$, $a_0(s)$ and $a_-(s)$ (these may be atmospheres of any suitable kind). Then the total number of linkages between $S(n)$ and $S(n+1)$ satisfies

$$\#\,\text{Linkages} = \sum_{s\in S(n)} a_+(s) = \sum_{s\in S(n+1)} a_-(s). \tag{4.1}$$

Dividing this by $c_n = \sum_{s\in S(n)} 1$ reveals that the expected positive and negative atmosphere sizes, evaluated with respect to the uniform distributions on $S(n)$ and $S(n+1)$, satisfy

$$\mathrm{E}_{(n)}\, a_+ = \frac{1}{c_n} \sum_{s\in S(n)} a_+(s) = \frac{c_{n+1}}{c_n} \frac{1}{c_{n+1}} \sum_{s\in S(n+1)} a_-(s) = \frac{c_{n+1}}{c_n}\, \mathrm{E}_{(n+1)}\, a_-, \tag{4.2}$$

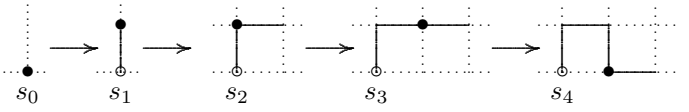


FIGURE 10. Drawing repeatedly from positive generalized atmospheres. Starting from the trivial walk, at each iteration a vertex • is chosen and a positive atmospheric edge is inserted to generate the next state in the sequence. The move is at each step chosen uniformly from those available.

so that

$$\frac{\mathrm{E}_{(n)}\, a_+}{\mathrm{E}_{(n+1)}\, a_-} = \frac{c_{n+1}}{c_n}. \tag{4.3}$$

Hence, an estimate of the ratio c_{n+1}/c_n (asymptotically unbiased) can be computed from random samples of self-avoiding walks of lengths n and $n+1$, generated by any convenient method; such as, for instance, the pivot algorithm mentioned above. This is one particular method of approximate enumeration that is based on atmospheric statistics. See references [**23, 42**] for more details and applications.

4.3. Generalized atmospheric Rosenbluth method (GARM). GARM uses positive *generalized* atmospheric moves, of the type shown in Figure 7, in a Rosenbluth-style method. (And optionally neutral moves as well; see below.) GARM generates a random sample of self-avoiding walks, with each walk grown independently as a sequence of ever-lengthening walks ('states'), and an associated sequence of weights. A 'counting formula' is used to estimate c_n. In any version of GARM it is required that (1) the trivial state (walk of length zero) communicate with each self-avoiding walk s by a sequence of positive atmospheric moves, (2) each positive atmospheric move be reversible to a negative one, and vice versa.

GARM is implemented as follows. Let s_0 be the trivial walk of length zero. Once s_j is known, find s_{j+1} by randomly and uniformly selecting a move from the positive atmosphere of s_j. In pictorial terms, simply apply the multivalued map f of Figure 9 an arbitrarily large number of times, obtaining as in Figure 10 a sequence of walks $s_0, s_1, s_2, \ldots$ that is a transient Markov chain (the associated weights $W_0, W_1, W_2, \ldots$ will be defined below). This conceptually infinite sequence, denoted ϕ, can be viewed as a point in a probability space. One may write $s_j(\phi)$ and $W_j(\phi)$. The symbol ϕ_L will signify the finite sequence $(s_0, \ldots, s_L)$, which may be a prefix of an infinite sequence ϕ.

Since the state s_{j+1} is obtained from s_j by randomly and uniformly selecting a positive atmospheric move, the conditional probability that s_{j+1} follows s_j in ϕ is

$$\Pr(s_{j+1}|s_j) = \frac{1}{a_+(s_j)}. \tag{4.4}$$

The probability of generating any specified sequence ϕ_L, from among those of a specified length L, is thus

$$\Pr(\phi_L) = \prod_{j=0}^{L-1} \Pr(s_{j+1}(\phi_L)|s_j(\phi_L)) = \prod_{j=0}^{L-1} \left[\frac{1}{a_+(s_j(\phi_L))}\right]. \tag{4.5}$$

The probability that ϕ_L terminates in a specified state $\tau \in S(L)$, i.e., that $s_L = \tau$, is

$$\Pr(s_L(\phi_L) = \tau) = \sum_{\substack{\phi_L \text{ s.t.} \\ s_L(\phi_L)=\tau}} \prod_{j=0}^{L-1} \left[\frac{1}{a_+(s_j(\phi_L))}\right]. \tag{4.6}$$

Weights $W_L(\phi)$ are assigned to each generated sequence ϕ of states, i.e., walks, by

$$W_L(\phi) = W_L(\phi_L) = \prod_{j=0}^{L-1} \left[\frac{a_+(s_j(\phi_L))}{a_-(s_{j+1}(\phi_L))}\right]. \tag{4.7}$$

Each factor in this product is the positive atmosphere size of the *current* state, divided by the negative atmosphere size of the *next* state. This definition generalizes the definition used in the original Rosenbluth method.

Note that in GARM, as in the RM, a walk may become trapped due to having an empty positive atmosphere, in which case it is effectively discarded (subsequent weights are set to zero, as in the RM). But if generalized atmospheres are used, this is rare. Attrition is not the serious problem it is in Rosenbluth sampling.

LEMMA 4.1. *Let τ be any state, i.e., any self-avoiding walk, of length n. Then the expectation of $W_n(\phi)$, times the indicator of the event that the length-n walk $s_n(\phi)$ is τ, equals unity. That is,*

$$\mathrm{E}\left[W_n(\phi)1_{s_n(\phi)=\tau}\right] = \sum_{\substack{\phi_n \text{ s.t.} \\ s_n(\phi_n)=\tau}} W_n(\phi_n)\Pr(\phi_n) = \sum_{\substack{\phi_n \text{ s.t.} \\ s_n(\phi_n)=\tau}} \prod_{j=0}^{n-1} \left[\frac{1}{a_-(s_{j+1}(\phi_n))}\right] = 1,$$

the expectation being computable over sequences ϕ_n of length n.

PROOF. The quantity

$$\sum_{\substack{\phi_n \text{ s.t.} \\ s_n(\phi_n)=\tau}} \prod_{j=0}^{n-1} \left[\frac{1}{a_-(s_{j+1}(\phi_n))}\right] \tag{4.8}$$

equals the probability that the walk τ can be reduced to the trivial walk s_0 composed of a single vertex by executing n randomly and uniformly chosen *negative* atmospheric moves. This probability is equal to unity, by the reversibility assumption satisfied by any version of GARM. □

Summing $\mathrm{E}\left[W_n(\phi)1_{s_n(\phi)=\tau}\right]$ over the c_n states $\tau \in S_n$ gives the 'counting formula'

$$\mathrm{E}\,W_n = \sum_{\tau\in S_n} \mathrm{E}\left[W_n(\phi)1_{s_n(\phi)=\tau}\right] = \sum_{\tau\in S_n} \prod_{j=0}^{n-1} \left[\frac{1}{a_-(s_{j+1}(\phi_n))}\right] = c_n, \tag{4.9}$$

the expectation being computable as an expectation over sequences ϕ_n of length n. Therefore, the mean weight $\langle W_n\rangle$ in a sample of walks grown by GARM, which is an estimator for $\mathrm{E}\,W_n$, is one for c_n.

The variances of weights in GARM do not increase as quickly with n as in the Rosenbluth method, but they do tend to increase. Much as in the RM, one may introduce pruning and enrichment moves to reduce the variance. GARM with pruning and enrichment proceeds, like PERM, by tracking the weight of a sequence of states being generated. If this weight grows too small, the sequence can be pruned, and if the weight grows too large, the sequence can be enriched.

If the implementation of GARM uses only endpoint atmospheric moves, then GARM will reduce to the RM. More generally, if GARM is implemented with pruning and enrichment, then it will generalize PERM; and it will reduce to PERM if it uses only endpoint atmospheric moves.

In Table 2, GARM estimates of c_n for the square lattice $\mathbb{Z}^2$ are compared to exact enumeration data from reference [**26**]. There were one million started sequences of states (i.e., self-avoiding walks), which were allowed to grow until they reached length 72, yielding an estimate $\langle W_n\rangle$ of c_n for each n.

The calculation of generalized atmospheres for walks of length n requires $O(n)$ computational effort; it follows that the effort in generating a single sequence of states, the final state of which is a walk of length n, is $O(n^2)$. This shows that GARM slows down significantly with increasing length of the generated walks. Improving this may be possible, by using a technique similar to the implementation of the pivot algorithm by Clisby [**3**].

GARM can also be implemented to include neutral atmospheric moves, provided that each neutral move is reversible to another such move. With this modification, the algorithm will choose randomly and uniformly from the available positive and neutral atmospheric moves, and the length of the walks in a sequence of generated walks will not necessarily increase monotonically. The probability of

n	c_n	GARM $\langle W_n \rangle$	flatGARM $\langle W_n \rangle$
0	1	1	1
1	4	4	4
2	12	12	12
3	36	36.0012	36.0049
4	100	100.0143	100.020
5	284	284.054	284.084
6	780	780.448	779.772
7	2172	2174.17	2172.42
8	5916	5919.34	5916.98
9	16268	16279.7	16263.7
10	44100	44121.1	44087.1
11	120292	1.20385×10^5	1.20297×10^5
12	324932	3.25203×10^5	3.24837×10^5
13	881500	8.82099×10^5	8.81427×10^5
14	2374444	2.37534×10^6	2.37490×10^6
15	6416596	6.41833×10^6	6.41986×10^6
16	17245332	1.72509×10^7	1.72540×10^7
17	46466676	4.64541×10^7	4.65168×10^7
18	124658732	1.24643×10^8	1.24788×10^8
19	335116620	3.35027×10^8	3.35552×10^8
20	897697164	8.96830×10^8	8.98939×10^8
25	123481354908	1.23346×10^{11}	1.23682×10^{11}
30	16741957935348	1.66939×10^{13}	1.67753×10^{13}
35	2252534077759844	2.23868×10^{15}	2.25720×10^{15}
40	300798249248474268	2.98010×10^{17}	3.01309×10^{17}
45	39992704986620915140	3.94580×10^{19}	4.00419×10^{19}
50	5292794668724837206644	5.21079×10^{21}	5.30212×10^{21}
55	6985017002775819546746O4	6.90023×10^{23}	6.99201×10^{23}
60	91895836025056214634047716	9.06769×10^{25}	9.20095×10^{25}
65	120662711363467257265478l0652	1.18578×10^{28}	1.20736×10^{28}
70	158078467825057188201748024363O	1.54860×10^{30}	1.58136×10^{30}
71	419089302090393505461912000591O	4.08355×10^{30}	4.19364×10^{30}

TABLE 2. Approximate enumeration in 2-D (GARM and flatGARM).

generating a sequence ϕ_L of walks, from among those of a specified length L, is

$$\prod_{j=0}^{L-1} \frac{1}{a_+(s_j(\phi_L)) + a_0(s_j(\phi_L))}, \tag{4.10}$$

and the associated weight is accordingly chosen to be

$$W_L(\phi) = W_L(\phi_L) = \prod_{j=0}^{L-1} \frac{a_+(s_j(\phi_L)) + a_0(s_j(\phi_L))}{a_-(s_{j+1}(\phi_L)) + a_0(s_{j+1}(\phi_L))}. \tag{4.11}$$

The expectation of $W_L(\phi_L)$, times the indicator of the event that ϕ_L ends on a state that is a length-n walk, when summed over all possible values L of the length of ϕ_L (i.e., $L \geqslant n$), will equal c_n. The proof of this is similar to the proof of Lemma 4.1. Note that if neutral atmospheric moves are possible, then the length L is no longer deterministically equal to n.

4.4. Flat-histogram GARM (flatGARM). Pruning and enrichment can be incorporated in GARM, much as in the original Rosenbluth method, to yield a 'flat' version of GARM called flatGARM [**43**]. In flatGARM as in flatPERM, the number of walks in a sample evolves non-deterministically, and the evolutions of the walks are coupled to one another.

If the j-th state $s_j = s_j(\phi)$ in a sequence ϕ of states (i.e., self-avoiding walks) generated by flatGARM has weight $W(s_j)$, then the algorithm calculates

$$r = \frac{W(s_j)}{\langle W_j \rangle}, \tag{4.12}$$

where $\langle W_j \rangle$ is the mean weight of the j-th state over all sequences in the sample being generated (including the current sequence). Calculating $W(s_j)$ to determine $\langle W_j \rangle$ requires the size of the negative atmosphere of the next state s_{j+1}, which has not been constructed yet. This difficulty may be overcome by extrapolating from the negative atmosphere size of the current state. For endpoint atmospheres one has exactly $a_-(s_{j+1}) = a_-(s_j) = 1$ unless s_j is the trivial state, and for generalized atmospheres it is not unreasonable to estimate $a_-(s_{j+1})$ by $a_-(s_j) + 1$. This estimate works well in practice.

Once the parameter r of equation (4.12) has been computed, pruning and enrichment are implemented as follows. If $r > 1$ then enrich the current sequence: Compute $c = \lceil r \rceil$ with probability $r - \lfloor r \rfloor$, and $c = \lfloor r \rfloor$ otherwise. Make c copies of the sequence, each with weight $W(s_j)/c$, and continue growing from each. If on the other hand $r < 1$, then with probability r, retain the sequence and update the weight $W(s_j)$ to $W(s_j)/r$. Otherwise, prune the sequence (with probability $1 - r$). Note that as in flatPERM, the calculation of any 'sample mean' is understood to include a division by the *original* sample size, i.e., the number of started walks. By this convention, the sample mean is unaffected by pruning and enrichment.

Running flatGARM produces some initial attrition of the started sequences due to pruning, but the parameter r is designed to produce a sample size that is more or less constant with n: the pruned sequences are eventually replaced by enriched ones. A major advantage over flatPERM is that correlations are suppressed by the enrichment process. While endpoint atmospheric moves in flatPERM leave the walk unchanged up to the enrichment point, generalized atmospheric moves quickly

update edges even along the early part of the walk, so that correlations between enriched copies soon become small.

Table 2 shows the results of a flatGARM estimation of c_n, based on one million started sequences, the growth of which reached states (i.e., walks) of length 72. These results are compared to the results of a GARM estimation, based on the same number of started walks. Clearly, flatGARM outperforms GARM for longer walks. Their CPU times are the same to within a few percent.

Microcanonical flatGARM sampling is implemented similarly to microcanonical flatPERM sampling, as described in Section 3.3. It also yields a flat distribution over length and energy.

5. Generalized atmospheric sampling (GAS)

5.1. The GAS algorithm. In Generalized Atmospheric Sampling (GAS), negative atmospheric moves are added to the mix of positive (and possibly neutral) atmospheric moves in GARM; and also weights $W_L(\phi) = W_L(\phi_L)$ that are qualitatively different from the GARM weights are introduced. The addition of negative moves prevents trapping, and has the potential of making the Markov chain recurrent (as in the Berretti–Sokal algorithm [**1**]), rather than transient.

If s is a self-avoiding walk, let $a_+(s)$, $a_0(s)$ and $a_-(s)$ be the sizes of its positive, neutral and negative atmospheres. As in GARM, reversibility is assumed: it is assumed that every positive atmospheric move is reversible to a negative atmospheric move, and vice versa, and that every neutral move is reversible to another such move. Also, any two walks must communicate along at least one sequence of moves; that is, the possible moves are assumed to act irreducibly on the infinite state space $S = \cup_{n=0}^{\infty} S(n)$ of all self-avoiding walks, and preferably aperiodically as well.

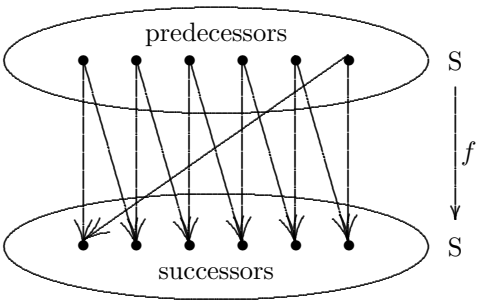


FIGURE 11. Generalized atmospheric moves in GAS define a multivalued map $f\colon S \to S$, represented by a digraph.

The possible atmospheric moves define linkages in S, as shown in Figure 11. They define a multivalued map $f\colon S \to S$, as in GARM. Iterating this map gives a method for growing a walk, though its length may decrease as well as increase.

Let $s_0 \in S$ be an initial state for the algorithm. Successors of s_j are states s_{j+1} that can be reached from s_j by performing a (positive, neutral or negative) atmospheric move. Once s_j has been selected, s_{j+1} is selected from amongst its successors, with nonzero probabilities not yet specified. This defines an irreducible, aperiodic Markov chain on S, yielding a sequence $s_0, s_1, s_2, \ldots$ of states.

The state space of the GAS algorithm is the infinite set S, but the algorithm can be modified as follows so that the state space is a finite set. Simply define all walks in S of length $n_{\max}$ to have empty positive atmospheres, and use $S_*(n_{\max}) = \cup_{n=0}^{n_{\max}} S(n)$ as a replacement for S.

The transition probabilities of the Markov chain remain to be specified. The following choice is made. Suppose that a state s_j has been reached. For some parameter $\beta > 0$, GAS will perform an atmospheric move with probabilities:

$$P_+(s_j) = \Pr(\text{positive move}) = \frac{\beta a_+(s_j)}{a_-(s_j) + a_0(s_j) + \beta a_+(s_j)}; \tag{5.1a}$$

$$P_0(s_j) = \Pr(\text{neutral move}) = \frac{a_0(s_j)}{a_-(s_j) + a_0(s_j) + \beta a_+(s_j)}; \tag{5.1b}$$

$$P_-(s_j) = \Pr(\text{negative move}) = \frac{a_-(s_j)}{a_-(s_j) + a_0(s_j) + \beta a_+(s_j)}, \tag{5.1c}$$

which are normalized to sum to unity. From among the possible moves of the selected type, a uniform choice is made.

Weights $W_L(\phi)$ are assigned to each generated sequence ϕ of states, i.e., self-avoiding walks, by

$$W_L(\phi) = W_L(\phi_L) = \left[\prod_{j=0}^{L-1} \left[\frac{a_-(s_j) + a_0(s_j) + \beta a_+(s_j)}{a_-(s_{j+1}) + a_0(s_{j+1}) + \beta a_+(s_{j+1})}\right]\right] \prod_{j=0}^{L-1} \beta^{\sigma(s_j, s_{j+1})}, \tag{5.2}$$

where

$$\sigma(s_j, s_{j+1}) = \begin{cases} -1, & \text{if } s_{j+1} \text{ follows } s_j \text{ through } a_+; \\ +1, & \text{if } s_{j+1} \text{ follows } s_j \text{ through } a_-; \\ 0, & \text{if } s_{j+1} \text{ follows } s_j \text{ through } a_0. \end{cases} \tag{5.3}$$

These are quite different from the Rosenbluth and GARM weights; for comparison, see equations (4.7) and (4.11). Let $P(\phi_L)$ be the number of positive atmospheric moves in ϕ_L, and $N(\phi_L)$ be the number of negative moves; so that $P(\phi_L) - N(\phi_L)$ is equal to n_L, the length of the final state s_L in ϕ_L. With these definitions the above product telescopes to give the much simplified expression

$$W_L(\phi) = W_L(\phi_L) = \left[\frac{a_-(s_0) + a_0(s_0) + \beta a_+(s_0)}{a_-(s_L) + a_0(s_L) + \beta a_+(s_L)}\right] \beta^{N(\phi_L) - P(\phi_L)}. \tag{5.4}$$

This telescoping is new (note that any sequence with $s_L = s_0$ has unit weight).

The probability of the Markov chain generating a particular length-L sequence $\phi_L = (s_0, s_1, \ldots, s_L)$ is given by

$$\Pr(\phi_L) = \left[\prod_{j=0}^{L-1} \left[\frac{1}{a_-(s_j) + a_0(s_j) + \beta a_+(s_j)}\right]\right] \beta^{P(\phi_L)}. \tag{5.5}$$

The expectation of the weight of a length-L sequence, times the indicator of the event that it ends in a specified state τ, is thus

$$\mathrm{E}\left[W_L(\phi) 1_{s_L(\phi)=\tau}(\phi)\right] = \sum_{\substack{\phi_L \text{ s.t.} \\ s_L(\phi_L)=\tau}} \left[\prod_{j=0}^{L-1} \left[\frac{1}{a_-(s_{j+1}) + a_0(s_{j+1}) + \beta a_+(s_{j+1})}\right]\right] \beta^{N(\phi_L)}, \tag{5.6}$$

where the summation over ϕ_L is over all length-L sequences starting from the trivial state s_0 and ending in state τ.

Now, reverse each step in the above ϕ_L to obtain a sequence

$$\phi'_L = (s'_0, \ldots, s'_L) = (s_L, \ldots, s_0), \tag{5.7}$$

which starts from $s_L = \tau$ and ends in the trivial state. Clearly, $N(\phi_L) = P(\phi'_L)$ and $P(\phi_L) = N(\phi'_L)$. Hence,

$$\mathrm{E}\left[W_L(\phi)1_{s_L(\phi)=\tau}\right] = \sum_{\substack{\phi'_L \text{ s.t.} \\ s'_L(\phi'_L)=s_0}} \left[\prod_{j=0}^{L-1}\left[\frac{1}{a_-(s'_j)+a_0(s'_j)+\beta a_+(s'_j)}\right]\right]\beta^{P(\phi'_L)}. \tag{5.8}$$

The summand is the probability that the sequence ϕ'_L starting in the state τ will terminate in the trivial state s_0, if the three sorts of atmospheric step are taken with the respective probabilities $P_+(s'_j)$, $P_0(s'_j)$, $P_-(s'_j)$, defined as in equations (5.1abc). Therefore,

$$\mathrm{E}\left[W_L(\phi)1_{s_L(\phi)=\tau}\right] = \Pr(s'_L = s_0 \mid s'_0 = \tau). \tag{5.9}$$

Let $|\tau| = n$, i.e., let n be the length of the self-avoiding walk τ. By the irreducibility and aperiodicity of the Markov chain, and the assumption of reversibility, the right side of equation (5.9) will be nonzero for all sufficiently large L.

If the parameter $\beta > 0$ of the Markov chain is sufficiently small (i.e., less than typical values of the quotient a_-/a_+, for all walk lengths n), the chain will be 'biased downward' in terms of the lengths of generated walks: informally, it will drift toward the zero-length (trivial) walk s_0. Moreover, it will be *recurrent* and *ergodic*, with a normalizable stationary density $\rho\colon S \to \mathbb{R}^+$ satisfying $\sum_{s\in S}\rho(s) = 1$. In this case, much more can be said. Irrespective of the choice of τ, equation (5.9) implies in this case that

$$\lim_{L\to\infty} \mathrm{E}\left[W_L(\phi_L)1_{s_L(\phi_L)=\tau}\right] = \rho(s_0), \tag{5.10}$$

which when summed over the c_n walks $\tau \in S_n$ becomes

$$\lim_{L\to\infty} \mathrm{E}\left[W_L(\phi_L)1_{|s_L(\phi_L)|=n}\right] = c_n\,\rho(s_0). \tag{5.11}$$

Therefore,

$$\frac{c_n}{c_m} = \lim_{L\to\infty} \frac{\mathrm{E}\left[W_L(\phi_L)1_{|s_L(\phi_L)|=n}\right]}{\mathrm{E}\left[W_L(\phi_L)1_{|s_L(\phi_L)|=m}\right]}, \tag{5.12}$$

which in particular, if $m = 0$, reduces to

$$c_n = \lim_{L\to\infty} \frac{\mathrm{E}\left[W_L(\phi_L)1_{|s_L(\phi_L)|=n}\right]}{\Pr(s_L(\phi_L) = s_0)}, \tag{5.13}$$

because as observed above, the weight W_L of any length-L sequence ϕ_L with $s_L = s_0$ equals unity. In fact,

$$\rho(s_0) = \lim_{L\to\infty} \Pr\left(s_L(\phi_L) = s_0\right). \tag{5.14}$$

An estimator for this limiting denominator is the fraction of time the Markov chain spends in the trivial state.

The GAS algorithm for estimating c_n/c_m (and hence c_n) is now clear: one runs the Markov chain on S for a large number L of iterations, and estimates the expected weights in equations (5.12) and (5.13) by *empirical means*, i.e.,

$$\mathrm{E}\left[W_L(\phi_L)1_{|s_L(\phi_L)|=n}\right] \approx L^{-1}\sum_{j=0}^{L-1} W_j(\phi_L)1_{|s_j(\phi_L)|=n}. \tag{5.15}$$

That is, one 'samples along the Markov chain.'

The resulting estimates will converge as they should, as $L \to \infty$, only if the Markov chain parameter $\beta > 0$ is sufficiently small, to ensure recurrence and ergodicity of the chain on its state space S. However, if for any choice of $n_{\max}$ the infinite space S is replaced by its truncation $S_*(n_{\max})$, which is finite, then in principle any positive value of β can be used.

5.2. Flat-histogram GAS (flatGAS). Using GAS, it may be difficult to obtain reliable estimates of c_n for large n, since the estimator has a large variance. This is because the Markov chain, if biased strongly 'downward' by having its drift parameter $\beta > 0$ chosen sufficiently small, may seldom visit states τ with $|\tau| = n$.

However, a flat histogram (over the lengths of the generated self-avoiding walks) version of GAS is possible, if the finite state space $S_*(n_{\max}) = \cup_{n=0}^{n_{\max}} S(n)$ is used. The aim is for the algorithm to sample states of each length $n \in [0, n_{\max}]$ roughly the same number of times. This is accomplished by a standard technique of importance sampling: altering the Markov chain transition probabilities (see equation (5.1)) by

n	c_n	GAS $\langle W_n\rangle$	flatGAS $\langle W_n\rangle$
0	1	1	1.00040
1	6	6.00013	6.00254
2	30	30.0131	29.9971
3	150	150.202	149.95
4	726	727.451	725.929
5	3534	3541.02	3535.28
6	16926	16958.1	16944.4
7	81390	81540.4	81497.5
8	387966	388757	388447
9	1853886	1.85830×10^6	1.85568×10^6
10	8809878	8.83188×10^6	8.81789×10^6
11	41934150	4.20464×10^7	4.19746×10^7
12	198842742	1.99482×10^8	1.99077×10^8
13	943974510	9.47294×10^8	9.45217×10^8
14	4468911678	4.48663×10^9	4.47232×10^9
15	21175146054	2.12653×10^{10}	2.11767×10^{10}
16	100121875974	1.00532×10^{11}	1.00112×10^{11}
17	473730252102	4.75742×10^{11}	4.73740×10^{11}
18	2237723684094	2.24647×10^{12}	2.23784×10^{12}
19	10576033219614	1.06104×10^{13}	1.05755×10^{13}
20	49917327838734	5.00652×10^{13}	4.98981×10^{13}
25	116618841700433358	1.16875×10^{17}	1.16462×10^{17}
30	27056990552545467 4614	2.71271×10^{20}	2.70128×10^{20}

TABLE 3. Approximate enumeration in 3-D (GAS and flatGAS).

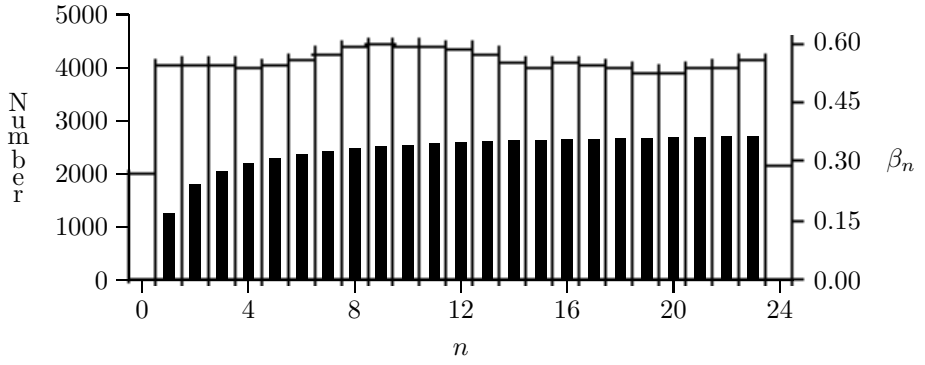


FIGURE 12. Flat histogram sampling in the flatGAS algorithm ($n_{\max} = 24$, $L = 100000$). Unfilled bars indicate the number of times walks of length n were visited (see left scale). Filled bars indicate β_n (see right scale). $\beta_0 = 1$ by definition and is omitted, since the probability of stepping from the walk of length 0 to one of length 1 is $1/2d$, independent of β_0. Also, $\beta_{24} = 0$.

making the parameter β dependent on the current state, and in particular on its length $n = n_j = |s_j|$. The aim is to reduce the expected drift to zero. The weights must also be altered, so as not to affect the expectation of the estimator while reducing its variance.

Since the expected downward drift in a state that is a self-avoiding walk of length n is roughly $\mathrm{E}_n\, a_+ / \mathrm{E}_n\, a_-$, the expectations being computed in stationarity over $S(n)$, it is reasonable to choose

$$\beta = \beta_{n_j} \approx \mathrm{E}_{n_j}\, a_+ / \mathrm{E}_{n_j}\, a_-, \tag{5.16}$$

where the quotient of expectations must somehow be estimated. With this choice, the Markov chain $\phi = (s_0, s_1, s_2, \dots)$ will induce a random walk on lengths, i.e., $(n_0, n_1, n_2, \dots)$, which should have a more or less uniform stationary distribution on $[0, n_{\max}]$. That is, the probability of a positive atmospheric move will be more or less equal to the probability of a negative one, for each $n \in (0, n_{\max})$.

In Figure 12 the flat histogram obtained from a run of this flatGAS Markov chain is shown. Here $n_{\max} = 24$, so the positive atmospheres of walks of length 24 were taken to be empty; and also $\beta_{24} = 0$. The other values β_n were taken to be empirical atmospheric ratios obtained in an earlier run (cf. equation (5.16)). The binning of states of each length $n \in [0, 24]$ in a sequence of length $L = 100000$ is indicated by the unfilled bars. The boundary states with $n = 0$ and $n = 24$ were visited about half as frequently as those with $n \in (0, 24)$, as one would expect.

It is easily checked that in flatGAS the appropriate definition of the weights is

$$W_L(\phi) = W_L(\phi_L) = \left[\frac{a_-(s_0) + a_0(s_0) + \beta_0 a_+(s_0)}{a_-(s_L) + a_0(s_L) + \beta_{|s_L|} a_+(s_L)}\right] \prod_{j=0}^{L-1} \beta_{|s_j|}^{\sigma(s_j, s_{j+1})}, \tag{5.17}$$

which is a modification of the definition used in GAS, equation (5.4). Estimation of c_n/c_m (and c_n) proceeds in flatGAS just as in GAS. The only really new feature is the technique used for estimating the drift parameters $\{\beta_n\}$. For this, it is useful to

generate N state sequences $\phi^{(m)}$, $m = 1, \ldots, N$, in succession, and update the $\{\beta_n\}$ on the basis of empirical atmospheric ratios, following each sequence. If the initial choices for the $\{\beta_n\}$ are erroneous, they will quickly improve, producing a flatter histogram on walk lengths $n \in [0, n_{\max}]$. The flatGAS algorithm differs from GAS in this self-tuning feature (in GAS, the sequences in a sample of N sequences would be independent).

In Table 3, GAS and flatGAS estimates of c_n for the cubic lattice $\mathbb{Z}^3$ are shown. Only endpoint atmospheres were used. The estimates were based on $N = 100$ sequences of length $L = 10^7$, with $n_{\max} = 71$, and the GAS algorithm used the constant parameter value $\beta = 0.212$. The flatGAS algorithm quickly settled down into flat histogram sampling. The observed errors show that the GAS and flatGAS data are scattered about the exact enumeration data (obtained from [**4**]). Overall, these relatively short runs produced estimates of c_n that deviate by less than 1% from exact data. More accurate results can be obtained by increasing either the length L of the sequences, or the number N of sequences.

GAS and flatGAS are in every respect, like Rosenbluth and flatPERM sampling, and GARM and flatGARM, *approximate* enumeration methods. It should be mentioned that the GAS algorithm with endpoint atmospheres is a generalization of the Berretti–Sokal dynamic Metropolis Monte Carlo algorithm [**1**], but incorporating sampling of weights along a Markov chain. This version of GAS is called GABS in reference [**24**], and its enhancement flatGABS is an especially efficient approximate enumeration tool for self-avoiding walks: it both performs atmospheric moves in $O(1)$ CPU time, and gives high quality estimates of c_n.

In GAS and flatGAS, the estimation of the microcanonical quantities $c_{n,e}$ proceeds similarly to the estimation of c_n, but now mean weights are tracked by the energy of the underlying sequence ϕ. Ratios of such mean weights are estimates of the ratios $c_{n,e_1}/c_{m,e_2}$, from which one may extract estimates of $c_{n,e}$.

6. Concluding remarks

GARM and flatGARM have been applied to the sampling of two-dimensional polygons in reference [**43**], and can easily be shown to be useful in sampling lattice trees and animals, as well as other lattice objects such as plaquette surfaces. Applying GARM to the sampling of additional objects is a matter of defining appropriate generalized atmospheres. In reference [**7**], GARM has been applied in a non-lattice context, to determine the size of classes of group elements.

Just as GARM and flatGARM are improvements on the original Rosenbluth method and flatPERM, so are GAS and flatGAS improvements on GARM and flatGARM. GAS introduces negative atmospheric moves into the mix of moves in GARM, while the flatGAS algorithm can self-tune to a flat histogram without the introduction of enrichment or pruning. Instead, it achieves flat-histogram sampling by tuning its single parameter β locally (for each n). More details on flatGAS can be found in reference [**24**].

Overall, the estimates of connective constants and critical exponents for self-avoiding walks using approximate enumeration data generated by GARM or GAS

do not approach the accuracy of estimates based on exact enumeration (see equations (1.9), (1.13), (1.14), and (1.16)). In reference [**24**], the flatGAS estimates

$$\mu = \begin{cases} 2.6383 \pm 0.0001, & \text{if } d = 2; \\ 4.684 \pm 0.001, & \text{if } d = 3, \end{cases} \tag{6.1}$$

$$\gamma = \begin{cases} 1.34 \pm 0.02, & \text{if } d = 2; \\ 1.15 \pm 0.02, & \text{if } d = 3, \end{cases} \tag{6.2}$$

were obtained using an endpoint atmosphere version of flatGAS (i.e., flatGABS). More efficient implementations of these algorithms, using techniques such as those of reference [**3**], should improve these estimates, and will be a natural next step in the examination of the properties of GARM and GAS algorithms.

Acknowledgements. The author is grateful to A. Rechnitzer for fruitful discussions.

References

[1] A. Berretti and A. D. Sokal, *New Monte Carlo method for the self-avoiding walk*, J. Statist. Phys. **40** (1985), 483–531.
[2] S. Caracciolo, M. S. Causo and A. Pelissetto, *High-precision determination of the critical exponent gamma for self-avoiding walks*, Phys. Rev. E **57** (1998), 1215–1218.
[3] N. Clisby, *Accurate estimate of the critical exponent ν for self-avoiding walks via a fast implementation of the pivot algorithm*, Phys. Rev. Lett. **104** (2010), article 055702, 4 pp.
[4] N. Clisby, R. Liang and G. Slade, *Self-avoiding walk enumeration via the lace expansion*, J. Phys. A **40** (2007), 10973–11017.
[5] B. Duplantier, *Polymer network of fixed topology: Renormalisation, exact critical exponent γ in two dimensions, and $d = 4 - \epsilon$*, Phys. Rev. Lett. **57** (1986), 941–944.
[6] P.-G. de Gennes, *Scaling Concepts in Polymer Physics*, Cornell Univ. Press, New York, 1979.
[7] M. Elder, E. Fusy and A. Rechnitzer, *Counting elements and geodesics in Thompson's group F*, J. Algebra, to appear. Available as arXiv:0902.0202.
[8] P. J. Flory, *The configuration of a real polymer chain*, J. Chem. Phys. **17** (1949), 303–310.
[9] P. J. Flory, *Statistical thermodynamics of semi-flexible chain molecules*, Proc. Roy. Soc. London Ser. A **234** (1955), 60–73.
[10] P. J. Flory, *Statistical Mechanics of Chain Molecules*, Wiley Interscience, New York, 1969.
[11] P. Grassberger, *Pruned–enriched Rosenbluth method: Simulation of θ-polymers of chain length up to* 1000000, Phys. Rev. E **56** (1997), 3682–3693.
[12] A. J. Guttmann and A. R. Conway, *Square lattice self-avoiding walks and polygons*, Ann. Comb. **5** (2001), 319–345.
[13] J. M. Hammersley, *Limiting properties of numbers of self-avoiding walks*, Phys. Rev. **118** (1960), 656–656.
[14] J. M. Hammersley, *The number of polygons on a lattice*, Proc. Cambridge Philos. Soc. **57** (1961), 516–523.
[15] J. M. Hammersley, *Generalization of the fundamental theorem on sub-additive functions*, Proc. Cambridge Philos. Soc. **58** (1962), 235–238.
[16] J. M. Hammersley and K. W. Morton, *Poor man's Monte Carlo*, J. Roy. Statist. Soc. Ser. B **16** (1954), 23–38.
[17] T. Hara and G. Slade, *The lace expansion for self-avoiding walks in five or more dimensions*, Rev. Math. Phys. **4** (1990), 235–327.
[18] T. Hara and G. Slade, *Critical behaviour of self-avoiding walks in five or more dimensions*, Bull. Amer. Math. Soc. (N.S.) **25** (1991), 417–423.
[19] T. Hara, G. Slade and A. D. Sokal, *New lower bounds on the self-avoiding-walk connective constant*, J. Statist. Phys. **72**, (1993), 479–517.
[20] H.-P. Hsu and P. Grassberger, *Polymers confined between two parallel plane walls*, J. Chem. Phys. **120** (2004), 2034–2041.

[21] H.-P. Hsu, W. Nadler and P. Grassberger, *Scaling of star polymers with* 1–80 *arms*, Macromolecules **37** (2004), 4658–4663.
[22] H.-P. Hsu, W. Nadler and P. Grassberger, *Simulations of lattice animals and trees*, J. Phys. A **38** (2005), 775–806.
[23] E. J. Janse van Rensburg and A. Rechnitzer, *Atmospheres of polygons and knotted polygons*, J. Phys. A **41** (2008), article 105002, 23 pp.
[24] E. J. Janse van Rensburg and A. Rechnitzer, *Generalized atmospheric sampling of self-avoiding walks*, J. Phys. A **42** (2009), article 335001, 29 pp.
[25] I. Jensen, *A parallel algorithm for the enumeration of self-avoiding polygons on the square lattice*, J. Phys. A **36** (2003), 5731–5745.
[26] I. Jensen, *Enumeration of self-avoiding walks on the square lattice*, J. Phys. A **37** (2004), 5503–5524.
[27] I. Jensen and A. J. Guttmann, *Self-avoiding walks, neighbour-avoiding walks and trials on semi-regular lattices*, J. Phys. A **31** (1998), 8137–8145.
[28] I. Jensen and A. J. Guttmann, *Self-avoiding polygons on the square lattice*, J. Phys. A **32** (1999), 4867–4876.
[29] H. Kesten, *On the number of self-avoiding walks*, J. Math. Phys. **4** (1963), 960–969.
[30] H. Kesten, *On the number of self-avoiding walks, II*, J. Math. Phys. **5** (1964), 1128–1137.
[31] J. Krawczyk, T. Prellberg, A. L. Owczarek and A. Rechnitzer, *Stretching of a chain polymer adsorbed at a surface*, J. Stat. Mech. Theory Exp. (2004), article P10004, 12 pp.
[32] G. F. Lawler, *Intersections of Random Walks*, Birkhäuser, Boston/Basel, 1991.
[33] J. C. Le Guillou and J. Zinn-Justin, *Accurate critical exponents from field theory*, J. de Physique **50** (1989), 1365–1370.
[34] B. Li, N. Madras and A. D. Sokal, *Critical exponents, hyperscaling, and universal amplitude ratios for two and three dimensional self-avoiding walks*, J. Statist. Phys. **80** (1995), 661–754.
[35] N. Madras and G. Slade, *The Self-Avoiding Walk*, Birkhäuser, Boston/Basel, 1993.
[36] N. Madras and A. D. Sokal, *Nonergodicity of local, length-preserving Monte Carlo algorithms for the self-avoiding walk*, J. Statist. Phys. **47** (1987), 573–595.
[37] N. Madras and A. D. Sokal, *The pivot algorithm: A highly efficient Monte Carlo method for the self-avoiding walk*, J. Statist. Phys. **50** (1987), 109–186.
[38] P. Nidras, *Grand canonical simulations of the interacting self-avoiding walk model*, J. Phys. A **29** (1996), 7929–7942.
[39] P. Nidras and R. Brak, *New Monte Carlo algorithms for interacting self-avoiding walks*, J. Phys. A **30** (1997), 1457–1469.
[40] G. L. O'Brien, *Monotonicity of the number of self-avoiding walks*, J. Statist. Phys. **59** (1990), 969–979.
[41] T. Prellberg and J. Krawczyk, *Flat histogram version of the pruned and enriched Rosenbluth method*, Phys. Rev. Lett. **92** (2004), article 120602, 4 pp.
[42] A. Rechnitzer and E. J. Janse van Rensburg, *Canonical Monte Carlo determination of the connective constant of self-avoiding walks*, J. Phys. A **35** (2002), L605–L612.
[43] A. Rechnitzer and E. J. Janse van Rensburg, *Generalized atmospheric Rosenbluth methods (GARM)*, J. Phys. A **41** (2008), article 442002, 8 pp.
[44] M. N. Rosenbluth and A. W. Rosenbluth, *Monte Carlo calculation of the average extension of molecular chains*, J. Chem. Phys. **23** (1955), 356–359.
[45] N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, published electronically at `http://www.research.att.com/~njas/sequences`, 2010.
[46] F. T. Wall and J. J. Erpenbeck, *New method for the statistical computation of polymer dimensions*, J. Chem. Phys. **30** (1959), 634–637.

Department of Mathematics and Statistics, York University, Toronto, Ontario, M3J 1P3, Canada
E-mail address: `rensburg@yorku.ca`
URL: `http://www.math.yorku.ca/Who/Faculty/Rensburg`

This page intentionally left blank

Contemporary Mathematics
Volume **520**, 2010

Fuchsian differential equations from modular arithmetic

Iwan Jensen

Abstract. Counting combinatorial objects and determining the associated generating functions can be computationally very difficult and expensive when using exact numbers. Doing similar calculations modulo a prime can be orders of magnitude faster. We use two simple polygon models to illustrate this: we study the generating functions of (singly) punctured staircase polygons and imperfect staircase polygons, counted by their extent along the main diagonal. For the former model this is equivalent to counting by the half-perimeter of the outer staircase polygon. We derive long series for these generating functions modulo a single prime, and then proceed to find Fuchsian ODEs satisfied by the generating functions, modulo this prime. Knowledge of a Fuchsian ODE modulo a prime will generally suffice to determine exactly its singular points and the associated characteristic exponents. We also present a procedure for the efficient reconstruction of the exact ODE, using results from multiple mod-prime calculations. Finally, we demonstrate how modular calculations can be used to factor Fuchsian differential operators.

1. Introduction

Counting the number of combinatorial objects p_n of size n (say, the number of self-avoiding polygons of perimeter $2n$, on the square lattice), and determining the corresponding generating function $F(x) = \sum p_n x^n$, is a fundamental pursuit of algebraic and enumerative combinatorics. A time-honored approach is to generate the series numerically (i.e., calculate p_n using a computer) and then try to guess the generating function using a symbolic package such as Gfun [**SZ**]. For typical combinatorial problems the number of objects grows exponentially with the size n, so computationally it can be extremely expensive to calculate p_n. Furthermore, various off-the-shelf search programs can quickly become stuck because of the often very large integer coefficients involved (2^n quickly becomes huge), and programs relying on exact formal calculations may easily run into memory barriers as well as time constraints.

In this article we review recent work, in which we have undertaken such calculations modulo a prime. This can be many orders of magnitude faster, and since the

2000 *Mathematics Subject Classification.* Primary 05A15; Secondary 82B41, 82B80.

Key words and phrases. Fuchsian differential equation, modular arithmetic, lattice polygon, singular point, critical exponent, characteristic exponent, rational reconstruction.

The author was supported in part by the Australian Research Council Grant DP0770705.

size of the coefficients is fixed and quite small, the problem of exhausting available memory is also somewhat alleviated. Specifically, our approach is to start from a series $F(x)$ known modulo a single prime, and then search for a linear differential equation with polynomial coefficients which has $F(x)$ as a solution (see Section 2). We illustrate the use of our methods by studying in Section 3 the generating functions $\mathcal{P}(x)$ and $\mathcal{I}(x)$ of punctured and imperfect staircase polygons on the square lattice. We calculate long series for these generating functions modulo a single prime, and then proceed to find ODEs modulo this prime. (That is, we calculate the coefficients of the polynomials in ODEs which $\mathcal{P}(x)$ and $\mathcal{I}(x)$ satisfy, modulo the prime.) We demonstrate in Section 3.2 how knowledge of an ODE modulo a single prime suffices to determine exactly its singular points and the associated characteristic exponents. That is, we need not know the ODE in exact arithmetic in order to find them exactly.

For $\mathcal{P}(x)$, the mod-prime ODE is sufficiently simple that we decided additionally to generate the exact series and from this calculate an exact ODE, which serves as a powerful check on the validity of the preceding modular calculations. (See Section 3.3.) For $\mathcal{I}(x)$, the mod-prime ODE is more complicated, so we developed a procedure outlined in Section 4 for the efficient reconstruction of an exact ODE using results from multiple mod-prime calculations. For this particular problem the multiple mod-prime approach is a factor of about 1000 times faster than first calculating the series exactly to the required order, and then finding an exact ODE using exact arithmetic.

In Section 5, we demonstrate how modular calculations can also be used to factor the differential operators in the calculated ODEs, in particular the order-11 one in the minimal ODE satisfied by $\mathcal{P}(x)$. Again, symbolic packages exist which do a decent job on ODEs of moderate size. The MAPLE package `DEtools` is very advanced and contains many useful routines, but for this order-11 ODE it simply fails, by running out of memory and time (the 'black-box' factorization routine `DFactor` quickly consumes many gigabytes of memory).

The use of modular calculations as outlined in this article promises to make it possible to find exact solutions, through numerical means, to many combinatorial problems which have hitherto been considered too computationally difficult. The prime example of this is probably our recent work on the 5-particle contribution $\tilde{\chi}^{(5)}$ to the square lattice Ising model [**BGHJ**]. The n-particle contributions $\tilde{\chi}^{(n)}$, which are combinatorial generating functions, can be expressed in terms of quite involved $(n-1)$-fold integrals over algebraic functions. The research outlined here stems from the studies of $\tilde{\chi}^{(3)}$ and $\tilde{\chi}^{(4)}$ by Zenine *et al.* [**ZBHM1, ZBHM2, ZBHM3**], where they found that these contributions are the solutions of quite high order (7 and 10, respectively) Fuchsian ODEs. Guttmann and Jensen [**GJ1, GJ2**] then proceeded to show that the generating functions for the combinatorial problems of punctured and three-choice polygons are solutions of 8th order Fuchsian ODEs. These studies needed moderately long series (a few hundred terms or so) in order to find the ODE, which was done using the exact series. This research has since progressed via the study of $\tilde{\chi}^{(5)}$ [**BGHJ**], which we found to be a solution of a high order ODE (the minimal order being 33), by using at least some 7400 terms of its series. The coefficients of the series and of the ODE were calculated modulo a single prime, and even attempting to generate the exact series (let alone finding an exact ODE) is completely beyond current computational resources.

The final piece of research reported on here, very briefly in Section 6, is our recent study [**BBGH**] of the factorization properties of the differential operator in the minimal-order ODE satisfied by $\tilde{\chi}^{(5)}$. We can factor such ODE operators, using modular calculations informed by knowledge of their characteristic exponents.

2. Fuchsian differential equations

The starting point of our approach is to calculate using an appropriate computer program a long series expansion of some function $F(x)$, which could be, for example, the generating function of a combinatorial problem. With the series coefficients known up to some order N, we look for a linear differential equation of order M such that $F(x)$ is an (approximate) solution, i.e.,

$$\sum_{k=0}^{M} P_k(x)\,\frac{\mathrm{d}^k}{\mathrm{d}x^k}F(x) = 0 + O(x^{N+1}), \tag{2.1}$$

where the $P_k(x)$ are polynomials. In what follows we shall always consider linear ODEs and generally leave out the word 'linear'. In order to make things as simple as possible, we limit the search to *Fuchsian* ODEs [**In**]. Such ODEs have only regular singular points. There are several reasons for searching for a Fuchsian ODE, rather than a more general differential equation. Computationally the Fuchsian assumption simplifies the search for a solution. From the general theory of Fuchsian equations [**In**] it follows that the degree of $P_k(x)$ is at most $D - M + k$, where D is the degree of $P_M(x)$. Thus any differential equation of Fuchsian type is constrained by two parameters, namely the order M and the degree D of the head polynomial $P_M(x)$. One may also argue, less precisely, that for most 'sensible' combinatorial models one would expect Fuchsian equations, as irregular singular points are characterized by explosive, super-exponential behavior. Such behavior is not normally characteristic of combinatorial problems. The point at infinity may be an exception to this somewhat imprecise observation. Recent work by Bostan *et al.* [**BBHM**] has thrown more light on why Fuchsian ODEs are so common in problems arising in statistical mechanics and enumerative combinatorics. One of their main observations is that many of the functions of interest can be written as n-fold integrals of algebraic integrands (they are, as pure mathematicians say, 'derived from geometry'). As shown in [**BBHM**], these functions are therefore necessarily solutions of ODEs which have rational coefficients, and moreover, are Fuchsian.

The Fuchsian condition on the ODE satisfied by $F(x)$ requires that all singular points be regular, and specifically that $x = 0$ and $x = \infty$ be regular. A form for the ODE that automatically satisfies this condition is $L_{MD}(F(x)) = 0$, where

$$L_{MD} = \sum_{i=0}^{M}\left(\sum_{j=0}^{D} a_{ij}\cdot x^j\right)\cdot\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^i, \qquad a_{M0}\neq 0, \quad a_{MD}\neq 0. \tag{2.2}$$

The condition $a_{M0} \neq 0$ (resp. $a_{MD} \neq 0$) ensures that $x = 0$ (resp. $x = \infty$) is regular. Note that it is the use of the operator $x\frac{\mathrm{d}}{\mathrm{d}x}$ (sometimes called Euler's operator), rather than just $\mathrm{d}/\mathrm{d}x$, which leads to the preceding simple conditions guaranteeing regularity of $x = 0$ and $x = \infty$, and to the equality of the degrees of the polynomials in front of the derivatives. A simple rearrangement of terms casts

$L_{MD}(F(x)) = 0$ into the form $\widehat{L}_{MD}(F(x)) = 0$, where

$$\widehat{L}_{MD} = \sum_{i=0}^{M} \left(\sum_{j=0}^{D} \hat{a}_{ij} \cdot x^{j+i} \right) \cdot \left(\frac{\mathrm{d}}{\mathrm{d}x} \right)^{i} \tag{2.3}$$

and the coefficients $\hat{a}_{ij}$ are linear combinations of the a_{ij}.

The only major difference between L_{MD} and $\widehat{L}_{MD}$ is the change from using Euler's operator $x\frac{\mathrm{d}}{\mathrm{d}x}$ to the standard operator $\frac{\mathrm{d}}{\mathrm{d}x}$. Either form can be used, but the form (2.2) has certain computational advantages. However, the form (2.3) was used in the original (and many subsequent) articles by Zenine *et al.* [**ZBHM1, ZBHM2, ZBHM3**]. It is also the form used by Guttmann and Jensen in their study of punctured and three-choice polygons [**GJ1, GJ2**]. The Fuchsian character of (2.3) is reflected in the decreasing degrees of the polynomials in front of successive derivatives. Both L_{MD} and $\hat{L}_{MD}$ contain $(M+1)(D+1)$ unknown coefficients.

It should be noted that there is no unique ODE for a given series. The series can be annihilated by many differential operators, i.e., be a solution of many ODEs. But among these ODEs there is one of minimal order m and degree D_m, and this ODE is unique. In terms of differential operators, the minimal-order differential operator appears as a *right factor* of any non-minimal-order differential operator. The minimal-order ODE may have a large number of 'apparent' singular points and can thus only be determined from a large number of series coefficients (generally speaking, $(m+1)(D_m+1)$ terms are needed). Other (non-minimal-order) ODEs, because they involve polynomials of smaller degrees, may require fewer series coefficients in order to be obtained. For any $M > m$, an ODE annihilating $F(x)$ (i.e., $L_{MD}(F(x)) = 0$), can be found for D sufficiently large, and if M is small enough we can choose M and D such that $(M+1)(D+1) < (m+1)(D_m+1)$. Among the non-minimal ODEs there will generally be one requiring the minimum number of terms. In a computational sense, one may view this as the 'optimal' ODE.

In previous articles such as [**ZBHM1, ZBHM2, ZBHM3, GJ1, GJ2**], the search for an ODE was done using exact series coefficients. The technique is to vary M and D until an $\widehat{L}_{MD}$ is found with $\widehat{L}_{MD}(F(x)) = 0 + O(x^{N+1})$, where from now on we shall drop the $O(x^{N+1})$. This condition leads to a set of linear equations for the unknown coefficients $\hat{a}_{ij}$. Naturally, one can always find such an ODE if the ODE is allowed to have at least N coefficients. The real problem is to find an ODE of order M and degree D with $F(x)$ (known only via the first N terms of its series) as a solution, such that the number of coefficients $K = (M+1)(D+1)$ of the ODE is less than N (and generally significantly less). The $N - K$ series terms not required in order to find the ODE are thus a strong check on (though obviously not a proof of) the correctness of the ODE. The only major computational trick used in these previous articles was that rather than solving the set of linear equations for the $\hat{a}_{ij}$ using *exact* arithmetic, it was done using floating-point arithmetic with very high precision (say, up to 1000 digits). Once a solution was found, the floating-point numbers were turned into rational numbers. Obviously the set of linear equations always yields a solution, that is, one can always find an ODE annihilating the first $K = (M+1)(D+1)$ terms of the series. To check if the ODE is actually correct (i.e., annihilates the full series) one can check numerically, at the floating-point level, whether $\widehat{L}_{MD}(F(x)) = 0$. However, a strong indication that the ODE is correct is for its coefficients to be fairly simple rational numbers. One accordingly factors $P_M(x)$ and checks whether the numbers of digits in the numerators of these

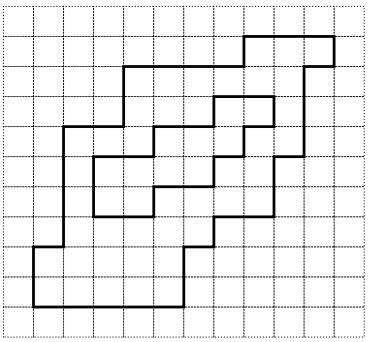
coefficients are 'small,' say, $\ll 1000$. One can perform a final check, using the ODE one has just obtained in exact arithmetic, to confirm that $\widehat{L}_{MD}(F(x)) = 0$.

Finding the ODE, if its size is large, can be very time consuming both in generating the series for $F(x)$ and in searching through values of M and D, looking for the ODE. In recent work [**BGHJ**], we have adopted a different and much more efficient strategy. Rather than performing the search using the exact series, we search only for a solution *modulo a specific prime* (in practice, we use the prime $p_r = 32749 = 2^{15} - 19$). The advantages of this are obvious. Firstly we only need generate a long series modulo a single prime (at least initially), and secondly, solving the system of linear equations determined by (2.2) largely amounts to finding whether or not the system has a zero determinant, which is easily done using Gaussian elimination. If a zero determinant is found, one can proceed to solve the system, which yields the ODE modulo the prime p_r. In [**BGHJ**] we were quite remarkably able to work with a series of 10000 terms for which we found several mod-prime ODEs requiring the determination of more than 7400 unknown coefficients. (As in the exact procedure, the remaining 2600 or so terms served as a powerful check on the correctness of the obtained ODE.) We emphasize that for this problem it would be quite impossible to calculate the *exact* series coefficients, since they become huge, growing as 2^n. Due to the size of the coefficients it would also be impossible to solve the set of linear equations for the coefficients of the ODE in either floating-point or exact arithmetic. In theory one must worry about possible false positive results, but we have never encountered this situation in practice (and in most cases, one can confirm any results by using a different prime). Below we give some further details of the procedure developed in [**BGHJ**].

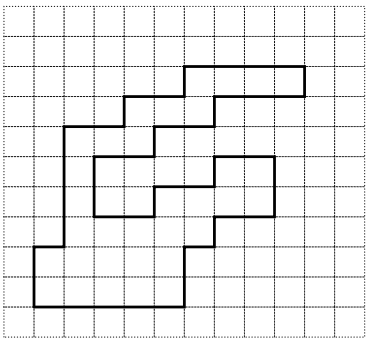
To determine the unknown coefficients a_{kj} of the polynomials in (2.2), we arrange the set of linear equations ($L_{MD}(F(x)) = 0$) in a specific order (see [**BGHJ**] for details). There exists a nontrivial solution if the determinant of the matrix of the system of $(M+1)(D+1)$ linear equations vanishes. We test this by standard Gaussian elimination, creating an upper triangular matrix U in the process. If we find that a diagonal element $U(K,K) = 0$ for some K, then a nontrivial solution exists. If $K < K_{MD} = (M+1)(D+1)$, we set to zero all a_{kj} in the ordered list beyond K. Of the remaining a_{kj} we set $a_{M0} = 1$, thus guaranteeing that $x = 0$ is a regular singular point, and determine the remaining coefficients by back-substitution. The K for which $U(K,K) = 0$ is the minimum number of series coefficients needed to find the ODE within the constraint of a given M and D. Obviously, $K \leqslant K_{MD} := (M+1)(D+1)$. Henceforth, D will always refer to the minimum degree D for which a solution was found for a given M. Then, for example, we can define a unique non-negative deviation Δ by $K = K_{MD} - \Delta = (M+1)(D+1) - \Delta$. In [**BGHJ**] we made a very striking empirical observation, namely that the numbers K depend on M and D via the simple linear relation

$$K = A \cdot M + B \cdot D + C = (M+1)(D+1) - \Delta, \tag{2.4}$$

where A, B, and C are constants depending on the particular series. We have no proof of this formula, but it has been found to work for all the problems we have studied. Furthermore, we have found that the constant B is reliably related to the minimum order possible for the ODE. In most cases the B found from (2.4) equals the order m of the minimal-order ODE, although there are exceptions to this rule; e.g., when a constant function is a solution of the ODE, so that $P_0(x) = 0$. If this is the case, then B equals $m - 1$, and one can obviously extend this to cases



Punctured staircase



Imperfect staircase

FIGURE 1. Examples of the types of polygons studied in this article.

where several of the lowest order polynomials are 0. The order m of the minimal-order ODE can thus be determined from non-minimal-order ODEs by using the remarkable formula (2.4). The constant A is the minimal possible degree of the ODE, i.e., the number of singular points, counted with multiplicity but excluding any 'apparent' singular points, and also excluding $x = 0$.

3. Punctured and imperfect staircase polygons

In recent articles, Guttmann and Jensen studied the problems of punctured staircase polygons [**GJ2**] and three-choice polygons [**GJ1**], and found in each case that the perimeter generating function can be expressed as the solution of a 8th order Fuchsian ODE.

A staircase polygon can be viewed as the intersection of two directed walks starting at the origin, moving only to the right or up, and terminating once the walks join at a vertex [**GP**]. The perimeter length of a staircase polygon is even. Let us denote by c_n the number of staircase polygons of perimeter $2n$. It is well known that $c_{n+1} = \frac{1}{n+1}\binom{2n}{n}$, i.e., that $c_{n+1} = C_n$, the n'th Catalan number. The associated half-perimeter generating function is

$$P(x) = \sum\nolimits_{n=2}^{\infty} c_n x^n = \frac{1 - 2x - \sqrt{1-4x}}{2} \sim \text{const} \times (1-\mu x)^{2-\alpha},$$

exhibiting both the connective constant $\mu = 4$, which determines the exponential growth of c_n as $n \to \infty$, and the critical exponent $\alpha = 3/2$. *Punctured* staircase polygons [**GJWE**] are staircase polygons with internal holes which are also staircase polygons (i.e., the polygons are mutually- as well as self-avoiding). In [**GJWE**] it was proved that the connective constant μ of k-punctured staircase polygons (i.e., ones with k holes) is the same as for unpunctured staircase polygons. Numerical evidence clearly indicates that the critical exponent α increases by $3/2$ per puncture. The closely related model of punctured discs was considered in [**JvRW**]. If punctured discs are counted by area, it was proved that the critical exponent increases by 1 per puncture. Here we study only the case of a *single* hole (see Figure 1), and we refer to these objects as punctured staircase polygons. The perimeter length of any staircase polygon is even, and thus the total perimeter of any punctured one

(i.e., its outer perimeter plus that of the hole) is also even. We denote by p_n the number of punctured staircase polygons of perimeter $2n$. The results of [**GJWE**] imply that the half-perimeter generating function $\sum_n p_n x^n$ has a simple pole at $x = x_c = 1/\mu = 1/4$, though the analysis in [**GJWE**] indicates that the critical behavior is more complicated than a simple algebraic singularity, and that logarithmic corrections to the dominant singular behavior are to be expected.

This was confirmed by a detailed analysis of the local solutions of the corresponding ODE, given in [**GJ2**]. Near the dominant singular point $x = x_c = 1/4$ the following singular behavior was found:

$$A(x)(1-4x)^{-1} + B(x)(1-4x)^{-1/2} + C(x)(1-4x)^{-1/2}\log(1-4x), \tag{3.1}$$

where $A(x)$, $B(x)$, and $C(x)$ are analytic in a neighbourhood of x_c. The ODE has other singular points. Near the singularity $x = x_- = -1/4$ on the negative x-axis, the behavior $D(x)(1+4x)^{13/2}$ was found, where $D(x)$ is analytic near x_-. Similar behavior is expected near the pair of singularities $x = \pm \mathrm{i}/2$, and finally at the roots of $1 + x + 7x^2$ the behavior $E(x)(1 + x + 7x^2)^2 \log(1 + x + 7x^2)$ is expected.

Three-choice self-avoiding walks on the square lattice were introduced by Manna [**Ma**], and can be defined as follows: Starting from the origin one can step in any direction; after a step upward or downward one can head in any direction (except backward); after a step to the left one can step forward or head downward, and similarly after a step to the right one can continue forward or turn upward. Alternatively put, one cannot make a right-hand turn after a horizontal step. As usual, one can define a polygon version of the walk model by requiring that the walk return to the origin. So a three-choice polygon [**GPO**] is simply a three-choice self-avoiding walk which returns to the origin, but has no other self-intersections. There are two distinct classes of three-choice polygons. The three-choice rule leads either to staircase polygons or to *imperfect* staircase polygons [**CGD**] (see Figure 1). Here we shall focus only on the case of imperfect staircase polygons, which as indicated in Figure 1 have exactly one notch or indentation. The ODE for the perimeter generating function of three-choice polygons has the same set of singularities as that for punctured staircase polygons, and the behavior at these singular points is almost the same, the only difference being that the first term in (3.1) is missing.

In the work reviewed below, the punctured and imperfect models are studied using an alternative counting variable, namely the 'length' (extent along the main diagonal) of the polygons. This 'length' is equal to the sum of the x- and y-coordinates of the point of the polygon furthest from the origin. For punctured staircase polygons this is equivalent to counting according to the half-perimeter of the outer staircase polygon (rather than the total perimeter of the outer and inner polygons combined). We denote the resulting generating functions for punctured and imperfect staircase polygons by $\mathcal{P}(x)$ and $\mathcal{I}(x)$, respectively.

3.1. Computer enumeration. The algorithms that we used to count the number of punctured and imperfect staircase polygons by 'length' are modified versions of the algorithm of Conway, Guttmann, and Delest [**CGD**] for the enumeration of imperfect staircase polygons. The two problems are similar, and consequently there are only minor differences between the algorithms. A detailed description of the algorithm we used to count imperfect staircase polygons can be found in [**GJ1**], and the minor changes required to count punctured ones are outlined in [**GJ2**]. The algorithms are based on transfer matrix techniques. This

TABLE 1. Number of terms of the series that are needed, to find an ODE of the form (2.2) or (2.3), modulo the prime $p_r = 2^{15} - 19$. M is the chosen order of the ODE, D is the degree of each of its polynomial coefficients, $K_{MD} = (M+1)(D+1)$, K is the number of terms predicted by (2.4) to be needed, and $\Delta = K_{MD} - K$.

$\mathcal{P}(x)$				
M	D	K_{MD}	K	Δ
11	53	648	646	2
12	31	416	413	3
13	23	336	334	2
14	20	315	310	5
15	17	288	286	2
16	16	289	284	5
17	15	288	282	6
18	14	285	280	5
19	13	280	278	2
20	13	294	287	7
21	13	308	296	12
22	12	299	294	5

$\mathcal{I}(x)$				
M	D	K_{MD}	K	Δ
14	92	1395	1300	95(2)
15	52	848	793	55(2)
16	39	680	637	43(3)
17	32	594	559	35(2)
18	28	551	520	31(2)
19	26	540	507	33(6)
20	24	525	494	31(6)
21	22	506	481	25(2)
22	21	506	481	25(3)
23	20	504	481	23(2)
24	20	525	494	31(10)
25	19	520	494	26(6)

entails bisecting the polygons by a line, and enumerating the number of polygons by moving the line 'forward' one step at a time. To count by length, all we need note is that the length is equal to the number of iterations or steps of the transfer matrix algorithm. This in fact makes these enumerations technically a little simpler than the original problem. The computed sequences of coefficients of $\mathcal{P}(x), \mathcal{I}(x)$ are available from the author. They appear as `A173408`, `A173409` in the OEIS [**Slo**].

3.2. ODEs for $\mathcal{P}(x)$ and $\mathcal{I}(x)$ modulo a prime. Starting with a series of 1000 terms for $\mathcal{P}(x)$ and one of 1500 terms for $\mathcal{I}(x)$, we were able to find underlying ODEs modulo the prime $p_r = 32749 = 2^{15} - 19$. In Table 1 we list the number of terms K of each series, which were required in order to find an ODE of order M and (minimal) degree D. One can easily check that the numbers follow the formula (2.4), the constants for $\mathcal{P}(x)$ being $A = 9$, $B = 11$, $C = -36$, and for $\mathcal{I}(x)$ being $A = 13$, $B = 13$, $C = -78$. However, for $\mathcal{I}(x)$ the resulting ODE always has $P_0(x) = 0$, which means that in this case the minimal order m equals not B but rather $B+1$, i.e., 14, and the ODE can be found using a reduced-size $M \times (D+1)$ matrix. In the rightmost column of Table 1, we indicate in parentheses the reduced value of the deviation Δ if this option is used. For both $\mathcal{P}(x)$ and $\mathcal{I}(x)$, we clearly see the substantial decrease in the number of terms K required to find the 'optimal' ODE ($K = 278$ and 481, respectively), as compared to the number of terms needed to find the minimal-order ODE ($K = 646$ and 1300). This illustrates the great utility of (2.4). From this formula we immediately get the minimal order, and we easily find the value of K for the minimal-order and optimal ODEs. This is very important in minimizing computational effort.

The possible singular points of the differential equation are given by the roots of the head polynomial $P_M(x)$, which factors into a polynomial $Q(x)$ exhibiting the true singular points as well as (in most cases) a polynomial exhibiting apparent

ones. As mentioned earlier, most of the calculated ODEs (including the minimal-order ODE) have many apparent singular points. The true ones can be found easily by calculating ODEs of different orders, factoring the polynomials $P_M(x)$ modulo p_r, and then deducing the singularity polynomials from the common factors. For punctured staircase polygons, we find (using the prime $p_r = 32749 = 2^{15} - 19$) that the common factors after factoring mod p_r are

$$(x + 8187)^4(x + 32748)^2(x + 26199)(x + 24562)(x + 10234),$$

where we recognize that since $-1/4 \bmod 32749 = 8187$, the first factor is just $(1 - 4x)^4$. Continuing in this way, one can easily deduce that the exact singularity polynomial is

$$Q_{\mathcal{P}}(x) = (1 - 16x)(1 - 5x)(1 - 4x)^4(1 - x)^2(1 + 4x). \tag{3.2}$$

For imperfect staircase polygons we similarly find that

$$Q_{\mathcal{I}}(x) = (1 - 16x)^2(1 - 5x)(1 - 4x)^6(1 - x)^3(1 + 4x). \tag{3.3}$$

The growth constant μ is given by the reciprocal of the position of the singular point closest to the origin, $x = 1/16$. So the coefficients in the power series for $\mathcal{P}(x)$ and $\mathcal{I}(x)$ grow asymptotically as 16^n, where n is 'length'; whereas if one counts polygons by total perimeter [**GJ1, GJ2**], the coefficients grow as 4^n.

It is possible from the method of Frobenius [**In**] to obtain from the indicial equation the characteristic exponents at the singular points. The indicial polynomial $P_{\mathrm{I}}(\lambda)$ at the singular point $x = x_s$ can be written as

$$P_{\mathrm{I}}(\lambda) = \sum_{k=0}^{M} \lim_{x \to x_s} \frac{(x - x_s)^{M-k} P_k(x)}{P_M(x)} \prod_{j=1}^{k} (\lambda - j + 1), \tag{3.4}$$

with the exponents at x_s obtained from the roots of $P_{\mathrm{I}}(\lambda)$. These calculations can be carried out using modular arithmetic, and we can thus find the exponents exactly (provided of course that they are integers or simple rational numbers). A simple MAPLE worksheet for calculating singular points and exponents for ODEs known modulo a prime is available from the author.

For example, for the minimal (order-11) ODE satisfied by $\mathcal{P}(x)$, we find at the singular point $x = 1/4$ that the indicial polynomial factored modulo p_r is

$$\begin{aligned}&\lambda(\lambda + 1)(\lambda + 16373)(\lambda + 16374)(\lambda + 16375)\\ &\quad \times (\lambda + 32743)(\lambda + 32744)(\lambda + 32745)(\lambda + 32746)(\lambda + 32747)(\lambda + 32748).\end{aligned}$$

Its roots are either integers or correspond to simple rational numbers; e.g., the factor $(\lambda + 16373)$ comes from $\lambda = 3/2 \bmod 32749$. In Tables 2 and 3 we list the exponents of the minimal-order ODEs for $\mathcal{P}(x)$ and $\mathcal{I}(x)$.

The results presented in this section cannot be guaranteed to be correct. After all, the calculations have been performed modulo only a single prime. However, due to the simplicity of the results for the singularities and exponents it is unlikely (though not impossible) that they are incorrect. To obtain a higher degree of confidence in such results, one can repeat the calculations for a different prime; and if one gets the same results, it is hard to imagine that they are not correct. In the following subsection we go one step further and find one of the two ODEs in exact arithmetic, thus confirming our results beyond any reasonable doubt.

TABLE 2. Exponents of the singular points of the minimal (order-11) Fuchsian ODE satisfied by $\mathcal{P}(x)$.

Singular point	Characteristic exponents
$x = 0$	1, 2, 3, 3, 10/3, 11/3, 4, 4, 5, 6, 8
$x = 1/16$	0, 1, 2, 3, 4, 5, 6, 13/2, 7, 8, 9
$x = 1/5$	0, 1/2, 1, 2, 3, 4, 5, 6, 7, 8, 9
$x = 1/4$	−1, −1/2, 0, 1/2, 1, 3/2, 2, 3, 4, 5, 6
$x = 1$	−2, −3/2, 0, 1, 2, 3, 4, 5, 6, 7, 8
$x = -1/4$	0, 1, 2, 3, 4, 5, 6, 13/2, 7, 8, 9
$x = \infty$	−3, −13/6, −2, −2, −11/6, −3/2, −5/4, −1, −3/4, 0, 1

TABLE 3. Exponents of the singular points of the minimal (order-14) Fuchsian ODE satisfied by $\mathcal{I}(x)$.

Singular point	Characteristic exponents
$x = 0$	0, 1, 2, 2, 7/3, 8/3, 3, 3, 3, 4, 5, 6, 7, 8
$x = 1/16$	0, 1, 2, 3, 4, 4, 5, 6, 13/2, 7, 8, 9, 10, 11
$x = 1/5$	0, 1/2, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12
$x = 1/4$	−1, −1/2, 0, 1/2, 1, 3/2, 2, 5/2, 3, 7/2, 4, 5, 6, 7
$x = 1$	−2, −3/2, −1, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
$x = -1/4$	0, 1, 2, 3, 4, 5, 6, 13/2, 7, 8, 9, 10, 11, 12
$x = \infty$	−2, −3/2, −7/6, −1, −1, −5/6, −1/4, 0, 0, 1/4, 1, 2, 3, 4

3.3. The ODE for $\mathcal{P}(x)$ in exact arithmetic. The minimal-order ODE for $\mathcal{P}(x)$ is small enough that the original method for finding the ODE in exact arithmetic should work. So we calculated the number of punctured staircase polygons up to length 750, exactly. The numbers become very large, so the calculation was performed using modular arithmetic [**Kn**]. This involves performing the calculation modulo several prime numbers p_i and then reconstructing the exact series coefficients at the end, by using the Chinese remainder theorem. We used primes of the form $p_i = 2^{30} - r_i$, where r_i is a small positive integer. Almost 100 primes of this form were needed to represent the integer coefficients correctly. The calculation for each prime used about 200MB of memory and 6 minutes of CPU time on a 2.8 GHz Xeon processor. The calculation of the ODE was done with MATHEMATICA using 1000 digit accuracy on a 2.5 GHz PowerPC G5 Macintosh computer. The calculation took about 40 minutes and used some 450MB of memory. Having found the exact ODE, we checked the results of the calculations summarized in Section 3.2; and we found all the singular points and associated exponents to be correct. We thus have complete confidence that the modular method for finding and analysing ODEs can yield exact results.

4. Reconstructing the exact ODE for $\mathcal{I}(x)$ from modular results

Calculating the exact minimal-order ODE for $\mathcal{I}(x)$ using the exact series coefficients would be a much more difficult task than it is for $\mathcal{P}(x)$. Since the size of the coefficients grows as 16^n, we would have to handle integers with some 1600 digits, using an array of size 1300^2 to solve the set of linear equations arising from

equation (2.2). We would thus expect to use more than four times the memory used to find the ODE for $\mathcal{P}(x)$, and an even larger multiple of the CPU time (and 1000 digits would, most likely, not provide sufficient accuracy to perform the floating-point calculations). So, such a calculation would probably stretch the capacity of that type of algorithm beyond its limit. Instead, we decided to use a different and (as we shall see) much more efficient approach. It is possible to reconstruct the exact ODE using the results from several mod-prime calculations. Depending on the precise approach taken, the ODE can be reconstructed exactly: using 18 primes in one case, but only 10 primes using a more efficient and better informed approach. Here we schematically outline the procedure for finding the exact minimal-order ODE. A MAPLE worksheet is available from the author.

Procedure for ODE reconstruction:

(1) Generate a long series modulo a single prime.
(2) Find ODEs of different orders and identify the constants A, B, and C of (2.4).
(3) Use this formula to identify both the minimal-order ODE and the optimal ODE requiring the least number of terms.
(4) Generate series for more primes p_i, long enough to find the optimal ODEs.
(5) Turn these ODEs into recurrences and generate longer series.
(6) Use these series to find the minimal-order ODEs mod p_i.
(7) Combine these modular results to obtain the exact minimal-order ODE, by using an algorithm for rational reconstruction [**Wa, Mo, CE**] to calculate the rational coefficients a_{ij} in (2.2).

In step (7) we used the built-in MAPLE routine `iratrecon`, which is based on work of Wang and Monagan [**Wa, Mo**]. This requires us first to use the Chinese remainder theorem to reconstruct the integers $b_{ij} := a_{ij} \bmod P$, where $P = \prod p_i$. The exact rational coefficients a_{ij} can then be calculated via the call $a_{ij} = \texttt{iratrecon}(b_{ij}, P)$. In this fashion one can find the rational number $a_{ij} = r/s$, provided that $2\,|r|\,s \leqslant P$.

We managed to reconstruct the exact ODE for $\mathcal{I}(x)$ using 18 primes of the form $p_j = 2^{30} - r_j$. Reconstructing the exact series coefficients using the Chinese remainder theorem up to the length needed to find the exact ODE by the original approach would have required about 10 times as many primes as the method used above. This is because in general, $P = \prod p_i$ must be larger than any of the integer coefficients we are trying to reconstruct; and since the coefficients in the series for $\mathcal{I}(x)$ grow as 16^n, the number of primes required to reconstruct a series of length N would be $4N/30$, or about 174 for $N = 1300$. We note that it only takes a few minutes to find the ODE modulo any given prime, and it also takes only a few minutes to reconstruct the exact ODE coefficient from the mod-prime ODEs.

Fewer than 18 primes are actually needed. In all the problems we have studied, the numerators happen to be much larger than the denominators. This means that we can modify the call to be $r/s = a_{ij} = \texttt{iratrecon}(b_{ij}, P, R, S)$, where R and S are positive integers such that $|r| \leqslant R$ and $0 < s \leqslant S$, with $2RS \leqslant P$. If we assume that $s < \sqrt{r}$ then we may choose $S = \sqrt[4]{P}$ and $R = P/(2S)$, and we can then find the a_{ij} using only 12 primes. This latter refinement clearly relies on the empirical observation that the denominators tend to be quite small, and this obviously need not be the case for all problems.

A further refinement is possible by generating the a_{ij} starting from a_{MD}. We then multiply all the residues by the denominator of a_{MD} modulo the respective primes. We go through the remaining coefficients by decreasing first j so as to generate all a_{Mj}. Whenever a non-integer rational number is encountered, we multiply all residues by its denominator (after this we found that the only remaining denominator was 9). We then repeat for $i = M - 1$, and so on until all a_{ij} have been exhausted. After this, the modified residues for the a_{ij} will be representations of integer coefficients, which we then reconstruct. This procedure can generate the exact integer coefficients of the ODE using only 10 primes.

We note that for $\mathcal{I}(x)$, this new procedure for finding the exact minimal-order ODE, based on multiple mod-prime calculations, is at least 1000 times faster than the original procedure described in Section 2 and used in Section 3.3 to find the exact ODE for $\mathcal{P}(x)$.

5. Factoring differential operators modulo a prime

Finally, we demonstrate how modular calculations can be used to factor linear differential operators. We use a method developed in [**BBGH**], which is similar to a method proposed by van Hoeij in [**vHo**]. The basic approach is to 'follow' the series associated to a specific characteristic exponent at a given singular point. Linear combinations of series with different exponents can be studied as well. The modular nature of our calculation is of great help in this, since with the series being known modulo a prime, the coefficient in the linear combination can take only a finite number of integer values; so 'guessing' the correct combination can be done by exhaustive search. For each series used as a candidate to 'break' the differential operator under consideration, we compute three (or more) ODEs, and from the ODE formula (2.4) we infer the minimal order.

For an ODE of minimal order m, let L_m denote the corresponding operator (henceforth the subscript on an operator will always denote its order). Consider a singular point $x = x_s$, and assume the exponents at this point are

$$\lambda_1^{k_1} < \lambda_2^{k_2} < \cdots < \lambda_p^{k_p}, \qquad \sum_{j=1}^{p} k_j = m, \tag{5.1}$$

where the superscript k_j on an exponent λ_j simply denotes its multiplicity. In our cases the exponents are either integers or rational numbers. Here we utilize only solutions which are analytic at the expansion point (say, $x = x_s = 0$). So in what follows we consider only integer exponents, and we denote them by n_j (with $n_1 < n_2 < \cdots < n_p$). Focusing on the highest exponent n_p, let us plug the series

$$S_p(x) = x^{n_p} + \sum_{k=n_p+1}^{\infty} a_k x^k \tag{5.2}$$

into the ODE. Requiring that $L_m(S_p(x)) = 0$ will determine all coefficients a_k. After calculating sufficiently many terms, we can find a new ODE satisfied by the particular series solution $S_p(x)$, which is by construction a solution of L_m. This new ODE will either have order m or have order $m_1 < m$. In the first case, either L_m is irreducible or the factor 'responsible' for annihilating the solution $S_p(x)$ is *not* the rightmost factor of L_m. In the second case, we must have a factorization

$$L_m = L_{m-m_1} \cdot L_{m_1}. \tag{5.3}$$

To summarize, the series associated to the highest exponent leads either to the original ODE (in the first case), or to a 'breaking' of it (in the second).

If the series $S_p(x)$ leads to the original ODE, we turn to the second-highest exponent n_{p-1}. In this case, a series $S_{p-1}(x)$ starting with $x^{n_{p-1}} + \cdots$, plugged into the original ODE, i.e., $L_m(S_{p-1}(x)) = 0$, will yield the expansion

$$S_{p-1}(x) = x^{n_{p-1}} + \sum_{k=n_{p-1}+1}^{n_p-1} a_k x^k + a_{n_p} x^{n_p} + \sum_{k=n_p+1}^{\infty} c_k x^k, \tag{5.4}$$

where all a_k with $k < p$ are fixed and the c_k's depend linearly on the free coefficient a_{n_p}, i.e., S_{p-1} is a one-parameter solution. This series $S_{p-1}(x)$ is a sum of a series starting with $x^{n_{p-1}} + \cdots$ and the series $a_{n_p} S_p(x)$, and is thus a linear combination of two formal series solutions, starting with $x^{n_{p-1}} + \cdots$ and $x^{n_p} + \cdots$, respectively. For *generic* values of the rational coefficient a_{n_p}, the series $S_{p-1}(x)$ will give rise to the original ODE. But for some values of the coefficient a_{n_p}, $S_{p-1}(x)$ may be the solution of an ODE of order less than m. If so, a factorization of L_m will be obtained.

To demonstrate how this procedure works in practice, we consider the order-11 operator of the minimal-order ODE for $\mathcal{P}(x)$ (denoted L_{11}). In this case we are fortunate to have access to both the exact ODE (through the calculation summarized in Section 3.3), and the exact series to any required length (in linear time). It is thus not a problem to generate the mod-prime series for any prime p_r we want. To find a 'breaking' linear combination we first form the series $S_{p-1}(x)$ for a given value of the free rational coefficient a_{n_p}, and then find the ODE annihilating $S_{p-1}(x)$. This procedure is carried out for all possible values of a_{n_p}; but since our calculations are done modulo p_r, this just means for all values in the interval $[1, p_r - 1]$. So when searching for a 'breaking' linear combination it is advisable to take the prime to be quite small, thus limiting the number of times the ODE solver must be used. (We use primes just above 1000 for this.) However, when reconstructing any factor exactly, larger primes are preferred, thus limiting the number of mod-prime ODEs required to calculate the factor in exact arithmetic. (In practice, as above we use primes just below 2^{30} for this.)

As displayed in Table 2, at the singular point $x = 0$ the exponents of the minimal ODE satisfied by $\mathcal{P}(x)$ are, with multiplicity,

$$1,\ 2,\ 3^2,\ 10/3,\ 11/3,\ 4^2,\ 5,\ 6,\ 8. \tag{5.5}$$

Substituting into the ODE the series that starts with $x^8 + \cdots$, i.e.,

$$S_8(x) = x^8 + \sum_{k=9}^{\infty} a_k x^k, \tag{5.6}$$

will determine all its coefficients. In fact, this procedure yields the expansion at $x = 0$ of $\mathcal{P}(x)$, thus leading to the original, minimal-order ODE. Of course this is not surprising, since the series for $\mathcal{P}(x)$, used to 'generate' the differential operator L_{11}, starts with x^8. The unique series $S_8(x)$ must be $\mathcal{P}(x)$ itself. So in this case, the series associated to the highest exponent cannot be used to 'break' L_{11}, since L_{11} is the minimal-order operator annihilating $\mathcal{P}(x)$.

Consider next a series that starts with $x^6+\cdots$, i.e., is associated to the second-highest exponent, namely

$$S(x) = x^6 + \sum_{k=7}^{\infty} a_k x^k. \tag{5.7}$$

We insert this series into the exact ODE for $\mathcal{P}(x)$ and then solve (term by term) the equations arising from $L_{11}(S(x)) = 0$. Doing this, we find that the coefficient a_7 is fixed, while the coefficient a_8 is undetermined and hence enters the series as a free parameter. The remaining coefficients are all expressed in terms of a_8. The resulting series is, after multiplication by 996 to ensure integer coefficients,

$$\begin{aligned} S(x) = 996x^6 + 6915x^7 + 996a_8x^8 + (13944a_8 - 333315)x^9 \\ + (126492a_8 - 3797208)x^{10} + (954168a_8 - 31074358)x^{11} + \ldots. \end{aligned}$$

The terms in $S(x)$ proportional to the free coefficient a_8 are the coefficients of the series $996S_8(x) = 996\mathcal{P}(x)$. We define $S_6(x)$ to be the series obtained from $S(x)$ by setting $a_8 = 0$. In order to break the operator L_{11}, we look at linear combinations $S_\alpha(x) = S_8(x) + \alpha S_6(x)$, where α is a rational number. For generic values of α the series $S_\alpha(x)$ is annihilated only by the full ODE of order 11. However, it is possible that for special values of α the series $S_\alpha(x)$ may be the solution of an ODE of order less than 11.

We do not know how such 'splitting' values of α can be obtained, except by exhaustive search. The use of modular calculations is very useful in searching for splitting values. The series $S_8(x)$ and $S_6(x)$ can be obtained modulo any prime p_r, and in the modular calculations α will thus take a value in the finite range $[1, p_r-1]$. If a rational splitting value of α exists, it can be found by looking for an underlying ODE of order less than 11 annihilating the series $S_\alpha(x)$. In the search we used the 'optimal' ODE, which (from Table 1) is of order 19 and degree 13, with $K = 278$. We used the prime $p_r = 1009 = 2^{10} - 15$ in our search. For each $\alpha \in [1, 1009]$ we calculated the series modulo p_r, and then looked for an annihilating ODE of order 19 and degree 13. For any value of α such an ODE exists, and for almost all values, $K = 278$. However, for the special values $\alpha = 18$, 132, and 962, we have $K = 274$, 256, and 138, respectively. The decrease in K is a sure sign that a simpler ODE annihilates $S_\alpha(x)$. In this particular case we find that the ODEs for $\alpha = 18$ and 132 are of order 10, while for $\alpha = 962$ the ODE is of order 3.

Working with the optimal ODE rather than the minimal-order ODE may appear counterintuitive, since we are attempting to find a factorization of L_{11} and not the order-19 operator O_{19} of the optimal ODE. However, in [**BBGH**] it is demonstrated that one can work with non-minimal-order ODEs because the minimal-order ODE is a right factor of the non-minimal-order ODE. In our case we have that there exists an 8th order operator O_8 such that $O_{19} = O_8 \cdot L_{11}$, and hence any right factor of O_{19} is also a right factor of L_{11}. We can confirm that the procedure works, by directly checking that the various factors obtained in exact arithmetic do indeed right-divide L_{11}.

We focus here only on the case $\alpha = 962$. The fact that the series is annihilated by a 3rd order operator means that we have found a right factor of L_{11} of order 3, and we thus have that $L_{11} = L_8 \cdot L_3$, where L_8 and L_3 are differential operators of orders 8 and 3, respectively. Going through the same procedure for a few more primes (5 in all), we find that in fact, $\alpha = 1/36646$. Having determined α exactly,

we then proceed to produce series for $S_\alpha(x)$ modulo several primes of the form $p_j = 2^{30} - r_j$, and calculate the minimal-order ODE and reconstruct L_3 exactly using the procedure described in Section 4, obtaining

$$L_3 = \sum_{i=0}^{3} P_i(x) \left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^i \tag{5.8}$$

with

$$\begin{aligned} P_3 = {} & x^3(1-4x)^2(1-x)^2(6640 - 115940x + 745572x^2 \\ & - 1660044x^3 - 3904387x^4 + 32245387x^5 - 81483743x^6 \\ & + 102350592x^7 - 63271990x^8 + 15728047x^9 - 1896930x^{10}), \end{aligned} \tag{5.9a}$$

$$\begin{aligned} P_2 = {} & x^2(1-4x)(1-x)(-39840 + 997500x - 10171096x^2 \\ & + 49748712x^3 - 65361298x^4 - 592676545x^5 + 3944925464x^6 \\ & - 12014127235x^7 + 21682463258x^8 - 23961110611x^9 \\ & + 15690708326x^{10} - 5774145597x^{11} + 1279604150x^{12} \\ & - 132785100x^{13}), \end{aligned} \tag{5.9b}$$

$$\begin{aligned} P_1 = {} & 2x(39840 - 1141920x + 13498966x^2 - 77265957x^3 \\ & + 109315350x^4 + 1533773209x^5 - 12778856043x^6 \\ & + 52423258998x^7 - 136624376823x^8 + 238960509306x^9 \\ & - 281352576354x^{10} + 218137005338x^{11} - 108012316206x^{12} \\ & + 34030365096x^{13} - 6798582940x^{14} + 607017600x^{15}), \end{aligned} \tag{5.9c}$$

$$\begin{aligned} P_0 = {} & 79680 + 2283840x - 26828612x^2 + 149243694x^3 \\ & - 144204350x^4 - 3690114536x^5 + 28993250188x^6 \\ & - 117936034106x^7 + 308125615678x^8 - 542161753964x^9 \\ & + 642034367908x^{10} - 499556504280x^{11} + 248708837380x^{12} \\ & - 80255885640x^{13} + 16105328880x^{14} - 1365789600x^{15}. \end{aligned} \tag{5.9d}$$

Next we use the command `rightdivision` from the MAPLE package `DEtools` to find exactly the factor $L_8 = \texttt{rightdivision}(L_{11}, L_3)$, and we then calculate the exponents of L_8 at the singularities of L_{11}. They are listed in Table 4. It should be noted firstly that the ODE corresponding to L_8 does not seem to be singular at $x = 1/5$ and $x = -1$. This can be confirmed by factoring the head polynomial of L_8, so L_3 is solely responsible for the singular behavior of L_{11} at these points. Secondly, it should be noted that all exponents of L_8 are rational.

The ODE corresponding to L_3 can be solved exactly, and it has the three independent solutions

$$F_1(x) = \frac{210x - 1718x^2 + 4271x^3 - 2836x^4 - 914x^5 + 1392x^5}{(1-4x)(1-x)^2}, \tag{5.10a}$$

$$F_2(x) = \frac{-2x + 10x^2 - 17x^3 + 18x^3}{\sqrt{1-4x}(1-x)^2}, \tag{5.10b}$$

$$F_3(x) = \frac{x(1-3x-x^2)^2}{(1-4x)(1-x)^{3/2}}\sqrt{1-5x}. \tag{5.10c}$$

TABLE 4. Exponents of the left factor L_8 of L_{11} (the order-11 operator in the minimal Fuchsian ODE satisfied by $\mathcal{P}(x)$), at the singular points of L_{11}.

Singular point	Characteristic exponents
$x = 0$	3, 3, 10/3, 11/3, 4, 4, 5, 8
$x = 1/16$	0, 1, 2, 3, 7/3, 4, 5, 6
$x = 1/5$	0, 1, 2, 3, 4, 5, 6, 7
$x = 1/4$	$-1/2$, 0, 1/2, 1, 2, 3, 4, 5
$x = 1$	0, 1, 2, 3, 4, 5, 6, 7
$x = -1/4$	0, 1, 2, 3, 7/2, 4, 5, 6
$x = \infty$	$-103/6$, -17, $-101/6$, $-65/4$, -16, $-63/4$, -15, -14

Having found these simple algebraic solutions it is natural for us to check if they appear as part of a direct sum decomposition of L_{11}. We do this by forming the series for the linear combinations $\mathcal{P}(x)+\alpha_i F_i(x)$, and then doing a search as above, to find the value (if any) of α_i, which makes the linear combination a solution of a lower order (actually order-10) ODE. We find $\alpha_1 = 1/253$ and $\alpha_2 = 1/6$, but we find no value α_3; so perhaps α_3 is not a rational number. We note that in addition to the solutions (5.10abc), L_{11} also has two simple polynomial solutions

$$Q_1(x) = x(1 - 16x^2) \qquad \text{and} \qquad Q_2(x) = x^2(1 - 4x). \tag{5.11}$$

Now, let L_9 denote the order-9 differential operator annihilating the 'reduced' series $\mathcal{P}_R(x) = \mathcal{P}(x) + F_1(x)/253 + F_2(x)/6$. The exponents at $x = 0$ of the corresponding ODE are

$$1,\ 2,\ 3^2,\ 10/3,\ 11/3,\ 4^2,\ 5. \tag{5.12}$$

By following the two solutions with exponents 5 and 4 we find a value of the combination coefficients that splits L_9 into a product of operators of order 7 and 2, $L_9 = L_7 \cdot L_2$. The operator L_2 can be found exactly using the method of Section 4, and hence solved. However, the two solutions are just $F_3(x)$ and $Q_1(x) - 3Q_2(x)$.

By acting with L_2 on $\mathcal{P}_R(x)$, we produce a series that is annihilated by the differential operator L_7. The exponents of L_7 at $x = 0$ are

$$2,\ 3^2,\ 10/3,\ 11/3,\ 4,\ 5. \tag{5.13}$$

Again we try to use the two highest exponents to split L_7, but this does not work in this case. When we plug the series starting with $x^4 + \cdots$ into the ODE, the coefficient of x^5 is a constant, so this series cannot be annihilated by the ODE. This is probably because the exponent 4 has multiplicity 2 in L_{11} (see Table 2), so there must be a solution of the form $x^4 \log(x)(1 + a_1 x + \cdots)$; and this logarithmic solution survives in L_7, which is why there is no solution of the form $x^4 + a_5 x^5 + \ldots$. Instead, we try using the series starting with $x^3 + \cdots$, and with this we find an order-1 factor such that $L_7 = L_6 \cdot L_1^{(a)}$, and going one step further we find a second order-1 factor so that $L_6 = L_5 \cdot L_1^{(b)}$. The ODE corresponding to $L_1^{(a)}$ has the exact solution

$$F(x) = \frac{x^3 P(x)}{(1 - x)^3} \tag{5.14}$$

with

$$P(x) = 1 - 23x + 281x^2 - 2595x^3 + 16426x^4 - 63202x^5 + 140322x^6 - 171410x^7 + 101936x^8 - 15904x^9.$$

Now $F(x)$ is not a solution of L_{11}, because the factor $L_1^{(a)}$ is not a right factor of L_9. However, we can use it to construct a solution $F_4(x) = x^3 + a_4x^4 + \cdots$ by solving $L_2(F_4(x)) = F(x)$, term by term. We find that in $L_2(F_4(x))$ the coefficient a_4 is a free parameter, and by fixing it at the value $a_4 = -9/2$ we obtain a new first order ODE, the solution of which,

$$F_4(x) = \frac{x^2(1 - 6x - 27x^2 + 140x^3 - 36x^4)}{(1 - 4x)(1 - x)^2}, \tag{5.15}$$

is linearly independent of the previous solutions. Similarly, we find that $L_1^{(b)}$ has a polynomial solution, which turns out to be just the previous solution $Q_2(x)$.

We find that the remaining operator L_5 cannot be factored by the method used above. To proceed further (if possible), we would have to use solutions associated to non-integer exponents at $x = 0$, or use expansions about other singularities such as $x = 1/4$ or $x = 1/16$. But this is work which we shall leave for the future.

6. Conclusion and outlook

In this article we have reported on recent progress in the problem of finding, numerically, the exact ODE underlying a series expansion of a given combinatorial generating function. We have shown how an ODE can be obtained from a series that is known modulo a single prime, and how the singular points and characteristic exponents can be obtained exactly (provided they are integers or simple rational numbers) from an ODE known modulo a single prime. The obvious advantage of this approach is that we need only calculate the series modulo one prime, thus saving a lot of CPU time; and opening up the possibility of studying large and complicated combinatorial problems such as that of the $\tilde{\chi}^{(5)}$ contribution to the susceptibility of the square lattice Ising model [**BBGH, BGHJ**].

We described in Section 4 a procedure for the efficient reconstruction of the exact ODE using results from multiple mod-prime calculations. For the particular problem of the generating function $\mathcal{I}(x)$ of imperfect staircase polygons, this approach is a factor of about 1000 times faster than first trying to generate the exact series and then finding the exact minimal-order ODE.

Additionally, we demonstrated in Section 5 how modular calculations can be used to factor linear differential operators. We showed that the differential operator L_{11} in the minimal ODE satisfied by the generating function $\mathcal{P}(x)$ can be factored as $L_{11} = L_8 \cdot L_3$ with L_3 being exactly solved, thus yielding three exact solutions. Two of these solutions can be subtracted from $\mathcal{P}(x)$ to leave us with a reduced series $\mathcal{P}_R(x) = \mathcal{P}(x) + F_1(x)/253 + F_2(x)/6$, which is the solution of an ODE based on a 9th order operator L_9, which in turn can be factored as $L_9 = L_5 \cdot L_1^{(b)} \cdot L_1^{(a)} \cdot L_2$. It is worth noting that L_9 is a right factor of L_{11}. The factorization carried out in Section 5 utilized only solutions analytic at $x = 0$, i.e., local solutions at $x = 0$ associated to the exponent zero. In future work we plan to extend the study of the ODEs for $\mathcal{P}(x)$ and $\mathcal{I}(x)$ by looking at local solutions associated to nonzero exponents, and local solutions around other singular points. It

may also be possible to utilize logarithmic solutions, such as those at $x = 0$, where two of the exponents of the ODE for $\mathcal{P}(x)$ have multiplicity two. (See Table 2.)

In this article we wanted to demonstrate how modular calculations can be used to speed up the search for the linear combinations of local solutions that split the differential operator into simpler factors. This may not be the most efficient approach. For example, for $\mathcal{P}(x)$ the only local solution at $x = 1/5$ that is singular, i.e., the only one that is associated to a non-integer exponent, starts with $(1-5x)^{1/2}$, and if we were to 'follow' this solution we would most likely find L_3 (or a simpler factor) directly, without the need for any guesswork. The approach of following singular local solutions has been used by B. Nickel to find many factors of the operator annihilating $\tilde{\chi}^{(5)}$ (several of these are reported in [**BBGH**]), and these factors were obtained without resorting to the type of guessing used in Section 5.

In Section 5, we had the luxury of being able to work with the exact minimal-order ODE satisfied by $\mathcal{P}(x)$, and thus with exactly known series. In the article [**BBGH**] the ODE for $\tilde{\chi}^{(5)}$ was only calculated modulo a few primes (and only one was really used for most calculations); and since the minimal-order ODE for $\tilde{\chi}^{(5)}$ requires almost 50000 terms while the 'optimal' ODE requires about 7400, it was impossible to work directly with the former. (Indeed, we would never have been able even to generate directly a series of 50000 terms.) Instead, this minimal-order ODE was always represented in the calculations by the optimal ODE or a nearly optimal one, and the factorization was carried out on these higher order ODEs, with the minimal order inferred from (2.4). These calculations of [**BBGH**] demonstrated the important fact that one can work at the level which is computationally most efficient.

The use of modular calculations as outlined in this article thus promises to make it possible to find exact solutions through numerical means to many combinatorial problems which have hitherto been considered too computationally difficult. The only real drawback (if one can call it that) is that *a priori* there is generally no way of knowing whether or not a series is a solution of an underlying (Fuchsian) ODE; and if so, how many terms one would need to uncover it.

Acknowledgements. The author would like to thank his co-workers S. Boukraa, A. J. Guttmann, S. Hassani, B. Nickel, J.-M. Maillard and N. Zenine for their many and varied contributions to this research. The calculations performed for this work used the computational resources of the Australian Partnership for Advanced Computing as well as the Victorian Partnership for Advanced Computing.

References

[BBGH] A. Bostan, S. Boukraa, A. J. Guttmann, S. Hassani, I. Jensen, J.-M. Maillard, and N. Zenine, *High order Fuchsian equations for the square lattice Ising model: $\tilde{\chi}^{(5)}$*, J. Phys. A **42** (2009), no. 27, article 275209, 32 pp.

[BBHM] A. Bostan, S. Boukraa, S. Hassani, J.-M. Maillard, J.-A. Weil, and N. Zenine, *Globally nilpotent differential operators and the square Ising model*, J. Phys. A **42** (2009), no. 12, article 125206, 50 pp.

[BGHJ] S. Boukraa, A. J. Guttmann, S. Hassani, I. Jensen, B. Nickel, J.-M. Maillard, and N. Zenine, *Experimental mathematics on the magnetic susceptibility of the square lattice Ising model*, J. Phys. A **41** (2008), no. 45, article 455202, 51 pp.

[CE] G. E. Collins and M. J. Encarnación, *Efficient rational number reconstruction*, J. Symbolic Comput. **20** (1995), no. 3, 287–297.

[CGD] A. R. Conway, A. J. Guttmann, and M. Delest, *The number of three-choice polygons*, Math. Comput. Modelling **26** (1997), no. 8–10, 51–58.

[GJ1] A. J. Guttmann and I. Jensen, *Fuchsian differential equation for the perimeter generating function of three-choice polygons*, Sém. Lothar. Combin. **54** (2005/07), article B54c, 14 pp.

[GJ2] A. J. Guttmann and I. Jensen, *The perimeter generating function of punctured staircase polygons*, J. Phys. A **39** (2006), no. 15, 3871–3882.

[GJWE] A. J. Guttmann, I. Jensen, Ling-Heng Wong, and I. G. Enting, *Punctured polygons and polyominoes on the square lattice*, J. Phys. A **33** (2000), no. 9, 1735–1764.

[GP] A. J. Guttmann and T. Prellberg, *Staircase polygons, elliptic integrals, Heun functions, and lattice Green functions*, Phys. Rev. E **47** (1993), no. 4, R2233–R2236.

[GPO] A. J. Guttmann, T. Prellberg, and A. L. Owczarek, *On the symmetry classes of planar self-avoiding walks*, J. Phys. A **26** (1993), no. 23, 6615–6623.

[In] E. L. Ince, *Ordinary Differential Equations*, Longmans, Green and Co., London, 1927.

[JvRW] E. J. Janse van Rensburg and S. G. Whittington, *Punctured discs on the square lattice*, J. Phys. A **23** (1990), no. 7, 1287–1294.

[Kn] D. E. Knuth, *Seminumerical Algorithms*, vol. 2 of *The Art of Computer Programming*, Addison–Wesley, Reading, MA, 3rd edition, 1997.

[Ma] S. S. Manna, *Critical behaviour of anisotropic spiral self-avoiding walks*, J. Phys. A **17** (1984), no. 16, L899–L903.

[Mo] M. B. Monagan, *Maximal quotient rational reconstruction: An almost optimal algorithm for rational reconstruction*, pp. 243–249 in *Proceedings of ISSAC 2004*, ACM Press, New York, 2004.

[SZ] B. Salvy and P. Zimmermann, GFUN: *A* MAPLE *package for the manipulation of generating and holonomic functions in one variable*, ACM Trans. Math. Software **20**, no. 2, (1994), 163–177.

[vHo] M. van Hoeij, *Factorization of differential operators with rational functions coefficients*, J. Symbolic Comput. **24** (1997), no. 5, 537–561.

[Slo] N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, published electronically at `http://www.research.att.com/~njas/sequences`, 2010.

[Wa] P. S. Wang, *Early detection of true factors in univariate polynominal factorization*, pp. 225–235 in *Computer Algebra: EUROCAL '83* (J. A. van Hulzen, ed.), no. 162 in Lecture Notes in Comput. Sci., Springer-Verlag, New York/Berlin, 1983.

[Wh] S. G. Whittington, *Anisotropic spiral self-avoiding walks*, J. Phys. A **18** (1985), no. 2, L67–L69.

[ZBHM1] N. Zenine, S. Boukraa, S. Hassani, and J.-M. Maillard, *The Fuchsian differential equation of the square Ising model $\chi^{(3)}$ susceptibility*, J. Phys. A **37** (2004), no. 41, 9651–9668.

[ZBHM2] N. Zenine, S. Boukraa, S. Hassani, and J.-M. Maillard, *Square lattice Ising model susceptibility: series expansion method, and differential equation for $\chi^{(3)}$*, J. Phys. A **38** (2005), no. 9, 1875–1899.

[ZBHM3] N. Zenine, S. Boukraa, S. Hassani, and J.-M. Maillard, *Ising model susceptibility: Fuchsian differential equation for $\chi^{(4)}$ and its factorization properties*, J. Phys. A **38** (2005), no. 19, 4149–4173.

DEPARTMENT OF MATHEMATICS AND STATISTICS, THE UNIVERSITY OF MELBOURNE, VICTORIA 3010, AUSTRALIA

E-mail address: `I.Jensen@ms.unimelb.edu.au`

URL: `http://www.ms.unimelb.edu.au/~iwan`

This page intentionally left blank

Contemporary Mathematics
Volume **520**, 2010

Random pattern-avoiding permutations

Neal Madras and Hailong Liu

Abstract. A pattern of length k is a permutation of $\{1, .., k\}$. This pattern is said to be contained in a permutation of $\{1, .., N\}$ (for $N \geqslant k$) if there is a subsequence of k elements of the (long) permutation that appears in the same relative order as the pattern. For a given pattern τ, let $\mathcal{S}_N(\tau)$ be the set of permutations of $\{1, .., N\}$ that do not contain τ. It is known that the cardinality of $\mathcal{S}_N(\tau)$ grows exponentially in N (rather than factorially), but little is known about the numerical value of the exponential growth rate. For example, which patterns of length 5 are easiest/hardest to avoid? We describe some Monte Carlo investigations into this and related problems. The design and implementation of these investigations raise some interesting mathematical questions. The Monte Carlo results lead to some new conjectures, including a description of what a "typical" 4231-avoiding permutation looks like. Numerical and rigorous results are presented.

1. Introduction

Pattern-avoiding permutations have been the subject of some extensive research by combinatorialists in the past couple of decades (see [**6**, Chapter 4] for an excellent introduction). In the present article, we look at some open problems in this field from the point of view of probability and experimental mathematics. Our primary goal is to show how to obtain some numerical data that is beyond the reach of exact enumeration methods. We raise some new questions and gain insights into some older ones.

For each positive integer N, let $\mathcal{S}_N$ denote the set of all permutations of the numbers $1, \ldots, N$. We represent a permutation as a string

$$\sigma \;=\; \sigma_1 \sigma_2 \ldots \sigma_N$$

where each σ_i is in $\{1, \ldots, N\}$, and $\sigma_i \neq \sigma_j$ whenever $i \neq j$ $(1 \leqslant i, j \leqslant N)$. It will often be helpful to visualize a permutation by considering the graph of σ_i as a function of i (see Figure 1).

2000 *Mathematics Subject Classification.* Primary 05A05; Secondary 05A16, 65C05.

Key words and phrases. Pattern-avoiding permutation, Stanley–Wilf limit, Monte Carlo.

N. Madras was supported in part by a Discovery Grant from NSERC.

H. Liu performed this research with support from an NSERC Undergraduate Summer Research Award while at York University.

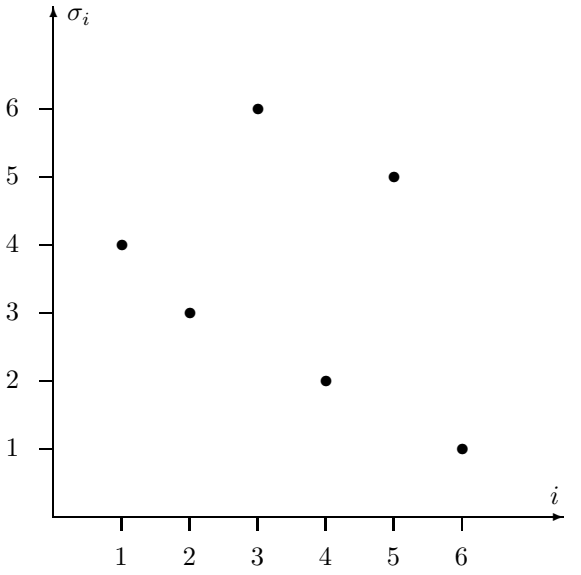


FIGURE 1. The graph of the permutation $\sigma = 436251$.

We now define the concept of "pattern," and what it means for a large permutation to contain or avoid a pattern.

DEFINITION 1.1. Let k be a positive integer.
(a) Let $\tau = \tau_1\tau_2\dots\tau_k$ be a permutation in $\mathcal{S}_k$. We say that a string of k distinct integers $a_1a_2\dots a_k$ *forms the pattern* τ if for each $i = 1,\dots,k$, a_i is the (τ_i)-th smallest element of $\{a_1,\dots,a_k\}$. For example, 7932 forms the pattern 3421 because 7 is the third smallest, 9 is the fourth smallest, etc.
(b) Let N be an integer with $N \geqslant k$. For $\sigma \in \mathcal{S}_N$ and $\tau \in \mathcal{S}_k$, we say that σ *contains the pattern* τ if some k-element subsequence $\sigma_{i(1)}\sigma_{i(2)}\cdots\sigma_{i(k)}$ of σ [where $1 \leqslant i(1) < i(2) < \cdots < i(k) \leqslant N$] forms the pattern τ. We say that σ *avoids the pattern* τ if σ does not contain τ. Let $\mathcal{S}_N(\tau)$ be the set of all permutations of $\{1,\dots,N\}$ that avoid τ.

For example, consider the permutation $\sigma = 436251$. Then σ contains the patterns 132 and 312, because of the subsequences 465 and 625. Also, $\sigma \in \mathcal{S}_6(123)$ because σ avoids the pattern 123, as is easy to see from the graph in Figure 1.

We are interested in $|\mathcal{S}_N(\tau)|$, the cardinality of $\mathcal{S}_N(\tau)$. The case $k = 2$ is trivial, since a permutation σ avoids 12 if and only if the elements of σ are arranged in decreasing order. Thus we see

$$|\mathcal{S}_N(\tau)| = 1 \quad \text{for every } \tau \in \mathcal{S}_2 \text{ and every } N \geqslant 2. \tag{1.1}$$

The case $k = 3$ is nontrivial, but it is well understood. It turns out [**19**] that

$$|\mathcal{S}_N(\tau)| = \frac{1}{N+1}\binom{2N}{N} \quad \text{for every } \tau \in \mathcal{S}_3 \text{ and every } N \geqslant 3. \tag{1.2}$$

The fact that the above right-hand side does not depend on τ can be partly explained by symmetry. Indeed, we know that $|\mathcal{S}_N(\tau)| = |\mathcal{S}_N(\omega)|$ if ω is τ's reversal

(i.e., $\omega_i = \tau_{k+1-i}$), or complement (i.e., $\omega_i = k+1-\tau_i$), or inverse, or any composition of these; thus, $|\mathcal{S}_N(123)| = |\mathcal{S}_N(321)|$ and $|\mathcal{S}_N(312)| = |\mathcal{S}_N(213)| = |\mathcal{S}_N(132)| = |\mathcal{S}_N(231)|$. The surprise in Equation (1.2) is that $|\mathcal{S}_N(123)| = |\mathcal{S}_N(312)|$. (It is less surprising, perhaps, that their common value is the N-th Catalan number!) Importantly, we see from Equation (1.2) that the pattern-avoiding permutations are very rare among the set of all permutations:

$$|\mathcal{S}_N(\tau)| \;<\; 2^{2N} \;\ll\; N! \;=\; |\mathcal{S}_N| \quad \text{for every } \tau \in \mathcal{S}_3 \;.$$

That is, the number of τ-avoiding permutations grows exponentially in N, as opposed to the factorial growth of the set of all permutations. In particular, we have

$$\lim_{N\to\infty} |\mathcal{S}_N(\tau)|^{1/N} \;=\; 4 \quad \text{for every } \tau \in \mathcal{S}_3 \;. \tag{1.3}$$

Does this property of exponential growth hold for all patterns of all lengths? The Stanley–Wilf Conjecture, formulated around 1980, predicted that the answer would be Yes (see [**6**] for some history). And indeed the answer is Yes, but a proof was discovered only recently [**16**]:

THEOREM 1.2 (Marcus and Tardos, 2004). *For every $k \geqslant 2$ and every $\tau \in \mathcal{S}_k$, the limit*

$$L(\tau) \;:=\; \lim_{N\to\infty} |\mathcal{S}_N(\tau)|^{1/N}$$

exists and is finite.

We call $L(\tau)$ the *Stanley–Wilf limit* of τ.

REMARK 1.3. As observed in [**3**], the hard part of proving Theorem 1.2 is proving finiteness, i.e., that $|\mathcal{S}_N(\tau)| < C^N$ for some finite C depending only on τ. Given this, the existence of the limit follows from the fact that, with some care, one can concatenate two τ-avoiding permutations to get a longer τ-avoiding permutation (see [**3**] or [**6**] for details), giving

$$|\mathcal{S}_N(\tau)|\,|\mathcal{S}_M(\tau)| \;\leqslant\; |\mathcal{S}_{N+M}(\tau)|.$$

This says that the sequence $\{|\mathcal{S}_N(\tau)|\}_N$ is *supermultiplicative* (or equivalently that the sequence $\{-\log|\mathcal{S}_N(\tau)|\}_N$ is *subadditive*). Fekete's Lemma (see [**9**], or [**15**, Lemma 1.2.2]) then implies that the limit exists (the limit being $+\infty$ if there is no exponential bound), and that

$$L(\tau) \;=\; \sup_{N\geqslant 1} |\mathcal{S}_N(\tau)|^{1/N}. \tag{1.4}$$

From Equations (1.1) and (1.3), we see that

$$L(\tau) = 1 \quad \text{for every } \tau \in \mathcal{S}_2, \text{ and}$$
$$L(\tau) = 4 \quad \text{for every } \tau \in \mathcal{S}_3 \;.$$

For longer patterns, only some values of $L(\tau)$ are known. For lengths 5 or more, our knowledge is rather sparse. One important result, due to Regev [**18**], says that for the increasing permutation $\tau = 12\ldots k$, it is the case that $L(\tau) = (k-1)^2$. More precisely,

$$|\mathcal{S}_N(12\ldots k)| \;\sim\; A_k N^{-k(k-2)/2}(k-1)^{2N} \quad \text{as } N\to\infty, \tag{1.5}$$

where A_k is an explicit constant (and "$f_N \sim g_N$" means that f_N/g_N converges to 1). The Stanley–Wilf limits for length 4 are mostly known, but one conspicuous gap remains. The present article was motivated by the challenge of filling this gap.

Regev's result (1.5) proves that $L(1234) = 9$. By reversal (i.e., reflecting the graph of a permutation σ in $\mathcal{S}_N$ about the vertical line $x = (N+1)/2$), it follows immediately that $L(4321)$ also equals 9. Less trivially, $L(2134) = 9$ [**4**]. However, other patterns of length 4 have different Stanley–Wilf limits: in particular, $L(1342) = 8$ [**6**]. From this it follows that $L(2431) = 8$ (by reversal), $L(1423) = 8$ (by inversion, i.e., reflecting the graph about the line $y = x$), and $L(4213) = 8$ (by complementation, i.e., reflecting the graph through the horizontal line $y = (N+1)/2$). In addition, $L(2413) = 8$ [**20**]. One can take care of most of the patterns of length 4 in this way, except for $\tau = 4231$ (and its reversal, etc.). Arratia [**3**] conjectured that $L(\tau) \leqslant (k-1)^2$ for every $\tau \in \mathcal{S}_k$. However, this conjecture was disproved by Albert *et al.* [**2**], who showed that $L(4231) \geqslant 9.47$. The best published upper bound for $L(4231)$ is 288 [**6**] (there was an error in an earlier proof [**5**] that claimed an upper bound of 36). This brings us to our first open problem:

Open Problem 1: Evaluate or estimate $L(4231)$.

Here, the word "estimate" can mean either "give a rigorous bound" or "find an approximation that is good in some sense". One way to approach this is to perform exact enumeration of $|\mathcal{S}_N(4231)|$ for several values of N. Equation (1.4) shows that $|\mathcal{S}_N(4231)|^{1/N}$ is a lower bound for $L(4231)$ for every N. But this alone does not work so well: exact values are known only up to $N = 25$ (see [**2**]), and $|\mathcal{S}_{25}(4231)|^{1/25} \approx 5.64$. The method of [**2**] for exact enumeration of $|\mathcal{S}_N(4231)|$ requires an exponential amount of effort: the complexity grows as $(1+\sqrt{2})^N$. Improvements may be possible, but it seems unlikely that anyone will know the exact value of $|\mathcal{S}_{50}(4231)|$ in the near future. For this reason, we shall turn our attention to approximations of $|\mathcal{S}_N(4231)|$ and related quantities for large values of N. We shall use Monte Carlo methods, which provide estimates together with error bars.

The following definition will be convenient for future discussion.

Definition 1.4. Let τ be a pattern in $\mathcal{S}_k$ and let σ be a permutation in $\mathcal{S}_N$ ($N \geqslant k$). A pattern map of τ into σ is an injection $\mathcal{M}\colon \{1,\dots,k\} \to \{1,\dots,N\}$ such that (a) $\mathcal{M}(i) < \mathcal{M}(i+1)$ for $i = 1,\dots,k-1$, and (b) $\sigma_{\mathcal{M}(1)} \dots \sigma_{\mathcal{M}(k)}$ forms the pattern τ.

For example, a pattern map of 132 into 436251 is given by $\mathcal{M}(1) = 1$, $\mathcal{M}(2) = 3$, and $\mathcal{M}(3) = 5$. Clearly a pattern map of τ into σ exists if and only if σ contains τ.

The article is organized as follows. Section 2 contains a brief introduction to Markov chain Monte Carlo (MCMC), a description of our approach, and some simple proofs of properties that ensure the validity of our methods. Section 3 discusses the efficiency of our algorithms. Our numerical results are presented in Section 4. Section 4.1 contains the results on $L(4231)$, together with a review of the statistical methodology used. Section 4.2 presents some results for patterns of length 5. Section 4.3 introduces a new area of investigation prompted by examining graphs of "typical" 4231-avoiding permutations.

2. Monte Carlo methods

Monte Carlo methods are mathematical experiments that use random numbers, typically with the goal of either (a) investigating typical behaviour of an inherently

random system, or (b) estimating some deterministic quantity by repeated simulation of a system with (possibly artificially added) randomness. (See, e.g., [**14**] for an introductory text.) The goals of this article will mostly be of type (b), i.e., estimating quantities related to $|\mathcal{S}_N(\tau)|$ by expressing these quantities in terms of the uniform probability distribution on $\mathcal{S}_N$ or $\mathcal{S}_N(\tau)$ or related sets. Our estimates are subject to random error, but careful statistical analysis can give an accurate description of the magnitude of this error. In Section 4.3 we shall pursue a goal of type (a), investigating properties of the uniform distribution on $\mathcal{S}_N(4231)$ for its own intrinsic interest.

To begin with, here is a simple Monte Carlo approach to our problem.

Direct Monte Carlo.
Fix N (the length of the permutation).
1. Generate a permutation σ uniformly at random from $\mathcal{S}_N$.
2. If $\sigma \in \mathcal{S}_N(\tau)$, then set $Y = 1$; otherwise, set $Y = 0$.

It is easy to perform Step 1 in time $O(N)$ (e.g., see [**7**, Section XIII.2]). It is less clear how easy it is to perform Step 2: what is the most efficient way to decide whether σ contains a given pattern? We shall return to this topic in Section 3.

Clearly, the probability that $Y = 1$ is exactly $|\mathcal{S}_N(\tau)|/|\mathcal{S}_N|$. Suppose we perform many (say, M) independent repetitions of this Direct Monte Carlo experiment. This produces a sequence of independent random variables $Y_1, \ldots, Y_M$, each having the same distribution as Y. Let $\bar{Y}_M$ be the sample average, i.e.,

$$\bar{Y}_M := \frac{Y_1 + \cdots + Y_M}{M}.$$

Then the expected value of $N!\,\bar{Y}_M$ is exactly $|\mathcal{S}_N(\tau)|$. However, this will not be very useful for large N, for the following reason. Let $p = |\mathcal{S}_N(\tau)|/N!$ be the probability that a random σ from Step 1 avoids τ. Unless M is at least of the order of $1/p$, it is very likely that none of the M random σ's will avoid τ, and hence $N!\,\bar{Y}_M$ will equal zero. Since $1/p$ grows faster than exponentially in N, we see that desirable values of M become too large to be practical as N grows. Thus the Direct Monte Carlo Method will not help us estimate $L(\tau)$ (although we do have a use for this method in Section 4.2).

Evidently, we need a less direct approach. Here is the one we shall use. First we explain the procedure, and then we explain why it helps us.

Indirect Monte Carlo.
Fix N (the length of the permutation).
1. Generate a permutation σ uniformly at random from $\mathcal{S}_N(\tau)$.
2. For each $i = 0, \ldots, N$, let $\sigma^{\dagger i}$ be the permutation in $\mathcal{S}_{N+1}$ obtained from σ by inserting "$N+1$" between the i^{th} and $(i+1)^{th}$ positions, i.e., $\sigma^{\dagger i} = \sigma_1 \ldots \sigma_i (N+1) \sigma_{i+1} \ldots \sigma_N$.
3. Let $W(\sigma)$ be the number of values of i for which $\sigma^{\dagger i} \in \mathcal{S}_{N+1}(\tau)$.

Example. If $\tau = 4231$, $N = 6$, and $\sigma = 315624$, then

$$\sigma^{\dagger 0} = 7315624, \qquad \sigma^{\dagger 1} = 3715624, \quad \sigma^{\dagger 2} = 3175624, \qquad \sigma^{\dagger 3} = 3157624,$$
$$\sigma^{\dagger 4} = 3156724, \qquad \sigma^{\dagger 5} = 3156274, \qquad \sigma^{\dagger 6} = 3156247.$$

It is easy to check that $\sigma^{\dagger 3}$, $\sigma^{\dagger 4}$, $\sigma^{\dagger 5}$, and $\sigma^{\dagger 6}$ avoid 4231, while the other three contain it; therefore $W(\sigma) = 4$.

It is not apparent how to perform Step 1 efficiently, but let us not worry about that yet. The following result helps explain the significance of W.

LEMMA 2.1. *Consider the random variable $W(\sigma)$ obtained as output of the Indirect Monte Carlo algorithm, where σ has been chosen uniformly at random from $\mathcal{S}_N(\tau)$. Its expected value satisfies*

$$E(W(\sigma)) \;=\; \frac{|\mathcal{S}_{N+1}(\tau)|}{|\mathcal{S}_N(\tau)|}. \tag{2.1}$$

PROOF. Observe that for each permutation ρ in $\mathcal{S}_{N+1}(\tau)$, there is a unique σ in $\mathcal{S}_N(\tau)$ and a unique i in $\{0, 1, \ldots, N\}$ such that $\sigma^{\dagger i} = \rho$. It follows that

$$\sum_{\sigma \in \mathcal{S}_N} W(\sigma) \;=\; |\mathcal{S}_{N+1}(\tau)|\,.$$

Dividing both sides by $|\mathcal{S}_N(\tau)|$ gives Equation (2.1) (since each σ has probability $1/|\mathcal{S}_N(\tau)|$ of being chosen in Step 1). □

Thus, if we repeat the Indirect Monte Carlo procedure many times, then the average of all the observed $W(\sigma)$'s is an estimate of $|\mathcal{S}_{N+1}(\tau)|/|\mathcal{S}_N(\tau)|$.

The above lemma makes the following result particularly relevant.

THEOREM 2.2. *For every pattern τ of length 2, 3, 4, or 5,*

$$\lim_{N\to\infty} \frac{|\mathcal{S}_{N+1}(\tau)|}{|\mathcal{S}_N(\tau)|} \;=\; L(\tau)\,. \tag{2.2}$$

For lengths 2 and 3, the proof is immediate from Equations (1.1) and (1.2). For lengths 4 and 5, the proof is due to Atapour and Madras (in preparation), who prove Equation (2.2) for infinitely many other patterns as well. However, at this time the following remains unsolved:

> Open Problem 2: Prove that Equation (2.2) holds for every pattern τ.

We note that Regev's result (1.5) proves that (2.2) holds for patterns $12\ldots k$. We remark that to prove Open Problem 2 in general, it suffices to show that the limit in Equation (2.2) exists.

In this article we shall study only patterns of length 4 and 5, so Theorem 2.2 will be sufficient for our purposes. Our plan to estimate $L(\tau)$ experimentally is the following.

> *Use the Indirect Monte Carlo algorithm to estimate the ratio $|\mathcal{S}_{N+1}(\tau)|/|\mathcal{S}_N(\tau)|$ for several (large) values of N, and then try to extrapolate the data to the limit as $N \to \infty$.*

(Similar methods have been used to estimate exponential growth rates in statistical mechanics; e.g., see [**10**, Eq. (17)].) We shall discuss the extrapolation methods later. But first, we have one gap to fill: how do we perform Step 1 in the Indirect Monte Carlo algorithm? We shall actually do something slightly different: we shall generate a random permutation σ from a distribution that is approximately uniform on $\mathcal{S}_N(\tau)$ (and this approximation can be made as good as desired, in principle). The method for doing this is *Markov chain Monte Carlo* (MCMC), which we now describe (see, e.g., [**14**] for more details and references).

Here is the general strategy of Markov chain Monte Carlo in our setting. Fix N and τ. We want to define a discrete-time Markov chain $\{X(t) : t = 0, 1, \dots\}$ whose *state space* is $\mathcal{S}_N(\tau)$ (i.e., $X(t) \in \mathcal{S}_N(\tau)$ for every $t = 0, 1, \dots$). We shall write the transition probabilities of this chain as $p(\cdot, \cdot)$:

$$p(\sigma, \tilde{\sigma}) := \Pr(\, X(1) = \tilde{\sigma} \,|\, X(0) = \sigma \,) \qquad [\sigma, \tilde{\sigma} \in \mathcal{S}_N(\tau)].$$

Moreover, we want the Markov chain to have three properties:

- symmetric [i.e., $p(\sigma, \tilde{\sigma}) = p(\tilde{\sigma}, \sigma)$];
- irreducible [i.e., for all $\sigma, \tilde{\sigma} \in \mathcal{S}_N(\tau)$, there exists $k > 0$ such that $\Pr(\, X(k) = \tilde{\sigma} \,|\, X(0) = \sigma \,) > 0$]; and
- aperiodic (for this, it suffices that $p(\sigma, \sigma) > 0$ for at least one σ; the full definition, which we shall not need, may be found in standard references such as [**8**] or [**11**]).

If these three properties hold, then the uniform distribution on $\mathcal{S}_N(\tau)$ is the unique equilibrium distribution of our Markov chain [**8, 11**]. In particular, the expected value of $W(X(t))$ [given any $X(0)$] converges to $E(W(\sigma))$ as given in Lemma 2.1 (here, E is with respect to uniform distribution on $\mathcal{S}_N(\tau)$). Moreover, by the Law of Large Numbers for Markov chains (e.g., [**17**]),

$$\lim_{M\to\infty} \frac{1}{M} \sum_{t=1}^{M} W(X(t)) \;=\; E(W(\sigma)) \;=\; \frac{|\mathcal{S}_{N+1}(\tau)|}{|\mathcal{S}_N(\tau)|} \quad \text{with probability 1.} \tag{2.3}$$

That is: if we run our chain for a sufficiently large number (M) of iterations, then the sample mean of $W(X(t))$ will be a good estimate of $|\mathcal{S}_{N+1}(\tau)|/|\mathcal{S}_N(\tau)|$. It is hard to say *a priori* how large M should be, so we shall deal with this question empirically; see Section 4.

Next we define our Markov chain. Here is an informal description. Given $X(0)$ in $\mathcal{S}_N(\tau)$, pick one of the N components of $X(0)$ and try to move it to another position. If the result still avoids τ, then we accept the result, and assign it to $X(1)$; but if the result contains τ, then we reject it, and $X(1)$ is defined to be $X(0)$. Then we repeat for $X(2)$, and so on. To make this description more formal, we define a function F as follows.

Algorithm to define the function $F\colon \mathcal{S}_N \times \{1, \dots, N\}^2 \to \mathcal{S}_N$.

Let $\sigma \in \mathcal{S}_N$ and $I, J \in \{1, \dots, N\}$.

1. Let ρ be the string obtained from σ by deleting the entry σ_I:
$$\rho \;=\; \rho_1 \dots \rho_{N-1} \;=\; \sigma_1 \dots \sigma_{I-1} \sigma_{I+1} \dots \sigma_N.$$
2. Let $\tilde{\rho}$ be the permutation obtained from ρ by inserting the number σ_I between the $(J-1)^{th}$ and J^{th} position:
$$\tilde{\rho} \;=\; \rho_1 \dots \rho_{J-1} \sigma_I \rho_J \dots \rho_{N-1}.$$
3. Then $F(\sigma, I, J) := \tilde{\rho}$.

Informally, F moves the I^{th} entry of σ to the J^{th} position. For example, suppose $\sigma = 1234567$. Then $F(\sigma, 4, 2) = 1423567$, $F(\sigma, 4, 7) = 1235674$, $F(\sigma, 4, 1) = 4123567$, and $F(\sigma, 4, 4) = \sigma$. Also observe that $F(\sigma, 4, 5) = 1235467 = F(\sigma, 5, 4)$.

Here is a formal algorithmic definition of our Markov chain.

Markov Chain Iteration.

If $X(t) \in \mathcal{S}_N(\tau)$, then the following algorithm gives $X(t+1)$.

1. Choose $I \in \{1, \dots, N\}$ uniformly at random.

2. Choose $J \in \{1, \dots, N\}$ uniformly at random.
3. Let $\theta = F(X(t), I, J)$.
4. If θ avoids τ, then set $X(t+1)$ to be θ. If θ contains τ, then set $X(t+1)$ to be $X(t)$.

PROPOSITION 2.3. *Fix a pattern τ of length $k \geqslant 3$, and fix $N > k$. The Markov chain defined by the above randomized algorithm is irreducible, symmetric, and aperiodic on $\mathcal{S}_N(\tau)$. Hence the unique equilibrium distribution of the chain is the uniform distribution on $\mathcal{S}_N(\tau)$.*

PROOF. First, the chain is clearly aperiodic because $F(\sigma, I, I)$ always equals σ. Next we show that the chain is symmetric. For $\sigma, \tilde{\sigma} \in \mathcal{S}_N(\tau)$, let $\mathcal{F}^*(\sigma, \tilde{\sigma})$ be the set of pairs $(I, J) \in \{1, \dots, N\}^2$ for which $F(\sigma, I, J) = \tilde{\sigma}$. Then $p(\sigma, \tilde{\sigma}) = |\mathcal{F}^*(\sigma, \tilde{\sigma})|/N^2$. Observe that for any pair (I, J) in $\mathcal{F}^*(\sigma, \tilde{\sigma})$, we have $F(\tilde{\sigma}, J, I) = \sigma$. Hence $|\mathcal{F}^*(\sigma, \tilde{\sigma})| = |\mathcal{F}^*(\tilde{\sigma}, \sigma)|$, and it follows that $p(\tilde{\sigma}, \sigma) = p(\sigma, \tilde{\sigma})$. Hence the chain is symmetric. (Note that there is a bit more subtlety here than may be apparent at first sight. For example, choosing a pair uniformly at random from $\{(I, J)\colon F(\sigma, I, J) \neq \sigma\}$ would not give a symmetric chain.)

Finally we prove irreducibility. Assume without loss of generality that the pattern τ satisfies $\tau_1 > \tau_k$. Let ψ be the increasing permutation $12 \dots N$ (i.e., $\psi_i = i$ for every i). Then ψ avoids τ. Irreducibility is a direct consequence of symmetry and the property that for every $\sigma \in \mathcal{S}_N(\tau)$, $\Pr(X(N-1) = \psi \,|\, X(0) = \sigma) > 0$. To prove this property, start with an arbitrary σ, and move the largest entry to the rightmost position; then move the second-largest entry to the second position from the right end; and so on. After $N - 1$ such moves we reach ψ, and it is not hard to see that none of these moves can create the pattern τ. Therefore the property holds, and the proof is complete. □

REMARK. As an alternative algorithm, we could replace Step 2 of the Markov Chain Iteration by "Choose $J \in \{1, \dots, N\} \setminus \{I\}$ uniformly at random." Since $F(\sigma, I, I)$ is always σ, this change would reduce some of the time spent on iterations that do not change σ. Proposition 2.3 holds for this variation. (The proof is identical except for aperiodicity, which can be shown as follows. Clearly there exists a permutation $\sigma' \in \mathcal{S}_N$ and an I and J such that σ' avoids τ but $F(\sigma', I, J)$ contains τ. Therefore $p(\sigma', \sigma') \geqslant [N(N-1)]^{-1}$.)

3. Algorithmic issues

One issue that we have hitherto evaded in our discussion of Monte Carlo is the following: *Given a pattern $\tau \in \mathcal{S}_k$ and a permutation $\sigma \in \mathcal{S}_N$, how does one check whether σ contains τ?* It turns out that this is a crucial issue in the implementation of the Monte Carlo method, since this checking is the slowest part of the process. Everything else in an iteration (whether it be one time step of the Markov chain, or one new permutation in the Direct Monte Carlo approach) can be accomplished in time of order N. But it is not clear how to check for a general pattern τ most efficiently.

At worst, searching σ for a given pattern of length k can be done in time $O(N^k)$, by checking all k-element subsequences of σ. For the Markov chain of Section 2, we can do a bit better. In Step 4, we can take advantage of the facts that (*i*) θ is obtained from $X(t)$ by moving a single entry, and (*ii*) $X(t)$ avoids τ. Therefore,

if θ contains τ, then the moved element must be a part of the pattern; or more formally, if there exists a pattern map $\mathcal{M}$ of τ into θ, then $\mathcal{M}(i) = J$ for some i. This observation validates the following algorithm:

Exhaustive search to check whether θ contains τ (Markov chain version).
(Assume that $\theta \in \mathcal{S}_N$, $\tau \in \mathcal{S}_k$, with $k < N$, and that
$\theta = F(\sigma, I, J)$ for some $\sigma \in \mathcal{S}_N(\tau)$ and $I, J \in \{1, \dots, N\}$.)
For each $l = 1, \dots, k$ in turn:
Loop through all possible values of $(i_1, i_2, \dots, i_k)$ with
$1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant N$ and $i_l = J$, and check whether
$\sigma_{i_1} \dots \sigma_{i_k}$ forms the pattern τ.

This algorithm can be performed in time $O(kN^{k-1})$. Similarly, in the Indirect Monte Carlo procedure, we check whether $\sigma^{\dagger i}$ avoids τ in time $O(N^{k-1})$:

Exhaustive search to check whether $\sigma^{\dagger i}$ contains τ.
(Assume that $\tau \in \mathcal{S}_k$, $\sigma \in \mathcal{S}_N(\tau)$, $0 \leqslant i \leqslant N$, and recall
$\sigma^{\dagger i} = (\sigma_1 \dots \sigma_i (N+1) \sigma_{i+1} \dots \sigma_N)$.)
Let h be the index such that $\tau_h = k$. Loop through all possible values of $(i_1, i_2, \dots, i_k)$ with $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant N$ and $i_h = i + 1$, and check whether $\sigma_{i_1} \dots \sigma_{i_k}$ forms the pattern τ.

This procedure works because if σ avoids τ and $\sigma^{\dagger i}$ contains τ, then the new "$N+1$" entry in the $(i+1)^{th}$ position of $\sigma^{\dagger i}$ must correspond to the largest entry of τ, i.e., any pattern map $\mathcal{M}$ of τ into $\sigma^{\dagger i}$ must have $\mathcal{M}(h) = i + 1$.

It is not immediately clear how to make significant improvements to the above algorithms for general patterns τ. Some progress has been made on this problem in [**1**]. Among other things, [**1**] describes clever algorithms for detecting any pattern of length 4 in time $O(N \log N)$. We learned about this work only in the final stages of preparation of this article, from one of the referees. We shall describe the algorithms that we actually used for the pattern 4231, which are simpler but less efficient (mainly $O(N^2)$, but see Proposition 3.3).

We begin with the fact that we can check 231-avoidance in $O(N)$ time. For this, we shall use the following algorithm, which employs a "stack" data structure. We can add items to and remove items from a stack on a "last in, first out" (LIFO) basis. (That is, any item added to the stack must be put on top of items that are already in the stack; only the top item can be removed from the stack at a given time; and we cannot change the order of items while they are in the stack.)

Greedy Stack Sort (GSS) Algorithm.
Given a permutation $\sigma = \sigma_1 \sigma_2 \dots \sigma_N$, this algorithm produces another permutation $\pi = \pi_1 \pi_2 \dots \pi_N$.

1. Start with an empty stack. Set $i = 1$ and $j = 1$. (The index i denotes the entry of σ currently being processed; the index j points to the entry of π that is next to be determined.)
2. If the stack is empty, then put σ_i in the stack (where it becomes the top item), and go to Step 4. Otherwise continue to Step 3.
3. Let T be the numerical value of the current top item in the stack. If σ_i is greater than T, then remove the top item from the stack, set π_j equal to T, increase j by 1, and go to Step 2.

Otherwise (i.e., if $\sigma_i < T$), put σ_i on top of the stack and go to Step 4.

4. If $i < N$, then increase i by 1 and go to Step 2. Otherwise (i.e., if $i = N$), continue to Step 5.
5. Set π_j equal to the top number in the stack, and remove the top item from the stack.
6. If $j < N$ then increase j by 1 and go to Step 5. Otherwise (i.e., if $j = N$), stop.

The following result is proven in [**12**, Exercise 5, Section 2.2.1]:

PROPOSITION 3.1. *Let $\sigma \in \mathcal{S}_N$, and let π be the permutation obtained after executing the GSS algorithm with input σ. Then π is the increasing permutation $12\dots N$ if and only if σ avoids the pattern* 231.

Observe that the running time of the GSS algorithm is $O(N)$, since each execution of Step 2, 3, or 5 moves an item either from σ to the stack or else from the stack to π, and exactly $2N$ such moves are performed in all. Thus we can check 231-avoidance in time $O(N)$. From this, it is easy to see that the following algorithm correctly checks 4231-avoidance:

Modified GSS Algorithm to check whether σ avoids 4231.
(Assume that $\sigma \in \mathcal{S}_N$.)
For each $m = 1, \dots, N-3$, do the following: Let $\sigma^{(m)}$ be the subsequence of $\sigma_{m+1}\sigma_{m+2}\dots\sigma_N$ consisting of those numbers that are less than σ_m. Let $\pi^{(m)}$ be the output obtained from running the GSS algorithm with input $\sigma^{(m)}$. If $\pi^{(m)}$ is the increasing permutation for every m, then σ avoids 4231.

Thanks to Proposition 3.1, the following result is clear.

PROPOSITION 3.2. *The Modified GSS Algorithm checks whether a given permutation of length N avoids the pattern* 4231 *in time $O(N^2)$.*

For our Indirect Monte Carlo with $\tau = 4231$, Proposition 3.1 leads to the following.

PROPOSITION 3.3. *Let $\sigma \in \mathcal{S}_N(4231)$. Then $W(\sigma)$ can be computed in time $O(N \log N)$.*

PROOF. Observe that if $\sigma^{\dagger i}$ avoids 4231, then $\sigma^{\dagger(i+1)}$ also avoids 4231 (this is true for any pattern τ in which τ_1 is the largest entry of τ). Therefore there is a unique nonnegative integer M such that $\sigma^{\dagger i}$ avoids 4231 for all $i > M$ and contains 4231 for all $i \leqslant M$. Then $W(\sigma) = N + 1 - M$. We can check whether a given i is greater than M by checking whether $\sigma^{\dagger i}$ avoids τ, and this can be done in time $O(N)$ since we only need to run the Modified GSS Algorithm with $m = i + 1$. Finally, we can find M by a bisection search that uses queries "Is i greater than M?" for at most $O(\log N)$ different values of i. □

4. Numerical results

4.1. Estimation of $L(4231)$. We chose forty values of N from 25 to 155, and for each N we ran the Markov chain of Section 2 to generate many random permutations from $\mathcal{S}_N(4231)$. More precisely, for each N, *we performed an initialization phase of 10^7 iterations, followed by a data collection phase consisting*

of 50 *batches of* 2000 *observations each, with* 5000 *iterations between each observation*. Although the approach is relatively standard among practitioners of Markov chain Monte Carlo, it is probably unfamiliar to some readers, so we describe it in the following paragraphs.

Initialization Phase: First, we chose an initial permutation $X(0)$ in $\mathcal{S}_N(4231)$, not necessarily from the uniform distribution. Theory tells us that if we run the Markov chain for a sufficiently large number of iterations T_0, then the distribution of $X(T_0)$ is close to the equilibrium distribution of the chain (in our case, this means that $X(T_0)$ is approximately uniformly distributed on $\mathcal{S}_N(4231)$). Ideally, one would like rigorous estimates of the T_0 needed for a given accuracy of approximation. This kind of time-to-equilibrium problem is difficult in general (especially if we want actual numbers instead of asymptotic estimates like $O(N^2)$), although substantial progress has been made on a variety of specific Markov chains (see, e.g., [**13, 17**]). Our approach in the present article is purely empirical. We hypothesized that $T_0 = 10^6$ would be large enough; and this hypothesis was consistent with some basic checks, such as plotting of batch means, as described in [**14**, Section 5.3]. (For significantly larger values of N, it would be wise to use more sophisticated statistical methods to check for slowly decaying correlations in the output.) For additional security, we increased T_0 to 10^7. So in our simulation runs, *we always discard the first* 10^7 *iterations to ensure that our Markov chain has "reached equilibrium"*. That is, we let $\sigma_{(0)} = X(10^7)$ and use $\sigma_{(0)}$ as our initial permutation when we start the data collection phase.

Data Collection Phase: Starting with the $\sigma_{(0)}$ produced by the initialization phase, we can assume that the chain is in (approximate) equilibrium, so we can treat our simulation as if it were coming from a Markov chain in equilibrium (in particular, it is stationary in the stochastic process sense). This is helpful for the statistical analysis, but it is still not as good as the observed output data $\{X(t)\}$ being independent. To deal with correlation in our output, we use the *Batch Means* method (e.g., see [**14**, Section 5.2.1]), as follows. Starting from $\sigma_{(0)}$, we run the chain for 50×10^7 iterations. The first 10^7 iterations form our first "batch"; the next 10^7 iterations form our second batch; and so on. Thus we obtain 50 batches of 10^7 iterations each. The batches do not overlap, but they are not independent. However, in light of the hypothesis that 10^6 iterations are enough to "forget the past" (as discussed in the previous paragraph), it is reasonable to expect that non-consecutive batches are almost independent, and that correlations between consecutive batches are not large. Within the k^{th} batch ($k = 1, \ldots, 50$), we can compute an estimate of $E(W(\sigma))$ by averaging the values of $W(X(t))$ for all $X(t)$ in this batch. However, a moment's thought will reveal that computing these 10^7 values of W is too much work, especially since $W(X(t))$ changes very slowly with t. To reduce our work significantly with just a small loss of statistical efficiency, we only compute $W(X(t))$ at every 5000^{th} iteration within the batch. Thus the k^{th} batch produces an estimate $\bar{W}^{[k]}$ of $E(W)$, defined by

$$\bar{W}^{[k]} := \frac{1}{2000} \sum_{t=1}^{2000} W(X(5000t))$$

and called the k^{th} *batch mean* (note that $2000 = 10^7/5000$ is the number of recorded observations per batch). Now the average of all the batch means produces our estimate $\hat{W}_N$ of $E(W) = |\mathcal{S}_{N+1}(4231)|/|\mathcal{S}_N(4231)|$, and the corresponding usual

N	$\hat{W}_N$	$\hat{V}_N$	N	$\hat{W}_N$	$\hat{V}_N$
25	8.10907	0.006533	50	9.05622	0.008599
26	8.16253	0.007520	53	9.13380	0.007805
27	8.23701	0.003612	57	9.19356	0.009365
28	8.28662	0.003903	61	9.29812	0.008750
29	8.33475	0.005219	66	9.38607	0.009959
30	8.37368	0.004448	70	9.44051	0.015951
31	8.43250	0.005861	73	9.51701	0.010449
32	8.46813	0.004838	80	9.55512	0.008491
33	8.50578	0.005521	85	9.64159	0.011583
34	8.55269	0.008820	91	9.68133	0.017521
35	8.61763	0.005403	97	9.72812	0.014874
36	8.65268	0.005977	103	9.82502	0.019832
37	8.68698	0.006668	112	9.85624	0.017183
38	8.71099	0.006462	116	9.87179	0.015270
39	8.74502	0.007992	125	9.96947	0.019569
40	8.78393	0.007117	134	9.99824	0.018018
41	8.80296	0.008329	140	9.99054	0.027671
42	8.84594	0.007498	145	10.07844	0.030349
43	8.88029	0.005821	150	10.15082	0.042680
46	8.96066	0.008746	155	10.17202	0.047939

TABLE 1. The Markov chain data generated for avoiding 4231. The estimates $\hat{W}_N$ (for $|\mathcal{S}_{N+1}(4231)|/|\mathcal{S}_N(4231)|$) and $\hat{V}_N$ (for the variance of $\hat{W}_N$), defined in Equation (4.1), are obtained from the batch means method after removing the first part of the output for initialization. See the text for details.

sample variance gives a useful estimate $\hat{V}_N$ of the variance of $\hat{W}_N$ (thanks to the approximate independence of the batches):

$$\hat{W}_N := \frac{1}{50}\sum_{k=1}^{50} \bar{W}^{[k]} \quad \text{and} \quad \hat{V}_N := \frac{1}{49}\sum_{k=1}^{50} (\bar{W}^{[k]} - \hat{W}_N)^2. \tag{4.1}$$

The values of $\hat{V}_N$ and $\hat{W}_N$ were computed from simulations of the Markov chain at 40 values of N from 25 to 155, based on 51×10^7 iterations for each value of N. The results are given in Table 1.

We are now ready to try to extrapolate the output of Table 1 to the $N \to \infty$ limit of $\hat{W}_N$, obtaining an estimate of $L(4231)$. A simple way to attempt this is to try to fit a line (or curve) to the graph of $y = \hat{W}_N$ versus $x = 1/N$; then the y-intercept will be our estimate of the limit. Indeed, there is a heuristic argument for this, as follows. Recalling the known result (1.5), it is reasonable to expect that for any pattern τ we have constants $A = A_\tau$ and $p = p_\tau$ such that

$$|\mathcal{S}_N(\tau)| \stackrel{?}{\sim} A\, N^{-p} L(\tau)^N \quad \text{as } N \to \infty. \tag{4.2}$$

Note that if (4.2) holds, then Equation (1.4) guarantees that p must be nonnegative. This suggests $|\mathcal{S}_{N+1}(4231)|/|\mathcal{S}_N(4231)| \approx L(4231)(1 + 1/N)^{-p}$. We choose to

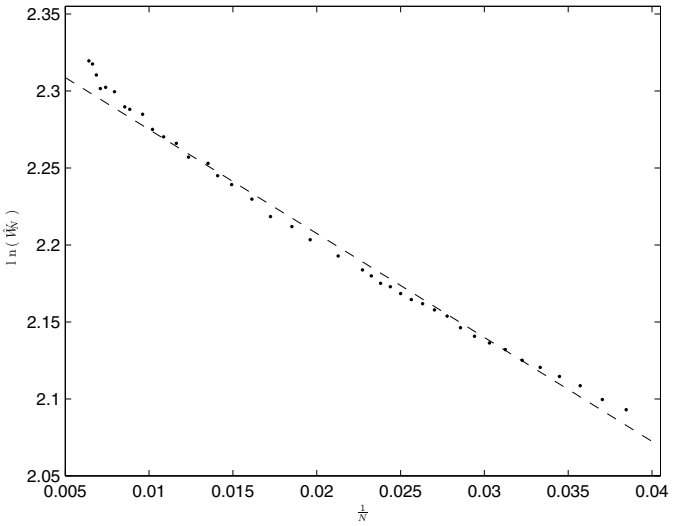


FIGURE 2. Plot of $\ln \hat{W}_N$ versus $1/N$, and the fitted line $\ln W = 2.3425 - 6.7518/N$.

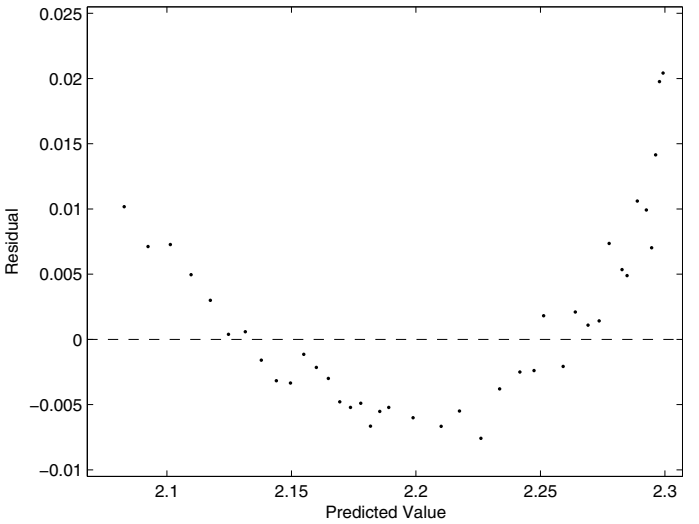


FIGURE 3. Residuals of Figure 2. Each solid circle in the graph is a point $(2.3425 - 6.7518/N,\ \ln \hat{W}_N - (2.3425 - 6.7518/N))$ for some N from Table 1.

take logarithms (to separate the interesting limit L from the "correction term"), obtaining $\ln(|\mathcal{S}_{N+1}(4231)|/|\mathcal{S}_N(4231)|) \sim \ln L(4231) - p\ln(1+1/N)$, which becomes

$$\ln\left(\frac{|\mathcal{S}_{N+1}(4231)|}{|\mathcal{S}_N(4231)|}\right) \approx \ln L(4231) - \frac{p}{N} \tag{4.3}$$

when we omit the $O(N^{-2})$ terms. (Of course, (4.3) is still just an educated guess!)

Accordingly, we first analyzed the data of Table 1 by performing a *weighted linear regression* (see for example [**15**, Section 9.2.1] for a brief relevant description) of the form $y = \beta_0 + \beta_1 x + \text{error}$, with $y = \ln \hat{W}_N$ and $x = 1/N$, using weights $1/\hat{V}_N$. See Figure 2. The regression line is $W = 2.3425 - 6.7518/N$, but it is not satisfactory. The plot of the data has visible curvature, which is strikingly shown in the residual plot, Figure 3. (The *residuals* are the quantities given by "data minus fit," i.e., $\ln \hat{W}_N - (2.3425 - 6.7518/N)$ in this case.)

Taking the curvature into account, we tried plotting $\ln \hat{W}_N$ against $1/\sqrt{N}$ (Figure 4). A line looked quite reasonable here (see also [**2**, Fig. 1]). The best linear fit is $\ln W = 2.4712 - 1.9175/\sqrt{N}$; the residuals, plotted in Figure 5, show considerably less pronounced curvature than those of Figure 3 (in particular, note the vertical scale).

To obtain an estimate of $L(4231)$ (corresponding to e^{β_0}), we removed the data points corresponding to the smaller values of N (on the grounds that the smaller values would show less asymptotic behaviour). Retaining only the largest 25 values of N ($40 \leqslant N \leqslant 155$), we obtain the fit $\ln W = 2.4657 - 1.8699/\sqrt{N}$, with a 95% confidence interval of $[2.461, 2.470]$ on the constant term β_0. This corresponds to a 95% confidence interval of $[11.71, 11.83]$ on $L(4231)$ (i.e., $\exp(\beta_0)$).

We also tried to account for the curvature of Figure 2 by fitting a quadratic curve, i.e., weighted linear regression of $\ln \hat{W}_N$ against $\beta_0 + \beta_1/N + \beta_2/N^2$. This also worked reasonably well, producing the fit $\ln W = 2.3706 - 9.8048/N + 68.624/N^2$ (the residuals are plotted in Figure 6).

As above, we omitted the smallest 15 of our 40 values of N to get our best extrapolation. This yielded the fit $\ln W = 2.3786 - 10.949/N + 103.28/N^2$, with a 95% confidence interval of $[2.371, 2.386]$ on the constant term β_0. This corresponds to a 95% confidence interval of $[10.71, 10.87]$ on $L(4231)$ (i.e., $\exp(\beta_0)$).

The last two models gave good fits, but rather different extrapolations for $L(4231)$. Without additional data (or rigorous analysis), it is impossible to tell which is "correct" (if either). We summarize our state of ignorance with the conservative "subjective" 95% confidence interval of $[10.71, 11.83]$ on $L(4231)$, which allows for model uncertainty as well as statistical error.

4.2. Avoiding patterns of length 5. It is known that the patterns of length 5 form 16 Wilf classes. (Two patterns τ and π are in the same Wilf class if $|\mathcal{S}_N(\tau)| = |\mathcal{S}_N(\pi)|$ for every N.) We list them in Table 2.

We wanted to investigate which classes have the largest and smallest Stanley–Wilf limits. In lieu of trying to estimate the limits, we settle for the easier problem of trying to rank the 16 values of $|\mathcal{S}_N(\tau)|$ as τ ranges over representatives from each class, for fixed N. For $N \leqslant 13$, exact enumeration is feasible [**21**]. For $N = 25$ and $N = 30$, we used the Direct Monte Carlo algorithm of Section 2 to estimate $p_N(k) := |\mathcal{S}_N(\tau^{[k]})|/N!$ for $\tau^{[k]}$ in the k^{th} Wilf class. For a given $\tau^{[k]}$, we form confidence intervals using classical statistics: if our estimate of $p_N(k)$ is $\bar{Y}_M$ (in the

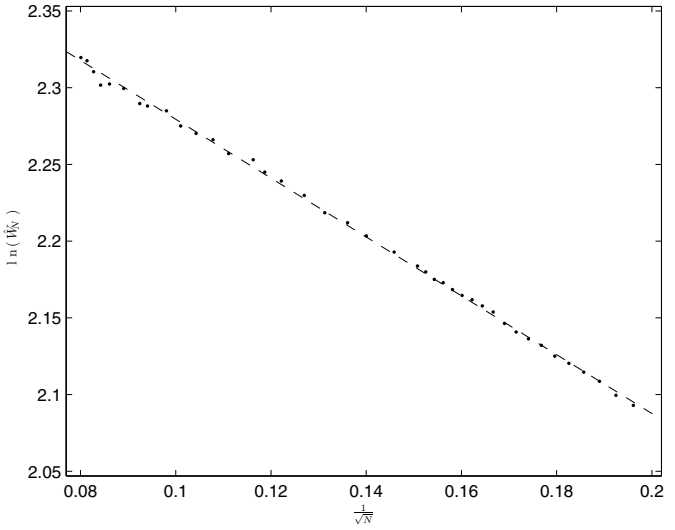


FIGURE 4. Plot of $\ln \hat{W}_N$ versus $1/\sqrt{N}$, and the fitted line $\ln W = 2.4712 - 1.9175/\sqrt{N}$.

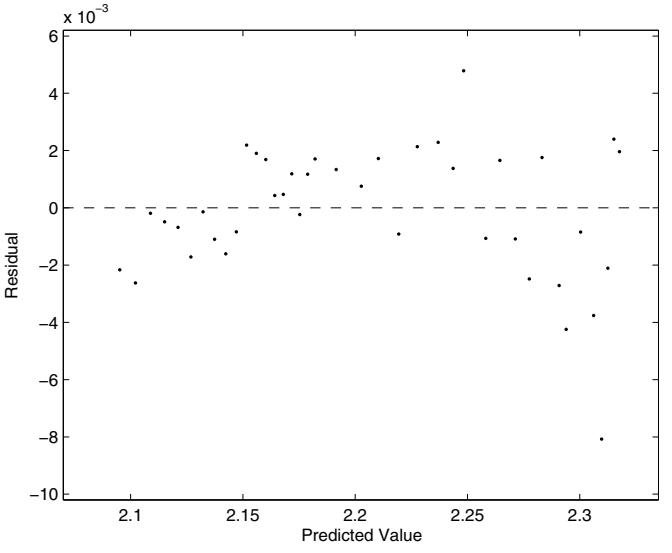


FIGURE 5. Residuals of Figure 4.

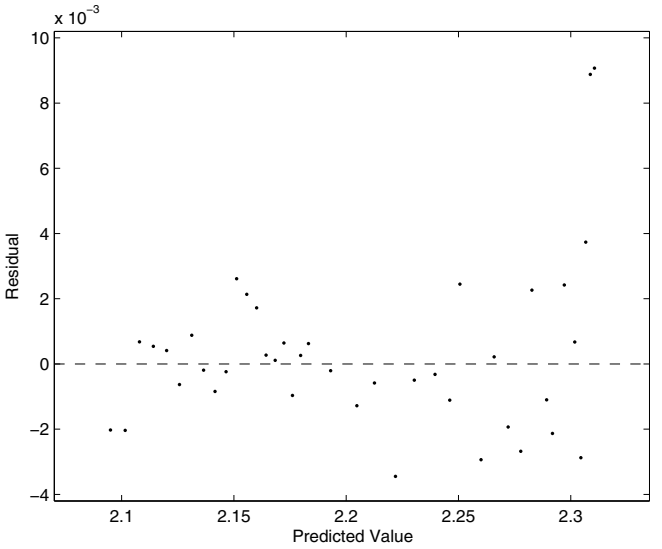


FIGURE 6. Residuals of the best fit of the form $\ln \hat{W} = \beta_0 + \beta_1/N + \beta_2/N^2$.

Class	Patterns
1	12345 54321 15432 23451 43215 51234 12543 32145 34521 54123
	12354 21345 45321 54312 21543 32154 34512 45123 21354 45312
2	12435 13245 53421 54231 13254 21435 45231 53412
3	12453 31245 35421 54213 21453 31254 35412 45213 12534 23145
	43521 54132 21534 23154 45132 43512
4	13425 14235 52431 53241
5	14253 31425 35241 52413 13524 24135 42531 53142
6	14325 52341
7	14523 32541 34125 52143
8	15234 23415 43251 51432 13452 25431 41235 53214
9	15243 32415 34251 51423 13542 24531 42135 53124
10	15423 32451 34215 51243 14532 23541 43125 52134
11	15324 24315 42351 51342 14352 25341 41325 52314
12	15342 24351 42315 51324
13	23514 25134 41532 43152 25413 31452 35214 41253
14	24153 31524 35142 42513
15	24513 31542 35124 42153 25143 32514 34152 41523
16	25314 41352

TABLE 2. The Wilf classes of patterns of length 5.

Class k	Estimate of $p_{25}(k)$	Estimate of $p_{30}(k)$
1	$3.90(1)\times10^{-4}$	$6.8(2)\times10^{-8}$
2	$4.36(1)\times10^{-4}$	$8.2(2)\times10^{-8}$
3	$2.70(1)\times10^{-4}$	$3.6(2)\times10^{-8}$
4	$2.96(1)\times10^{-4}$	$4.4(2)\times10^{-8}$
5	$2.50(1)\times10^{-4}$	$3.5(2)\times10^{-8}$
6	$5.10(2)\times10^{-4}$	$11.3(3)\times10^{-8}$
7	$2.46(2)\times10^{-4}$	$3.4(2)\times10^{-8}$
8	$2.32(1)\times10^{-4}$	$2.9(2)\times10^{-8}$
9	$2.15(1)\times10^{-4}$	$2.7(2)\times10^{-8}$
10	$2.52(1)\times10^{-4}$	$3.1(2)\times10^{-8}$
11	$2.21(1)\times10^{-4}$	$2.9(2)\times10^{-8}$
12	$2.88(1)\times10^{-4}$	$4.0(2)\times10^{-8}$
13	$1.79(1)\times10^{-4}$	$2.0(2)\times10^{-8}$
14	$1.60(1)\times10^{-4}$	$1.5(1)\times10^{-8}$
15	$2.06(1)\times10^{-4}$	$2.5(2)\times10^{-8}$
16	$1.42(1)\times10^{-4}$	$1.4(1)\times10^{-8}$

TABLE 3. Monte Carlo data for patterns of lengths 25 and 30. The number in parentheses gives the statistical uncertainty in the lowest significant digit shown. E.g., the 95% confidence interval on $p_{25}(7)$ is $[2.44 \times 10^{-4}, 2.48 \times 10^{-4}]$.

notation of Section 2, with M independent repetitions), then the (approximate) 95% confidence interval has endpoints $\bar{Y}_M \pm 1.96\sqrt{\bar{Y}_M(1-\bar{Y}_M)/M}$.

We used $M = 10^9$ repetitions for $N = 25$, and $M = 10^8$ for $N = 30$. We used exhaustive search to check pattern avoidance, which became significantly more time-consuming as N increased from 25 to 30. The results are shown in Table 3.

The 16 confidence intervals for $N = 25$ were all pairwise disjoint. This was not true for $N = 30$, so we did longer Monte Carlo runs for selected classes until the confidence intervals were all disjoint. This allowed us to be confident in our rankings of the values of $p_N(k)$ for each N. The rankings are shown in Figure 7.

We observe that the ranks change with N. This was known for $N = 12$ and $N = 13$ from exact enumeration [**21**], and our results show that this does not appear to be an isolated phenomenon. Nevertheless, class 6 seems like a good bet for the largest limit ($L(14325)$). The smallest limit is not as clear-cut, but class 16 ($L(25314)$) is our best guess from the data so far. Preliminary Markov chain Monte Carlo studies (with $N < 30$) suggest $L(14325) \approx 19$ and $L(25314) \approx 13$. (Recall from (1.5) that $L(12345) = 16$.)

Open Problem 3: Is it true that $L(25314) \leqslant L(\tau) \leqslant L(14325)$ for all patterns τ of length 5?

This problem becomes even more intriguing when we show the patterns graphically in Figure 8.

4.3. What does a typical 4231-avoiding pattern look like? To try to improve our understanding of the properties of 4231-avoiding permutations, we

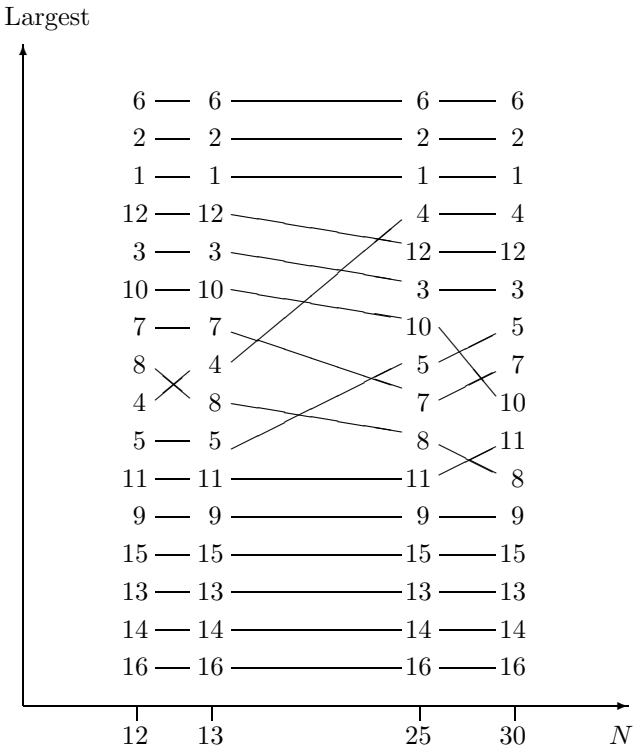


FIGURE 7. The 16 Wilf classes ranked according to the estimates of $|\mathcal{S}_N(\tau)|$ for different patterns τ of length 5, for $N = 12, 13, 25, 30$. Classes with larger estimates are higher up. E.g., class 12 has the fifth largest estimate for $N = 12$ and 13, and the sixth largest for $N = 25$ and 30.

wanted to see what a randomly chosen 4231-avoiding permutation looked like. To this end, we ran our Markov chain for 10^7 iterations and plotted the graph of the final permutation. We repeated this many times. Figures 9 and 10 are typical of the graphs that we saw. We were surprised by the amount of structure that the graphs displayed. (In particular, the graphs highlight the left-to-right maxima and right-to-left minima that were key to the discussion of 4231-avoidance in [**5**].)

Afterwards, we learned that Andrew Rechnitzer (unpublished) had seen similar structure in shorter 4231-avoiding permutations ($N \approx 60$) generated by a very different Monte Carlo method. All of our graphs had a similar appearance, which led us to conjecture that some kind of asymptotic shape exists in the sense of the following definition.

DEFINITION 4.1. An upper envelope for the pattern τ is a non-decreasing function $f\colon [0,1] \to [0,1]$ with the following property: for every $\epsilon > 0$, there exists $\alpha > 0$

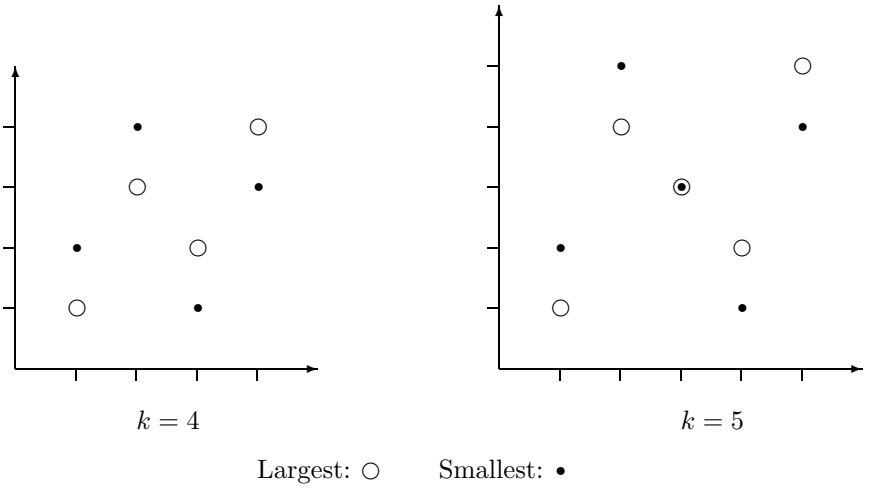


FIGURE 8. Graphs of patterns with largest and smallest Stanley–Wilf limits for $k = 4$ [1324 and 2413 respectively—proved; note that $L(1324) = L(4231)$] and for $k = 5$ [14325 and 25314 respectively—conjectured]. Removing the central point(s) in the $k = 5$ picture gives the $k = 4$ picture.

such that

$$P_N\left(\sigma_i > N\left[f\left(i/N\right) + \epsilon\right] \text{ for some } i \in \{1, \ldots, N\}\right) < e^{-\alpha N}$$

for all sufficiently large N (where P_N is the uniform distribution over $\mathcal{S}_N(\tau)$).

Clearly any upper envelope satisfies $f(x) \geqslant x$. This definition begs various questions: What is the smallest upper envelope for 4231? Can we say more about the distribution of random variables such as σ_1 or $\sigma_{N/2}$? Do other patterns besides 4231 have interesting upper envelopes? Some progress on these questions is being made by Atapour and Madras (in preparation).

Open Problem 4: What can be said about the upper envelopes for general patterns τ?

In the special case $\tau = 321$, it is not hard to use the characterization that any 321-avoiding permutation can be partitioned (not necessarily uniquely) into two increasing subsequences to show that the smallest upper envelope for $\mathcal{S}_N(321)$ is $f(x) = x$. That is, graphs of 321-avoiding permutations are usually very close to the diagonal. Dealing with 4231-avoiding permutations is harder, perhaps partly because they have no analogous convenient "constructive" characterization.

We conclude with a modest result along these lines.

PROPOSITION 4.2. *Let P_N be the uniform distribution on $\mathcal{S}_N(4231)$. Then $P_N(\sigma_1 > 0.71N)$ decays exponentially in N.*

PROOF. We first introduce some notation. For a permutation σ and a set of integers A, let $\sigma|^A$ be the subsequence of σ_i's in σ such that $\sigma_i \in A$. Also, let $\mathcal{H}_N(j) := \{\sigma \in \mathcal{S}_N(4231) : \sigma_1 = j\}$.

Suppose $\sigma \in \mathcal{H}_N(j)$. Then $\sigma|^{[1,j-1]}$ must belong to $\mathcal{S}_{j-1}(231)$. Also, $\sigma|^{[j+1,N]}$ must be a permutation of $\{j+1, \ldots, N\}$ that avoids 4231. Observe that the map

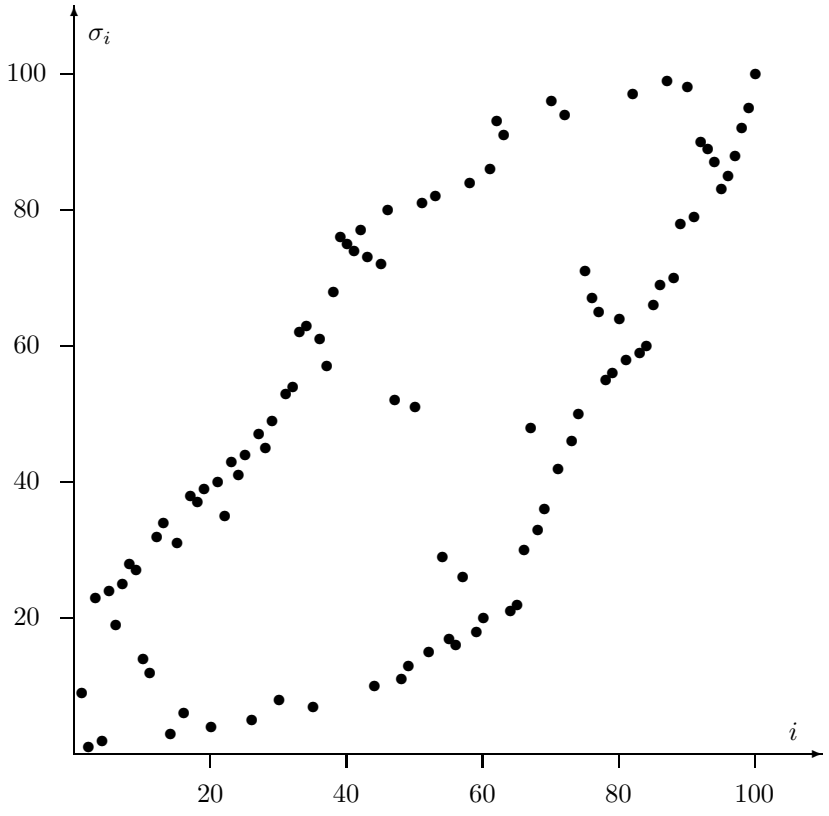


FIGURE 9. The graph of a random permutation σ in $\mathcal{S}_{100}(4231)$.

$\sigma \mapsto (\sigma|^{[1,j-1]}, \sigma|^{[j+1,N]})$ is at most $\binom{N-1}{j-1}$-to-1 on $\mathcal{H}_N(j)$. Therefore

$$\begin{aligned} |\mathcal{H}_N(j)| &\leqslant |\mathcal{S}_{j-1}(231)|\,|\mathcal{S}_{N-j}(4231)| \binom{N-1}{j-1} \\ &\leqslant 4^{j-1} L(4231)^{N-j} \binom{N-1}{j-1} \quad \text{[by (1.3) and (1.4)]}. \end{aligned}$$

Write L for $L(4231)$, and $\lfloor x \rfloor$ for the greatest integer less than or equal to x. For $1/2 \leqslant \alpha < 1$, we have

$$P_N(\sigma_1 \geqslant \alpha N) \;=\; \frac{\sum_{j \geqslant \alpha N} |\mathcal{H}_N(j)|}{|\mathcal{S}_N(4231)|} \;\leqslant\; \frac{N\,L^N}{4\,|\mathcal{S}_N(4231)|} \left(\frac{4}{L}\right)^{\alpha N} \binom{N-1}{\lfloor \alpha N \rfloor - 1}.$$

With the help of Stirling's formula, we see that $\limsup_N P_N(\sigma_1 \geqslant \alpha N)^{1/N} < 1$ if $4^\alpha/(\alpha^\alpha(1-\alpha)^{1-\alpha}) < L^\alpha$. Since $L \geqslant 9.47$ [**2**], a calculation shows that this inequality holds for $\alpha = 0.71$. □

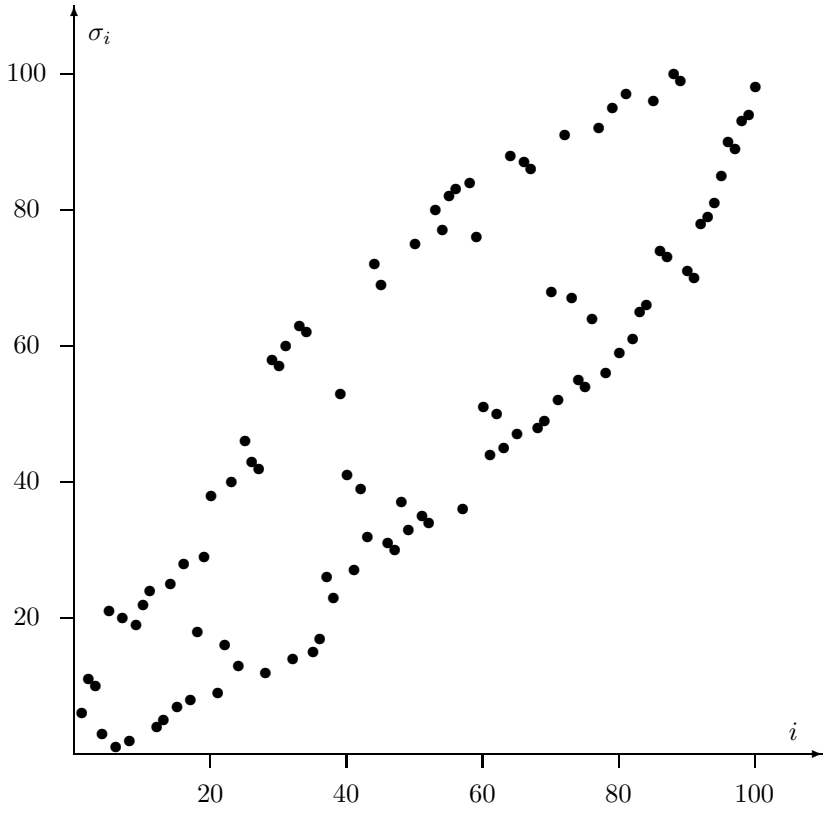


FIGURE 10. The graph of another random permutation σ in $\mathcal{S}_{100}(4231)$.

We remark that [**6**, Lemma 4.25] proves (in the notation of Proposition 4.2) that $P_N(\sigma_1 > \sigma_N)$ decays exponentially. This does not in itself seem to imply a result of the form of Proposition 4.2.

Acknowledgements. The first author wishes to thank Andrew Rechnitzer and Mike Zabrocki for introducing him to this field and for helpful discussions. He also acknowledges the hospitality of the Fields Institute for Research in Mathematical Sciences during the writing of this article. We also thank the referees for their comments.

References

[1] M. H. Albert, R. E. L. Aldred, M. D. Atkinson, and D. A. Holton, *Algorithms for pattern involvement in permutations*, pp. 355–367 in *Algorithms and Computation (Christchurch, 2001)*, no. 2223 in Lecture Notes in Comput. Sci., Springer-Verlag, New York/Berlin, 2001.

[2] M. H. Albert, M. Elder, A. Rechnitzer, P. Westcott, and M. Zabrocki, *On the Stanley–Wilf limit of 4231-avoiding permutations and a conjecture of Arratia*, Adv. in Appl. Math. **36** (2006), 96–105.

[3] R. Arratia, *On the Stanley–Wilf conjecture for the number of permutations avoiding a given pattern*, Electronic J. Combin. **6** (1999), note N1, 4 pp.

[4] J. Backelin, J. West, and G. Xin, *Wilf equivalence for singleton classes*, Adv. in Appl. Math. **38** (2007), 133–148.
[5] M. Bóna, *Permutations avoiding certain patterns: The case of length 4 and some generalizations*, Discrete Math. **175** (1997), 55–67.
[6] M. Bóna, *Combinatorics of Permutations*, Chapman & Hall/CRC, Boca Raton, FL, 2004.
[7] L. Devroye, *Nonuniform Random Variate Generation*, Springer-Verlag, New York/Berlin, 1986.
[8] W. Feller, *An Introduction to Probability Theory and Its Applications*, Volume I, Wiley, New York, 3rd edition, 1968.
[9] M. Fekete, *Über die Vierteilun der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzähligen Koeffiizienten*, Math. Z. **17** (1923), 228–249.
[10] E. J. Janse van Rensburg and A. R. Rechnitzer, *High precision Monte Carlo determination of the growth constant of square lattice trees*, Phys. Rev. E **67** (2003), article no. 036116, 9 pp.
[11] S. Karlin and H. M. Taylor, *A First Course in Stochastic Processes*, Academic Press, New York, 1975.
[12] D. E. Knuth, *The Art of Computer Programming, Volume 1: Fundamental Algorithms*, Addison–Wesley, Reading, MA, 2nd edition, 1973.
[13] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov Chains and Mixing Times*, American Mathematical Society, Providence, RI, 2009. With a chapter by J. G. Propp and D. B. Wilson.
[14] N. Madras, *Lectures on Monte Carlo Methods*, no. 16 in Fields Institute Monographs, American Mathematical Society, Providence, RI, 2002.
[15] N. Madras and G. Slade, *The Self-Avoiding Walk*, Birkhäuser, Boston/Basel, 1993.
[16] A. Marcus and G. Tardos, *Excluded permutation matrices and the Stanley–Wilf conjecture*, J. Combin. Theory Ser. A **107** (2004), 153–160.
[17] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*, Springer-Verlag, New York/Berlin, 1993.
[18] A. Regev, *Asymptotic values for degrees associated with strips of Young diagrams*, Adv. in Math. **41** (1981), 115–136.
[19] R. Simion and F. W. Schmidt, *Restricted permutations*, European J. Combin. **6** (1985), 383–406.
[20] Z. Stankova, *Forbidden subsequences*, Discrete Math. **132** (1994), 291–316.
[21] Z. Stankova and J. West, *A new class of Wilf-equivalent permutations*, J. Algebraic Combin. **15** (2002), 271–290.

Department of Mathematics and Statistics, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada
E-mail address: madras@mathstat.yorku.ca
URL: http://www.math.yorku.ca/Who/Faculty/Madras

Department of Statistics and Actuarial Science, University of Waterloo, 200 University Avenue, Waterloo, Ontario N2L 3G1, Canada
E-mail address: h59liu@math.uwaterloo.ca

Contemporary Mathematics
Volume **520**, 2010

Analytic combinatorics in *d* variables: An overview

Robin Pemantle

ABSTRACT. Let $F(\mathbf{Z}) = \sum_{\mathbf{r}} a_{\mathbf{r}} \mathbf{Z}^{\mathbf{r}}$ be a rational generating function in the d variables $Z_1, \ldots, Z_d$. Asymptotic formulae for the coefficients $a_{\mathbf{r}}$ may be obtained via Cauchy's integral formula in $\mathbb{C}^d$. Evaluation of this integral is not as straightforward as it is in the univariate case. This paper discusses geometric techniques that are needed for evaluation of these integrals and surveys classes of functions for which these techniques lead to explicit and effectively computable asymptotic formulae.

1. Introduction

A body of work in the last decade addresses the problem of estimating the coefficients of a multivariate generating function. The survey paper [**PW08**] is filled with examples and practical advice on how to extract asymptotics from such a generating function. It focuses on the theoretically easiest cases, wherein lie most known combinatorial examples. By contrast, the present overview is concerned with the theoretical structure of the enterprise and focuses on the boundaries of knowledge in the more difficult sub-cases. In particular, if we go beyond the *combinatorial* case (all coefficients are nonnegative real), as is necessary for instance with quantum random walks and with the diagonal applications in [**RW08**], then locating the dominating critical points can be much more difficult; see Section 1.4, and equation (2.4) and following. Central results from a number of papers are collected here. Proofs are included, sketched or omitted, according to the extent that they enhance understanding or give an alternative to the published argument. The context of the multivariate problem begins with a summary of the comparatively well understood univariate case.

1.1. Analytic combinatorics in one variable. Analytic combinatorics is the application of analytic methods to problems in combinatorial enumeration. This typically occurs as follows. A combinatorial class is defined whose size depends on a parameter $n = 0, 1, 2, 3, \ldots$. Let C_n denote the size of the n^{th} class. The description of the class, often recursive in nature, allows for the derivation of the

2000 *Mathematics Subject Classification.* Primary 05A16; Secondary 41A60, 41A63.

Key words and phrases. Rational function, generating function, Morse theory, Cauchy integral, Fourier–Laplace integral.

This work was supported in part by NSF grant no. DMS-063821.

generating function $F(z) := \sum_{n=0}^{\infty} a_n z^n$, where most often $a_n = C_n$ or $C_n/n!$. To apply analytic methods, the formal power series F must be convergent in some domain and the analytic properties of the function it represents must be understood, either because F is represented as some combination of elementary functions or because estimates on F or $|F|$ may somehow be obtained. Cauchy's integral formula expresses a_n exactly as an integral $(2\pi\mathrm{i})^{-1} \int z^{-n-1} F(z)\,\mathrm{d}z$. In order to evaluate the integral, complex contour methods must be brought to bear. The method of *singularity analysis*, described at length in the recent book [**FS09**], provides tools for analyzing this integral asymptotically as $n \to \infty$. The outcome depends on the behavior of F near its singularities of smallest modulus. If F is poorly behaved, for example failing to have any extension beyond its disk of convergence, circle methods such as Darboux' give asymptotic bounds on a_n. In the case of algebraic or logarithmic singularities, entire asymptotic developments may be carried out; see [**FO90**] for a description of how analytic information about F near its dominant singularity may be converted, nearly automatically, to asymptotic information about $\{a_n\}$. The subclass of rational functions is particularly simple, resulting in a limited number of types of asymptotic behavior: finite sums of terms $p_\gamma(n)\gamma^n$, where γ is a positive real number and p_γ is a polynomial (or, in the periodic case, a quasipolynomial).

1.2. Several variables. Consider now a generating function $F(Z_1, \ldots, Z_d) = \sum_{\mathbf{r}} a_{\mathbf{r}} \mathbf{Z}^{\mathbf{r}}$ in several variables, where $\mathbf{r}$ ranges over d-tuples of (usually nonnegative) integers and $\mathbf{Z}^{\mathbf{r}}$ stands for the monomial $Z_1^{r_1} \cdots Z_d^{r_d}$. Such a function arises in combinatorial applications from counting problems in which the class to be counted is naturally indexed by several parameters. Examples abound in which such generating functions turn out to be elementary functions; a long list may be found, for example, in [**PW08**]. Interesting univariate generating functions are at least of algebraic complexity, often transcendental. In the multivariate realm, interesting applications abound, with rational generating functions whose analyses are nontrivial. In fact a great proportion of combinatorial applications lead to rational functions or to no closed form at all. Indeed, the scarcity of compelling examples appears chiefly responsible for the slow development of multivariate generating function analysis outside the realm of rational generating functions. Furthermore, the main technical difficulties are already encountered with rational functions. In any case, the main thrust of multivariate analytic techniques to date is the rational case (though singularity analysis is sometimes possible for implicitly defined algebraic or D-finite generating functions), and this will be assumed until the last section of the present survey.

Let

$$F(\mathbf{Z}) = \frac{P(\mathbf{Z})}{Q(\mathbf{Z})} = \sum_{\mathbf{r}} a_{\mathbf{r}} \mathbf{Z}^{\mathbf{r}} \tag{1.1}$$

be a d-variable rational generating function. Our objective is to estimate the coefficients $a_{\mathbf{r}}$ asymptotically. As in the univariate case, the coefficients $\{a_{\mathbf{r}}\}$ may be recovered from F via the multivariate Cauchy integral formula [**Hör90**, (2.2.3)]

$$a_{\mathbf{r}} = \frac{1}{(2\pi\mathrm{i})^d} \int_T \mathbf{Z}^{-\mathbf{r}} F(\mathbf{z}) \frac{\mathrm{d}\mathbf{Z}}{\mathbf{Z}}\,, \tag{1.2}$$

where the torus T is a product of sufficiently small circles about the origin in each coordinate and $\mathrm{d}\mathbf{Z}/\mathbf{Z}$ is $(Z_1 \cdots Z_d)^{-1}$ times the holomorphic volume form $\mathrm{d}z_1 \wedge \cdots \wedge \mathrm{d}z_d$.

The purpose of this note is to give an overview of the analytic and geometric techniques necessary for the evaluation of (1.2). In more than one variable, the asymptotic analysis of generating functions is much less well understood than in the univariate case. Effective algorithms to produce asymptotics exist only for certain subclasses. Multivariate rational functions exhibit a wide range of asymptotic behaviors, which are not yet fully classified. Some of the complex contour methods that are necessary for the evaluation of this integral when $d = 1$ have higher-dimensional counterparts. These involve deforming the contour to pass through or near points of stationary phase on the pole variety $\mathcal{V} := \{\mathbf{Z} : Q(\mathbf{Z}) = 0\}$. Typically, asymptotics as $|\mathbf{r}| \to \infty$ with $\hat{\mathbf{r}} := \mathbf{r}/|\mathbf{r}| \to \mathbf{r}_*$ will be determined by the geometry of $\mathcal{V}$ near a *dominating* point $\mathbf{Z}_*$ that depends on $\mathbf{r}_*$. The dominating point will be one of more points from a finite set of *critical* points for the log-linear function $-\sum_{j=1}^{d} (r_*)_j \log |Z_j|$ on $\mathcal{V}$. The easiest case is when $\mathbf{Z}_*$ is a smooth point of $\mathcal{V}$ and lies on the boundary of the domain of convergence of the power series $\sum_{\mathbf{r}} a_{\mathbf{r}} \mathbf{Z}^{\mathbf{r}}$; such critical points are called *minimal* in the terminology of [**PW02, PW04**]. This case is analyzed in [**PW02**] via elementary methods. The case where $\mathcal{V}$ is a normal crossing in a neighborhood of $\mathbf{Z}_*$ is analyzed in [**PW04**] via elementary methods and in [**BP04**] via multivariate residues. When $\mathcal{V}$ has a singularity at $\mathbf{Z}_*$ that is not a local self-intersection, the analysis is more difficult. The subclass of products of powers of locally quadratic and locally linear divisors is analyzed in [**BP08**]; this class contains the generating functions arising in connection with some well known random tiling models [**CEP96, PS05**]. The elementary methods of [**PW02, PW04**] appear sometimes to be *ad hoc* but can be better understood in light of the apparatus introduced in later work, such as [**BP08**] and [**BBBP08**]. A second purpose of this note, therefore, is to re-cast the earlier analyses in a Morse-theoretic framework, thereby explaining the choices of reparametrization of the integrals and the forms of the results.

1.3. Notation. The following conventions are in place in order to achieve some consistency of notation and make the interpretations of variables visually obvious. The dimension is always d. Boldface is used for vectors and lightface for their coordinates; thus $\mathbf{Z} := (Z_1, \ldots, Z_d)$. A logarithmic change of variables is often required, in which case a corresponding lower case variable will be employed, for example $\mathbf{Z} = \exp(\mathbf{z}) := (\exp(z_1), \ldots, \exp(z_d))$; functions such as exp, log and absolute value, when applied to vectors, are taken coordinatewise. In the logarithmic coordinates, it is sometimes required to separate the real and imaginary parts of the vector $\mathbf{z}$; we shall denote these by $\mathbf{z} := \mathbf{x} + \mathrm{i}\mathbf{y}$. In the exponential space, we have no need for this and in low dimensions will sometimes use (X, Y, Z) in place of (Z_1, Z_2, Z_3). Unitized vectors will be denoted with a hat: $\hat{\mathbf{r}} := \mathbf{r}/|\mathbf{r}|$, where some norm is understood; a number of norms are useful depending on the application; instances are the euclidean norm, the L^1-norm and the pseudo-norm $|\mathbf{r}| := |r_d|$.

For asymptotics, the big-O, little-o and asymptotic equivalence notation $\sim$ will be employed; thus $f \sim g$ if and only if $f = (1 + o(1))g$. In the case of an asymptotic series development, $f \sim \sum_n b_n g_n$ will mean that for all N we have $\left| f - \sum_{n=0}^{N-1} b_n g_n \right| = O(g_N)$. This is slightly nonstandard because it allows some

of the coefficients b_n to vanish, but we shall use it only when infinitely many are nonvanishing.

The most basic quantitative estimate on $\{a_{\mathbf{r}}\}$ is the exponential growth rate in a given direction. Define the rate in the direction $\mathbf{r}_*$ by

$$\beta(\mathbf{r}_*) := \lim |\mathbf{r}|^{-1} \log |a_{\mathbf{r}}| \tag{1.3}$$

if such a limit exists, where the limit is as $|\mathbf{r}| \to \infty$ with $\hat{\mathbf{r}} \to \mathbf{r}_*$. One can force this to be well defined by taking a limsup instead of a limit. In fact there are a number of natural reasons, discussed in the next section, why one would not expect a limit to exist. In some cases, the limit will exist but fail to behave as expected for certain non-generic choices of $\mathbf{r}_*$. For this reason, we define a slightly more general limsup exponential growth rate by allowing $\hat{\mathbf{r}}$ to vary in a neighborhood $\mathcal{N}$ of $\mathbf{r}_*$ and taking the infimum over such neighborhoods:

$$\overline{\beta}(\mathbf{r}_*) := \inf_{\mathcal{N}} \limsup_{\mathbf{r}\to\infty,\, \hat{\mathbf{r}}\in\mathcal{N}} |\mathbf{r}|^{-1} \log |a_{\mathbf{r}}| \ . \tag{1.4}$$

EXAMPLE 1.1. The generating function $F(x,y) := (x-y)/(1-x-y)$ enumerates differences between consecutive binomial coefficients:

$$a_{ij} = \binom{i+j-1}{i-1} - \binom{i+j-1}{j-1} .$$

By symmetry, $a_{nn} = 0$ for all n, so that if $\mathbf{r}_*$ is the diagonal direction then $\beta(\mathbf{r}_*)$ exists and is equal to $-\infty$; whereas $\overline{\beta}(\mathbf{r}_*)$ is the logarithmic growth rate $\lim_{n\to\infty}(2n)^{-1}\log\binom{2n}{n} = \log 2$.

1.4. Organization of remainder of paper. Section 2 is concerned with computing the exponential rate. A function $\beta_Q(\mathbf{r}_*)$ is introduced that is always an upper bound for $\overline{\beta}$ (Proposition 2.2) and is often equal to $\overline{\beta}$. The formulation of β_Q and the dominating points $\mathbf{Z}_*$, as well as the proof of Proposition 2.2, are the central topics of Section 2.

Section 3 is concerned with the case where the dominating point $\mathbf{Z}_*$ is a smooth point of the variety $\mathcal{V}$. In this case explicit formulae are known for the leading term, and the entire asymptotic series is effectively computable. There are sometimes difficulties in selecting the dominating point from among a finite set of *critical points*, which we denote by $\mathsf{mincrit}$. Results are discussed in two special cases when $\mathcal{V}$ is everywhere smooth: the case where $d = 2$ and the *combinatorial* case where $a_{\mathbf{r}} \geqslant 0$. In the latter case, $\mathsf{mincrit}$ is always nonempty.

Section 4 catalogues a number of results that hold when $\mathbf{Z}_*$ is not a smooth point of $\mathcal{V}$. The next simplest geometry is that of a self-intersection or *multiple point*. This is discussed in Sections 4.1–4.3. After this, one might expect the next simplest case to be an algebraic curve ($d = 2$) with a cusp or other nore complicated singular point. However, as shown in [**BP08**], singularities in dimension 2 other than self-intersections are non-hyperbolic and cannot therefore contribute to the asymptotic expansion. Section 4.4 concentrates therefore on a three-dimensional example.

Returning to the problem of the exponential rate, Section 5 addresses the conjectured behavior of $\overline{\beta}$ in cases not covered by the results in the remainder of the paper. A modified version $\hat{\beta}_Q$ of β_Q is formulated that agrees with β_Q when the dominating critical point $\mathbf{Z}_*$ is minimal. Counterexamples from Sections 2–4 are

catalogued, after which a weak converse to the upper bound in Proposition 2.2 (with β_Q replaced by $\hat{\beta}_Q$) is conjectured.

2. Exponential estimates

The crudest estimate of $a_{\mathbf{r}}$ that is still informative is the exponential rate of growth or decay. If we are unable to compute or estimate $\overline{\beta}$, then our quantitative understanding of $\{a_{\mathbf{r}}\}$ is certainly quite poor! In statistical mechanical models, $\beta(\mathbf{r}_*)$ has an interpretation as an *entropy function* (cf. [**Ell85**, Section II.4]), or *large deviation rate*. In combinatorial applications, β is the exponential growth rate of a *partition function*, this being a (weighted) sum over a combinatorial class. In this section, we discuss how to "read off" a rate function $\beta_Q(\mathbf{r}_*)$ from the denominator of $F = P/Q$, which is always to be an upper bound for $\overline{\beta}$ (Proposition 2.2) and is often equal to $\overline{\beta}$.

2.1. Multidimensional contour deformations. The integrand in (1.2) that we denote by $\omega := \mathbf{Z}^{-\mathbf{r}} F(\mathbf{Z})\, \mathrm{d}\mathbf{Z}/\mathbf{Z}$ is holomorphic on the domain

$$\mathcal{M} := \mathbb{C}^d \setminus (\mathcal{V} \cup \{\mathbf{Z} : Z_1 \cdots Z_d = 0\}) \ .$$

This is an open subset of $\mathbb{C}^d$, hence a real $(2d)$-manifold. Any holomorphic d-form has $\mathrm{d}\omega = 0$, from which it follows by Stokes' Theorem that $\int_{\mathcal{C}} \omega$ depends only on the homology class of $\mathcal{C}$ in $H_d(\mathcal{M})$. Intuitively, this says that any deformation of T within $\mathcal{M}$ leaves the integral unchanged; technically, homology is weaker than homotopy, which means that there are equivalent chains of integration not obtained via deformation, though these are not usually needed (this is briefly discussed in Section 5).

Define a height function $h = h_{\mathbf{r}_*} : \mathcal{M} \to \mathbb{R}$, depending on $\mathbf{r}_*$, as the dot product of $-\mathbf{r}_*$ with the coordinatewise log-modulus:

$$h(\mathbf{Z}) = -\mathbf{r}_* \cdot \log |\mathbf{Z}| \ .$$

Let $\mathcal{M}^a$ denote the set $\{\mathbf{Z} \in \mathcal{M} : h(\mathbf{Z}) \leqslant a\}$ of points up to height a and let $\iota = \iota_a$ denote the inclusion map of $\mathcal{M}^a$ into $\mathcal{M}$; for $b < a$, the homology group $H_d(\mathcal{M}^b)$ maps naturally into $H_d(\mathcal{M}^a)$ via ι_*. The following estimates for any $a > b$ are immediate.

$$\int_{\mathcal{C}} \omega = O(e^{a|\mathbf{r}|}), \quad \text{if } [\mathcal{C}] \in H_d(\mathcal{M}^a); \tag{2.1a}$$

$$\left| \int_{\mathcal{C}} \omega - \int_{\mathcal{D}} \omega \right| = O(e^{b|\mathbf{r}|}), \quad \text{if } [\mathcal{C}] = [\mathcal{D}] \in H_d(\mathcal{M}^a, \mathcal{M}^b) \, . \tag{2.1b}$$

We may interpret (2.1b) as saying that ω has well defined integrals on relative homology classes in $H_d(\mathcal{M}^a, \mathcal{M}^b)$, the value of the integral being taken to be an equivalence class under differences by $O(e^{b|\mathbf{r}|})$.

Our deformations will be guided by the following heuristic. The chief difficulty in estimating such an integral is that the integrand may be much bigger than the integral, with rapid oscillation leading to significant cancellation. To address these problems we therefore attempt to:

> Deform the chain of integration so as to minimize the maximum over the chain of the modulus of the integrand.

To obtain asymptotic estimates, we must do this simultaneously for many values $\mathbf{r}$. If $|\mathbf{r}| \to \infty$ with $\hat{\mathbf{r}} \to \mathbf{r}_*$, then the exponential factor in the integral will be maximized where h is maximized. This suggests that we deform the contour so as to minimize $\sup_{\mathbf{Z}\in\mathcal{C}} h(\mathbf{Z})$. The deformations are constrained to lie in $\mathcal{M}$, that is, to avoid $\mathcal{V}$. Because of this, the minimizing contour is not achieved in $\mathcal{M}$ but is rather a limit of contours in $\mathcal{M}$, and touches $\mathcal{V}$ at one or more points $\mathbf{Z}_*$. A little calculus shows that $\mathbf{Z}_*$ must be a *stationary phase* point for h on $\mathcal{V}$, that is, dh restricted to $\mathcal{V}$ must vanish at $\mathbf{Z}_*$. At a stationary phase point, locally there is no cancellation due to oscillation, which justifies the prior assertion that minimizing the maximum modulus solves the oscillation problem as well. The remainder of the heuristic is that the minimized integral will be tractable. We shall see that this occurs in many families of cases, provided that we are careful with the interpretation of the integral on the limiting chain, which is not in $\mathcal{M}$.

2.2. Laurent series. It costs little and includes more applications if we extend the scope from power series to Laurent series. Formal Laurent series $\sum_{\mathbf{r}\in\mathbb{Z}^d} a_{\mathbf{r}}\mathbf{Z}^{\mathbf{r}}$ are not as nice as formal power series because there is no well defined formal multiplication. However, for Laurent series expansions of rational functions, convergence will occur on certain domains, allowing formal operations to be defined by the corresponding analytic operations, and allowing analytic methods still to be used.

Corresponding to each rational function are a number of Laurent series, each convergent on a different domain. The following facts about Laurent series and amoebas of polynomials may be found in [**GKZ94**, Chapter 6]; for a complete proof of Cauchy's formula in a poly-annulus, see [**Ran86**] or [**Pem09a**, Section 8.2].

Let $\mathrm{ReLog}(\mathbf{Z}) := \log|\mathbf{Z}| = (\log|Z_1|, \ldots, \log|Z_d|)$ denote the coordinatewise log-modulus of $\mathbf{Z}$. If Q is any polynomial in d variables, let $\mathsf{amoeba}(Q)$ denote the image of its zero set $\mathcal{V}$ under ReLog. Two examples with $d = 2$ are given in Figure 1.

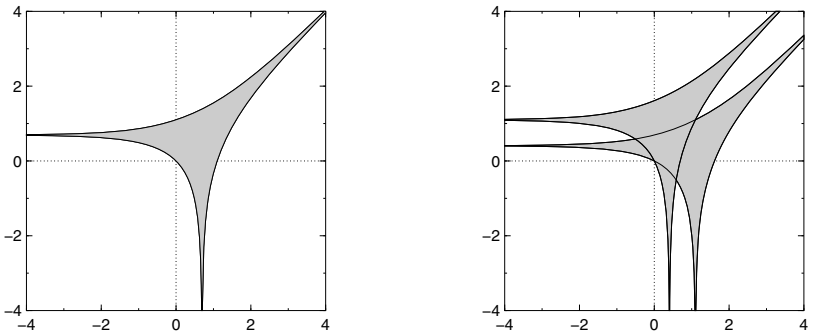


FIGURE 1. Two amoebae.
(a) $\mathsf{amoeba}(2 - X - Y)$. (b) $\mathsf{amoeba}((3 - X - 2Y)(3 - 2X - Y))$.

A general description of the amoeba and its relation to the various Laurent expansions for rational functions F with denominator Q is given by the following proposition.

PROPOSITION 2.1 ([**BP08**, Proposition 2.2]). *The connected components of* $\mathbb{R}^d \setminus \mathsf{amoeba}(Q)$ *are convex open sets. The components are in bijective correspondence with Laurent series expansions for* $1/F$*, as follows. For any Laurent series*

expansion of $1/F$, the open domain of convergence is precisely $\mathrm{ReLog}^{-1} B$ *where B is a component of $\mathbb{R}^d \setminus \mathsf{amoeba}(Q)$. Conversely, if B is such a component, a Laurent series $1/F = \sum a_{\mathbf{r}} \mathbf{Z}^{\mathbf{r}}$ convergent on B may be computed by the formula*

$$a_{\mathbf{r}} = \frac{1}{(2\pi \mathrm{i})^d} \int_{\mathbf{T}} \mathbf{Z}^{-\mathbf{r}-\mathbf{1}} \frac{1}{F(\mathbf{Z})} \, \mathrm{d}\mathbf{Z} \,,$$

where $\mathbf{T}$ is the torus $\mathrm{ReLog}^{-1}(\mathbf{x})$ *for any $\mathbf{x} \in B$. Changing variables to $\mathbf{Z} = \exp(\mathbf{z})$ and $\mathrm{d}\mathbf{Z} = \mathbf{Z}\,\mathrm{d}\mathbf{z}$ gives*

$$a_{\mathbf{r}} = \frac{1}{(2\pi \mathrm{i})^d} \int_{\mathbf{x}+\mathrm{i}\mathbf{t}} e^{-|\mathbf{r}|(\hat{\mathbf{r}}\cdot\mathbf{z})} \frac{1}{f(\mathbf{z})} \, \mathrm{d}\mathbf{z} \,, \tag{2.2}$$

where $f = F \circ \exp$ and $\mathbf{t}$ is the torus $\mathbb{R}^d/(2\pi\mathbb{Z})^d$. We remark that the separation of $\mathbf{r}$ into $|\mathbf{r}|\ \hat{\mathbf{r}}$ will be convenient when we send $\mathbf{r}$ to ∞ with $\hat{\mathbf{r}}$ held roughly constant.

The integral in (2.2) is of Fourier–Laplace type: $\int_{\mathcal{C}} e^{-\lambda\phi(\mathbf{z})} f(\mathbf{z})\,\mathrm{d}\mathbf{z}$ for some phase function ϕ, amplitude function f and chain of integration $\mathcal{C}$. The term "Fourier–Laplace" is used because the distinction between Fourier-type integrals (ϕ is purely imaginary on $\mathcal{C}$) and Laplace-type integrals (ϕ is real and nonnegative on $\mathcal{C}$) vanishes when the chain $\mathcal{C}$ is deformed in complex d-space (see [**PW09**] for details). The coefficients $\{a_{\mathbf{r}}\}$ may be viewed as a kind of Fourier transform of the logarithmic generating function f. A rigorous version of this appears in [**BP08**, Section 6]. In the present paper we shall use this interpretation only to give a second viewpoint on various formulae. This is because of the considerable technical difficulties in dealing with nonconvergent Fourier integrals as well as with discretization of the Fourier parameter.

Given a component B of the complement of the amoeba of Q, and given a real unit vector $\mathbf{r}_*$, define

$$\beta_Q(\mathbf{r}_*) := \inf\{-\mathbf{r}_* \cdot \mathbf{x} : \mathbf{x} \in \overline{B}\}. \tag{2.3}$$

Unless the closure of B fails to be strictly convex, and as long as $-\mathbf{r}_* \cdot \mathbf{x}$ is bounded from below on B, there is a unique point of $\overline{B}$ at which this minimum is attained. This point $\mathbf{x}_*$ is called the *minimizing point* for $\mathbf{r}_*$ and lies on the common boundary of B and $\mathsf{amoeba}(Q)$. If we choose only contours of the form $\mathbf{x} + \mathrm{i}\mathbf{t}$ then h will be constant on our contour, and it is clear that the maximum height is minimized when $\mathbf{x} \to \mathbf{x}_*$. Sending $\mathbf{x} \to \mathbf{x}_*$ in (2.2) immediately implies the following proposition.

PROPOSITION 2.2. *Let $F = P/Q$ be a rational function, and let $\sum_{\mathbf{r}} a_{\mathbf{r}} \mathbf{Z}^{\mathbf{r}}$ be the Laurent expansion of F corresponding to the component B of $\mathsf{amoeba}(Q)^c$. Define $\beta_Q \in [-\infty, \infty)$ by* (2.3)*. Then for any real unit vector $\mathbf{r}_*$,*

$$\overline{\beta}(\mathbf{r}_*) \leqslant \beta_Q(\mathbf{r}_*) \,.$$

REMARK 2.3. Computation of β_Q is semi-algebraic and hence effective; see, for instance, [**The02**, Section 2.2]).

2.3. Stratified spaces. Any algebraic variety admits a *Whitney stratification*. This is a partition into finitely many manifolds $\{S_\alpha\}$, called *strata*, satisfying two conditions. The first is that for distinct α, β, either S_α is disjoint from the closure of S_β or contained in it. The second is a condition on the limits of tangent spaces at points on the boundary of S_β; the reader is referred to [**PW09**, Definition 2.1] or [**GM88**, Section I.1.2] for a statement of this condition and its consequences.

DEFINITION 2.4 (critical and minimal points). A smooth function f on a stratified space X is said to have a *critical* point at p if $\mathrm{d}f|_S(p) = 0$ where S is the stratum of X containing p. In other words, p must be a critical point for the restriction of f to S.

The set of critical points of each stratum is algebraic, with membership defined by the *critical point equations*. These say that $\mathbf{x} \in S$ and that $\nabla h(\mathbf{x})$ is orthogonal to the tangent space $T_{\mathbf{x}}(S)$. When S is a k-dimensional stratum and the ambient space has dimension n, there are $n-k$ equations for $\mathbf{x} \in S$ and k equations for $\nabla h(\mathbf{x}) \perp T_{\mathbf{x}}(S)$. Thus the set of critical points of S is zero-dimensional for any S. If S has complex structure, then $n = 2d$, $k = 2\ell$, and there are $d - \ell$ complex equations for $\mathbf{x} \in S$ and ℓ complex equations for $\nabla h(\mathbf{x}) \perp T_{\mathbf{x}}(S)$.

Given $F = P/Q$, a component B of $\mathsf{amoeba}(Q)^c$, a direction $\mathbf{r}_*$ and a minimizing point $\mathbf{x}_*$ as above, define the set of *minimal* critical points by

$$\mathsf{mincrit}(Q, \mathbf{r}_*) := \{\mathbf{Z}_* \in \mathcal{V} : \operatorname{ReLog} \mathbf{Z}_* = \mathbf{x}_* \text{ and } \mathbf{Z}_* \text{ is a critical point for } h_{\mathbf{r}_*} \text{ on } \mathcal{V}\}. \tag{2.4}$$

A consequence of Theorem 3.1 in the next section is that Proposition 2.2 is sharp when there is minimal critical point $\mathbf{Z}_*$ at which $\mathcal{V}$ is smooth.

PROPOSITION 2.5. *Let $F = P/Q$ be a rational Laurent series and suppose there is a minimal critical point $\mathbf{Z}_*$ that is a smooth point of $\mathcal{V}$ with $P(\mathbf{Z}_*) \neq 0$. Then $\overline{\beta}(\mathbf{r}_*) = \beta_Q(\mathbf{r}_*)$.*

The following partial converse to this will be proved in Section 2.4 along with Theorem 2.8. Note that the computation of $\mathsf{mincrit}(\mathbf{r}_*)$ is algebraic and effective.

PROPOSITION 2.6. *If $\mathsf{mincrit}(\mathbf{r}_*)$ is empty then $\overline{\beta}(\mathbf{r}_*) < \beta_Q(\mathbf{r})$.*

In the remaining cases, when $\mathsf{mincrit}(\mathbf{r}_*)$ contains no smooth point but is not empty, it can be difficult to tell whether $\overline{\beta} = \beta_Q$. In most cases this can be resolved by computing the normal cone $\mathbf{N}_* = \mathbf{N}_*(\mathbf{Z}_*)$ associated to each $\mathbf{Z}_* \in \mathsf{mincrit}$. A self-contained definition of this cone is too lengthy to give here, but the gist is as follows.

DEFINITION 2.7 (normal cones). Let B be a component of $\mathsf{amoeba}(Q)^c$, let $\mathbf{x}_*$ be the minimizing point for $\mathbf{r}_*$ and let $\mathbf{Z}_* \in \mathsf{mincrit}(\mathbf{r}_*)$. Let $\mathbf{K} = \mathbf{K}(\mathbf{r}_*)$ denote the (geometric) tangent cone to B at $\mathbf{x}_*$, that is, $\mathbf{y} \in \mathbf{K}$ if and only if $\hat{\mathbf{y}}$ is the limit of normalized secants $(\mathbf{b} - \mathbf{x}_*)/|\mathbf{b} - \mathbf{x}_*|$. Denote by $\mathbf{N}_*(\mathbf{r}_*)$ the (outward) dual cone to $\mathbf{K}$, that is, the cone of vectors $\mathbf{v}$ such that $\mathbf{v} \cdot \mathbf{b} \leqslant 0$ for all $\mathbf{b} \in \mathbf{K}$. It is shown in [**BP08**, Definition 2.13] that for each $\mathbf{Z}_*$ there is a naturally defined cone $\mathbf{K}(\mathbf{Z}_*)$ that contains $\mathbf{K}(\mathbf{r}_*)$. Let $\mathbf{N}_*(\mathbf{Z}_*)$ denote the dual cone to $\mathbf{K}(\mathbf{Z}_*)$. Note that by duality, $\mathbf{N}_*(\mathbf{Z}_*) \subseteq \mathbf{N}_*(\mathbf{r}_*)$.

At the moment, the best known sufficient criterion for $\overline{\beta} < \beta_Q$ is given in the following result, proved in Section 2.4; see Section 5 for a discussion and conjecture as to how close this criterion is to being sharp.

THEOREM 2.8 (upper bound). *Given $F, P, Q, B, \mathbf{r}_*$ and $\mathbf{x}_*$ as above, let $h = h_{\mathbf{r}_*}$ and $c := h(\mathbf{x}_*)$, and let $\mathcal{C}$ be the chain of integration in* (1.2)*. Then*

(i) *There is an $\epsilon > 0$ such that the cycle $\mathcal{C}$ is homologous in $H_d(\mathcal{M}^{c+\epsilon}, \mathcal{M}^{c-\epsilon})$ to a sum of relative cycles $\mathcal{C}(\mathbf{Z}_*)$ supported in arbitrarily small neighborhoods of points $\mathbf{Z}_* \in \mathsf{mincrit}(\mathbf{r}_*)$ also satisfying $\mathbf{r}_* \in \mathbf{N}_*(\mathbf{Z}_*)$;*

(ii) *If* $\mathsf{mincrit}(\mathbf{r}_*)$ *is empty, or contains only points* $\mathbf{Z}_*$ *with* $\mathbf{r}_* \notin \mathbf{N}_*(\mathbf{Z}_*)$, *then there is an* $\epsilon > 0$ *and a neighborhood* $\mathcal{N}$ *of* $\mathbf{r}_*$ *such that*

$$|a_{\mathbf{r}}| = O\left(e^{(c-\epsilon)|\mathbf{r}|}\right) ,$$

as $\mathbf{r} \to \infty$ *with* $\hat{\mathbf{r}} \in \mathcal{N}$. *It follows in this case that*

$$\overline{\beta} < \beta_Q .$$

There are natural examples in which $\mathsf{mincrit}$ is empty; see for instance Example 3.6 below, or [**BP08**, Example 2.19]. On the other hand, nonnegativity of the coefficients $\{a_{\mathbf{r}}\}$ is sufficient to assure that part (*ii*) of Theorem 2.8 does not apply: $\mathsf{mincrit}$ is nonempty and contains a point $\mathbf{Z}_*$ with $\mathbf{r}_* \in \mathbf{N}_*(\mathbf{Z}_*)$.

PROPOSITION 2.9 (combinatorial case). *Suppose* $a_{\mathbf{r}} \geqslant 0$ *for all* $\mathbf{r}$. *Then for each* $\mathbf{r}_*$ *with unique minimizing* $\mathbf{x}_* \in \partial B$, *the real point* $\mathbf{Z}_* := \exp(\mathbf{x}_*)$ *is in* $\mathsf{mincrit}(\mathbf{r}_*)$ *and satisfies* $\mathbf{r}_* \in \mathbf{N}_*(\mathbf{Z}_*)$.

PROOF. Meromorphicity of F together with nonnegativity of F implies the presence of some pole of F on the torus $\mathrm{ReLog}^{-1}(\mathbf{x})$ for each $\mathbf{x} \in \partial B$; nonnegativity of the coefficients then implies that the positive real point $\mathbf{Z} := \exp(\mathbf{x})$ is in $\mathcal{V}$; see [**PW02**, Theorem 6.1] for further details on this step. We see from this that $\mathbf{Z}_* \in \mathcal{V}$, and in fact that the entire image of ∂B under the exponential map is in $\mathcal{V}$. It follows from the theory of hyperbolic functions that $\mathbf{K}(\mathbf{Z}_*) \subseteq \mathbf{K}(\mathbf{r}_*)$ and hence $\mathbf{N}_*(\mathbf{Z}_*) = \mathbf{N}_*(\mathbf{r}_*)$. It is automatic from the definition of $\mathbf{x}_*$ as a minimizing point for $\mathbf{r}_*$ that $\mathbf{r}_* \in \mathbf{N}_*(\mathbf{r}_*)$. Therefore, $\mathbf{r}_* \in \mathbf{N}_*(\mathbf{Z}_*)$, as desired. By [**BP08**, Proposition 2.22], this also implies that $\mathbf{Z}_* \in \mathsf{mincrit}(\mathbf{r}_*)$, finishing the proof. □

EXAMPLE 2.10 (large deviations). Let $\{p(\mathbf{r}) : \mathbf{r} \in \mathbb{Z}^{d-1}\}$ be a collection of nonnegative numbers summing to one. Assume that this probability distribution has finite moment generating function:

$$\phi(\mathbf{u}) := \sum_{\mathbf{r}} p(\mathbf{r}) e^{\mathbf{u}\cdot\mathbf{r}} < \infty, \qquad \text{for all } \mathbf{u} \in \mathbb{R}^{d-1} .$$

Let

$$F(\mathbf{Z}) := \frac{1}{1 - Z_d\, \phi(Z_1, \ldots, Z_{d-1})}$$

be the spacetime generating function for a random walk with steps governed by p. Thus, if $\mathbb{P}$ denotes the law of such a walk with partial sums $\{S_n\}$, and $\mathbf{Z}^{(\mathbf{r},n)}$ denotes $Z_1^{r_1} \cdots Z_{d-1}^{r_{d-1}} Z_d^n$, then

$$F(\mathbf{Z}) = \sum_{n=0}^{\infty} \sum_{\mathbf{r}} \mathbb{P}(S_n = \mathbf{r}) \mathbf{Z}^{(\mathbf{r},n)} .$$

The real surface $Q = 0$ is the graph over $\mathbb{R}^{d-1}$ of $1/\phi$; the region $\{\mathbf{Z} : Z_d < 1/\phi(Z_1, \ldots, Z_{d-1})\}$ is a component of the complement of the amoeba of Q; and each $\mathbf{Z} \in \partial B$ is a smooth point of $\mathcal{V}$ and is in $\mathsf{mincrit}(\mathbf{r}, 1)$, where $\mathbf{r}$ is the mean of the *tilted* distribution defined by

$$p_{\mathbf{Z}}(\mathbf{r}) = \frac{\mathbf{Z}^{\mathbf{r}} p(\mathbf{r})}{\sum_{\mathbf{r}'} \mathbf{Z}^{\mathbf{r}'} p(\mathbf{r}')} .$$

The preceding facts are not hard to show and may be found in [**Pem09b**]. It follows that for any $\mathbf{r}$ that is the mean of a tilted distribution $p_{\mathbf{Z}}$, the point $\mathbf{x} = \log \mathbf{Z}$ satisfies (2.3). Thus $\beta_Q(\mathbf{r}) = -\mathbf{r} \cdot \mathbf{x}$, and it follows from Proposition 2.5 that

$\overline{\beta}(\mathbf{r}) = -\mathbf{r}\cdot\mathbf{x}$. We may view this as a large deviation principle for sums of i.i.d. integer vectors with small-tailed distributions:

$$\limsup_{n\to\infty} \frac{1}{n}\log \mathbb{P}\left(\frac{S_n}{n} \in B(\mathbf{r},\epsilon)\right) \to -\mathbf{r}\cdot\mathbf{x}$$

as $\epsilon \downarrow 0$, where $B(\mathbf{r},\epsilon)$ is the ball of radius ϵ centered at $\mathbf{r}$. Subexponential decay occurs exactly when $\mathbf{Z} = \mathbf{1}$, corresponding to $\mathbf{r} = \nabla Q(\mathbf{1}) = \nabla\phi$, which we recognize to be the mean of the (untilted) distribution.

2.4. Proofs of criteria implying $\overline{\beta} < \beta_Q$. Theorem 2.8 and its corollary, Proposition 2.6, are proved in [**BP08**, Corollary 5.5]. The proof is via a direct construction of a homotopy between $\mathcal{C}$ and the local cycles in conclusion (i) of the theorem, the second conclusion and the proposition both following from the first conclusion. Of greater interest, however, is the Morse-theoretic argument which points the way to the proof. This argument does not appear in the published proof because it relies on the following conjecture, which has not been verified.

Conjecture 2.11 (compactification conjecture). There exists a compact space $\mathcal{V}^\dagger$ such that

(i) $\mathcal{V}$ embeds as a dense subset of $\mathcal{V}^\dagger$;
(ii) h extends to a continuous function mapping $\mathcal{V}^\dagger$ to the extended real line $[-\infty,\infty]$.

The conclusion of this conjecture is required as a hypothesis for the fundamental lemma of stratified Morse theory. It is also required that h be a Morse function, but Morse functions are generic, so this second requirement may be bypassed by taking a limit of Morse perturbations of h. Assuming this conjecture, Theorem 2.8 may be proved as follows.

Proof of Theorem 2.8 via Conjecture 2.11. The fundamental lemma of Morse theory states that the inclusion of $\mathcal{M}^b$ into $\mathcal{M}^a$ is a homotopy equivalence if there are no critical values in $[b,a]$. This homotopy equivalence deforms any chain in $\mathcal{M}^a$ into a homologous chain in $\mathcal{M}^b$. Thus we may lower the maximum height of $\mathcal{C}$ at least until a critical value for h is encountered. Let p be a critical point for h on $\mathcal{V}$ and let $c := h(p)$. Let D be a ball around p in $\mathcal{M}^c$ containing no other critical points, choose $\epsilon > 0$ small enough so that h has no critical values in $(c-\epsilon, c)$, and define the local space $\mathcal{M}^p_{\text{loc}}$ to be the topological pair $(\mathcal{M}^{c-\epsilon}\cup D, \mathcal{M}^{c-\epsilon})$. A further consequence of the fundamental lemma is that $H_d(\mathcal{M}^c, \mathcal{M}^{c-\epsilon})$ is a direct sum of the groups $H_d(\mathcal{M}^p_{\text{loc}})$. While the Morse-theoretic proof in [**GM88**] uses the gradient flow, these consequences are proved in [**BP08**] using a construction from [**ABG70**] involving hyperbolicity of Q at points on the boundary of $\mathsf{amoeba}(Q)$. In any case, the direct sum decomposition is exactly what is needed to finish the proof of Theorem 2.8. □

3. Smooth case

When $\mathsf{mincrit}$ is a single smooth point of $\mathcal{V}$ and a certain nondegeneracy assumption is satisfied, the form of the asymptotics for $\{a_{\mathbf{r}}\}$ is what is commonly called *Gaussian* or *Ornstein–Zernike*:

$$a_{\mathbf{r}} \sim C(\hat{\mathbf{r}})\,|\mathbf{r}|^{(1-d)/2}\,\mathbf{Z}_*^{-\mathbf{r}}\,. \tag{3.1}$$

The nondegeneracy assumption is that the Gaussian curvature of $\log \mathcal{V}$ not vanish, which is the same as the nonvanishing of the Hessian determinant in (3.3).

3.1. Formula in coordinates. The following theorem identifies the function $C(\hat{\mathbf{r}})$ and extends to finitely many critical points. The theorem was first proved in [**PW02**, Theorem 3.5], under the extra assumption that the only points of $\mathcal{V}$ in $\mathrm{ReLog}^{-1}(\mathbf{x}_*)$ are critical. Their proof gave an explicit deformation of the chain of integration to a sum $\sum_{\mathbf{Z}_* \in \mathsf{mincrit}} \mathcal{C}(\mathbf{Z}_*)$ of relative classes in $H_d(\mathcal{M}_{\mathrm{loc}}^{\mathbf{Z}_*})$, though they did not use this terminology. The extra assumption was removed in [**BP08**], via an existence proof which replaced the elementary deformations.

THEOREM 3.1. *Let $F = P/Q$ be a d-variate rational Laurent series corresponding to the component B of $\mathsf{amoeba}(Q)^c$, and let $\mathbf{x}_*$ be the minimizing point for $h = h_{\mathbf{r}_*}$. Suppose that $\mathsf{mincrit}(Q, \mathbf{r}_*)$ is nonempty and that at every $\mathbf{Z}_* \in \mathsf{mincrit}$, both the gradient of Q and the Gaussian curvature of $\log \mathcal{V}$ at $\log \mathbf{Z}_*$ are nonvanishing. Then there are relative homology classes $\mathcal{C}(\mathbf{Z}_*) \in H_d(\mathcal{M}_{\mathrm{loc}}^{\mathbf{Z}_*})$ such that*

$$a_{\mathbf{r}} = \sum_{\mathbf{Z}_* \in \mathsf{mincrit}} \int_{\mathcal{C}(\mathbf{Z}_*)} \omega \,, \tag{3.2}$$

where the equality is of equivalence classes up to $O(e^{(\beta_Q - \epsilon)|\mathbf{r}|})$, as in (2.1b). *At each point $\mathbf{Z}_* \in \mathsf{mincrit}$, if P is nonzero, the corresponding summand of* (3.2) *is given asymptotically by*

$$\int_{\mathcal{C}(\mathbf{Z}_*)} \omega \;\sim\; \Phi(\mathbf{Z}^*) := (2\pi r_d)^{(1-d)/2} (\mathbf{Z}_*)^{-\mathbf{r}} \mathcal{H}^{-1/2} \frac{P(\mathbf{Z}_*)}{(Z_d \, \partial Q/\partial Z_d)(\mathbf{Z}_*)} \,. \tag{3.3}$$

Here $\mathcal{H}$ is the determinant of the Hessian matrix of the parametrization of $\log \mathcal{V}$ by the coordinates $(z_1, \ldots, z_{d-1})$ near the point $\mathbf{z}_* := \log \mathbf{Z}_*$.

REMARK 3.2. The decomposition (3.2) holds whether or not the curvature (and hence $\mathcal{H}$) vanishes, or P vanishes. In fact, when $\mathcal{H}$ does not vanish, the corresponding summand of (3.2) may be expanded in an asymptotic series is descending powers of r_d:

$$\int_{\mathcal{C}(\mathbf{Z}_*)} \omega \;\sim\; \mathbf{Z}_*^{-\mathbf{r}} \sum_{n=0}^{\infty} b_n(\mathbf{Z}_*) r_d^{(1-d)/2-n} \,.$$

When $P(\mathbf{Z}_*)$ is also nonzero, then b_0 is nonzero so it is the leading term and agrees with (3.3):

$$b_0 = (2\pi)^{(1-d)/2} \mathcal{H}^{-1/2} \frac{P(\mathbf{Z}_*)}{(Z_d \, \partial Q/\partial Z_d)(\mathbf{Z}_*)} \,.$$

If $\mathcal{H}$ vanishes, an asymptotic expansion exists in decreasing fractional powers of r_d and possibly $\log r_d$; see [**Var77**].

A familiar example from [**PW02, PW08**] gives a concrete illustration of Theorem 3.1.

EXAMPLE 3.3. Let $F = 1/(1 - x - y - xy)$ be the generating function for the Delannoy numbers. This example is worked in [**PW08**, Section 4.2]. The variety $\mathcal{V}$ where $Q := 1 - x - y - xy = 0$ is smooth. In two variables, letting $\mathbf{r} := (r, s)$, the

critical point equations are

$$Q = 0\,,$$
$$sx\frac{\partial Q}{\partial x} - ry\frac{\partial Q}{\partial y} = 0\,.$$

Plugging this into a Gröbner basis package with $Q = 1 - x - y - xy$ yields precisely two solutions:

$$\mathbf{Z}_\pm := \left(\frac{\pm\sqrt{r^2+s^2} - s}{r}, \frac{\pm\sqrt{r^2+s^2} - r}{s} \right), \tag{3.4}$$

where the same sign choice is taken in both coordinates. The positive point is easily shown to be minimal (directly, or via Proposition 2.9), so $\mathbf{x}_* = \operatorname{ReLog} \mathbf{Z}_+$ and $\mathsf{mincrit}(\mathbf{r}_*) = \{\mathbf{Z}_+\}$. We have $\beta_Q = -\hat{\mathbf{r}} \cdot \operatorname{ReLog} \mathbf{Z}_+$ and

$$a_{\mathbf{r}} \sim (2\pi)^{-1/2} \sqrt{\frac{rs}{2\pi\sqrt{r^2+s^2}(r+s-\sqrt{r^2+s^2})^2}} \cdot \left(\frac{\sqrt{r^2+s^2}-r}{s} \right)^{-s} \left(\frac{\sqrt{r^2+s^2}-s}{r} \right)^{-r}.$$

Here the initial factor is computed by plugging in the values $\mathbf{Z}_\pm$ in (3.4) for $\mathbf{Z}_*$ on the right-hand side of (3.3), and simplifying. Nonvanishing of this quantity verifies the hypothesis of nonvanishing curvature of $\log \mathcal{V}$ at the point $\mathbf{z}_* := \log |\mathbf{Z}_*|$.

Proof of Proposition 2.5. At a smooth $\mathbf{Z}_* \in \mathsf{mincrit}$ we have an asymptotic series for $\int_{\mathcal{C}(\mathbf{Z}_*)} \omega$, either of the form (3.3) or the more general form in the subsequent remark. If there are more points in $\mathsf{mincrit}$, the corresponding summands in (3.2) will have different phases and therefore will not be able to cancel the contribution from $\mathbf{Z}_*$, except along a sublattice. □

Example 3.4 (local large deviations and CLT). This example shows why the asymptotics in Theorem 3.1 are known as Gaussian. Continuing Example 2.10, we suppose the random walk to be aperiodic. Theorem 3.1 gives the asymptotic value of $\mathbb{P}(S_n = \mathbf{v})$. Letting $\mathbf{r} := \mathbf{v}/n$, after solving for $\mathbf{Z}$ and $\mathbf{x}$ we obtain

$$\mathbb{P}(S_n = \mathbf{v}) \sim (2\pi)^{-d/2} K^{-1/2}(\mathbf{r})\, n^{-d/2} e^{-n\beta_Q(\mathbf{r})}\,,$$

where K is the determinant of the covariance matrix for the tilted distribution $p_{\mathbf{Z}}$. This estimate is uniform as long as $\mathbf{r}$ stays within a compact subset of the set of tilted means, which is just the interior of the convex hull of the support of p.

Let $\boldsymbol{\mu}$ be the mean of the distribution p. If $|\mathbf{v} - n\boldsymbol{\mu}| = o(n)$ then the determinants of the tilted covariance matrices are all $K_0 + o(1)$, where K_0 is the determinant of the untilted covariance matrix. In this regime, therefore, $\mathbb{P}(S_n = \mathbf{v})$ is proportional to $e^{-n\beta_Q(\mathbf{r})}$. The function β_Q reaches its maximum of zero at $\boldsymbol{\mu}$. Letting H denote the quadratic Taylor term, we then have

$$n\beta_Q\left(\frac{\mathbf{v}}{n}\right) = n\left[H\left(\frac{\mathbf{v}}{n} - \boldsymbol{\mu}\right) + O\left(\left| \frac{|\mathbf{v}|}{n} - \boldsymbol{\mu} \right| \right)^3 \right]$$
$$= \frac{H(\mathbf{v} - n\boldsymbol{\mu})}{n} + O\left(\frac{|\mathbf{v} - n\boldsymbol{\mu}|^3}{n^2} \right).$$

Therefore, as long as $|\mathbf{v} - n\boldsymbol{\mu}|$ is $o(n^{2/3})$, we have a uniform local central limit estimate

$$\mathbb{P}(S_n = \mathbf{v}) \sim C n^{-d/2} e^{-H(\mathbf{v} - n\boldsymbol{\mu})}\,, \tag{3.5}$$

where the quadratic Taylor term H is represented by the inverse of the covariance matrix and the normalizing constant C is given by $(2\pi)^{-d/2}K_0^{-1/2}$.

3.2. Coordinate-free formula. In Theorem 3.1, it is sufficient that any partial derivative of Q not vanish: in the formula (3.3) for Φ, z_d may then be replaced by z_j for any j such that $\partial Q/\partial Z_j \neq 0$. Although the explicit coordinate choice makes (3.3) useful for computing, this observation prompts us to rewrite the quantity in (3.3) in a more canonical way. The Hessian determinant looks like, and is, a curvature. To avoid discussing Gaussian curvature in any case beyond that of a real hypersurface, we give the coordinate-free formula only in a special case, arising in the study of quantum random walks. The important feature of this generating function is that $\log \mathcal{V}$ has a large intersection (co-dimension 1) with $\mathrm{i}\mathbb{R}^d$. More specifically, the following *torality* hypothesis is satisfied (see [**BBBP08**, Proposition 2.1]):

$$\mathbf{Z} \in \mathcal{V} \text{ and } |Z_1| = \cdots = |Z_{d-1}| = 1 \quad \Longrightarrow \quad |Z_d| = 1\,. \tag{3.6}$$

This next result is stated and proved in [**BBBP08**, Theorem 3.3].

THEOREM 3.5. *Suppose* $\mathbf{x}_* = \mathbf{0}$ *and* $\mathsf{mincrit}$ *is non-empty and that the torality hypothesis* (3.6) *is satisfied. Then, with* $|\cdot|$ *denoting the euclidean norm,*

$$a_{\mathbf{r}} = \pm \left(\frac{1}{2\pi\,|\mathbf{r}|}\right)^{(d-1)/2} \sum_{\mathbf{Z}\in\mathsf{mincrit}} \mathbf{Z}^{-\mathbf{r}} \frac{P(\mathbf{Z})}{|\nabla_{\log} Q(\mathbf{Z})|} \frac{1}{\sqrt{|\mathcal{K}(\mathbf{Z})|}}\, e^{-\mathrm{i}\pi\tau(\mathbf{Z})/4} + O\left(|\mathbf{r}|^{-d/2}\right).$$

Here, $\nabla_{\log} Q$ *is the logarithmic gradient* $(Z_1\,\partial Q/\partial Z_1, \ldots, Z_d\,\partial Q/\partial Z_d)$, $\mathcal{K}$ *is the Gaussian curvature, and* τ *is the difference between the numbers of positive and negative eigenvalues of the Hessian matrix. The estimate is uniform as* $\mathbf{r}_*$ *varies over compact subsets of the set of unit vectors* $\hat{\mathbf{r}}$ *in* $\mathbb{R}^d$ *for which* $\mathcal{K} \neq 0$ *and* $\mathsf{mincrit}(\hat{\mathbf{r}})$ *is nonempty.*

EXAMPLE 3.6. A *quantum random walk* (QRW) on $\mathbb{Z}^{d-1}$ with unitary coin U was defined in [**ADZ93**], where U is any matrix in the unitary group of rank $2(d-1)$. (See also [**ABN$^+$01, Kem03**].) Starting with a single particle at the origin, let $a(\mathbf{r}, n)$ denote the amplitude of finding the particle at position $\mathbf{r} \in \mathbb{Z}^{d-1}$ at time n (technically, one must also fix the starting and ending chiralities $(i, j) \in \{1, \ldots, 2(d-1)\}^2$, which will be assumed, but not explained). The spacetime generating function is

$$F(\mathbf{x}, y) := \sum_{\mathbf{r}\in\mathbb{Z}^{d-1},\, n\geqslant 0} a(\mathbf{r}, n)\, \mathbf{x}^{\mathbf{r}} y^n\,,$$

with $(\mathbf{x}, y) \in \mathbb{R}^d$. It is shown in [**BBBP08**, equation (2.2)] that $F = P/Q$ with

$$Q = \det\left(I - y\, M(\mathbf{x})\, U\right)\,,$$

where $M(\mathbf{x})$ is the diagonal matrix of order $2(d-1)$ with entries $x_1, x_1^{-1}, \ldots, x_{d-1}, {x_{d-1}}^{-1}$. The torality hypothesis (3.6) is verified as [**BBBP08**, Proposition 2.1].

The origin is on the common boundary of $\mathsf{amoeba}(Q)$ and a component B of $\mathsf{amoeba}(Q)^c$. For any $(\mathbf{r}_*, 1) \in \mathbb{R}^d$ not outwardly normal to a support hyperplane of B at $\mathbf{0}$, the origin is not a minimizing point, whence the amplitudes $a(\mathbf{r}, n)$ decay exponentially as $n \to \infty$ with $(\mathbf{r}, n)/n = (\mathbf{r}/n, 1) \to (\mathbf{r}_*, 1)$. When $(\mathbf{r}_*, 1)$ is inside the normal cone to B at $\mathbf{0}$, it may be verified in a number of cases that $\mathcal{V}$ is

smooth; see [**BBBP08**, Section 4] for several families of smooth QRW generating functions with $d-1=2$. It then follows from Proposition 2.5 and Theorem 2.8 that the amplitudes decay exponentially in direction $(\mathbf{r}_*, 1)$ if and only if $(\mathbf{r}_*, 1)$ is not the lognormal direction to any point $\mathbf{Z}_*$ of $\mathcal{V}_1 := \mathcal{V} \cap \mathrm{ReLog}^{-1}(\mathbf{0})$. Vectors $\mathbf{r}_* \in \mathbb{R}^{d-1}$ in the direction of which the decay is not exponential are said to be in the (asymptotically) *feasible region*. If there is no exponential decay and $\mathcal{K}$, the curvature of $\log \mathcal{V}_1$ at $\mathbf{z}_* := \log \mathbf{Z}_*$, is nonzero, the summands Φ in (3.3) are of order $n^{-(d-1)/2}$ with magnitudes proportional to the $-1/2$-power of the curvature. The curvature vanishes on a co-dimension 1 set including not only the boundary of the feasible region, but also certain interior curves.

Typically, there is more than one summand, with the phases of the summands related in a complicated way. This results in the Moiré patterns visible in Figure 2. Note that the set of directions $\mathbf{r}_* \in \mathbb{R}^{d-1} = \mathbb{R}^2$ of non-exponential decay is not always convex. On the other hand, the normal cone to B is the dual to the tangent cone, hence convex. We conclude that there are $(\mathbf{r}_*, 1)$ for which $\mathbf{0}$ is the minimizing point, but for which mincrit is empty.

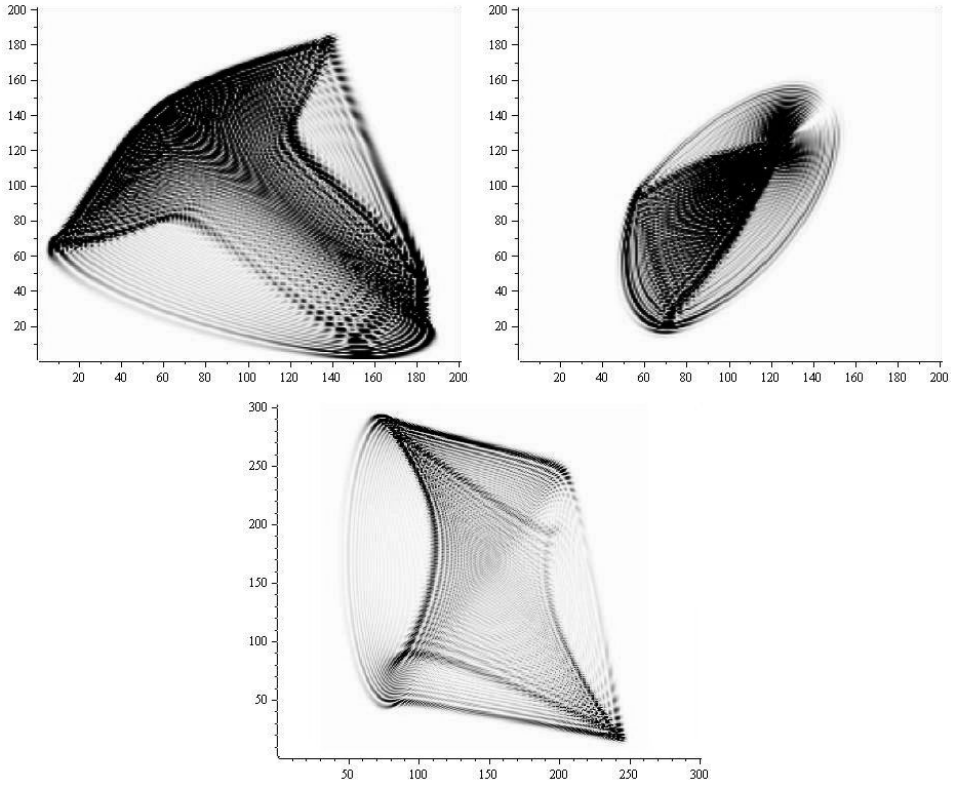


FIGURE 2. Intensity plot of squared amplitudes for three QRWs on $\mathbb{Z}^{d-1} = \mathbb{Z}^2$ at time $n = 200$; the asymptotic behavior is already clearly visible.

3.3. No minimal points. The statement and the original proof of Theorem 3.1 require the dominating point(s) $\mathbf{Z}_*$ to be minimal. In order to determine asymptotics in directions $\mathbf{r}_*$ for which $\mathsf{mincrit}(\mathbf{r}_*)$ is empty, it is useful to employ a residue form in place of the univariate residues employed in [**PW02**]; this difference is somewhat cosmetic, but still quite helpful.

Define the *residue form* of a meromorphic form $\omega := (P/Q)\,\mathrm{d}\mathbf{Z}$ on $\mathcal{V} := \{Q = 0\}$ by

$$\mathrm{Res}(\omega) := \iota^* \eta\,,$$

where η is any solution to

$$\mathrm{d}Q \wedge \eta = P\,\mathrm{d}\mathbf{Z}$$

and ι is the inclusion of $\mathcal{V}$ into $\mathbb{C}^d$. It is shown in [**DeV10**, Proposition 2.6] that such an η exists and that $\iota^*\eta$ is independent of the choice of η. The form $\mathrm{Res}(\omega)$ is also known in [**AGZV88**, Chapter 7] as the *Gel'fand–Leray* form of Q.

LEMMA 3.7 (Cauchy–Leray Residue Theorem). *Any transverse intersection of $\mathcal{V}$ with a homotopy between $\mathcal{C}$ and infinity has the same homology class $\psi \in H_d(\mathcal{V}, \mathcal{V}^\lambda)$ as long as λ is less than all critical values of h. For this class ψ,*

$$\int_{\mathcal{C}} \omega = 2\pi\mathrm{i} \int_{\psi} \mathrm{Res}(\omega)\,. \tag{3.7}$$

PROOF. See [**DeV10**, Theorem 2.8]). □

This result may be coupled with a formalization of the saddle point method that parallels the construction in the proof of Theorem 2.8.

LEMMA 3.8 (saddle point method). *Let $\mathcal{V}$ be a complex d-manifold and h a proper Morse function on $\mathcal{V}$ which is the real part of a complex analytic function. Let $\mathcal{C}$ be any d-cycle on $\mathcal{V}$. Then there is a unique critical value c and cycle $\mathcal{C}' \in \mathcal{V}^c$ such that:*

(i) *$\mathcal{C}$ projects to zero in $H_d(\mathcal{V}, \mathcal{V}^{c+\epsilon})$;*
(ii) *$[\mathcal{C}] = [\mathcal{C}']$ in any $H_d(\mathcal{V}, \mathcal{V}^{c-\epsilon})$;*
(iii) *$\mathcal{C}' = \sum_p \mathcal{C}_p$, where p runs over some subset of the critical points at height c and $\mathcal{C}_p$ is diffeomorphic to a d-ball.*

REMARKS. (1) We call the set of p appearing in the sum the *contributing critical points* and the height c the *minimax height*. (2) If h has isolated critical points then the Morse property is not actually needed: h is a limit of Morse perturbations h_ϵ and any weak limit of the resulting cycles $\mathcal{C}'_\epsilon$ satisfies (i) and (ii); the third conclusion must be generalized to allow a wedge of balls. (3) When $\mathcal{V}$ is not smooth but is smooth in a neighborhood of $\mathsf{mincrit}$, the same construction gives a relative cycle $\psi \in H_d(\mathcal{V}^c, \mathcal{V}^{c-\epsilon})$ where c is the common height of the points in $\mathsf{mincrit}$; this is good enough to produce asymptotics.

PROOF. The first two facts are quite general. If

$$\varnothing = X_0 \subseteq X_1 \subseteq X_2 \subseteq \cdots \subseteq X_n$$

is any filtration and the homology dimension of all spaces is d, then for any $\sigma \in H_d(X_n)$ there is a unique least j such that σ is homologous to a cycle supported on X_j, and if $j > 0$ then σ is nonzero in $H_d(X_j, X_{j-1})$. Choosing $X_j := \mathcal{V}^{c_j}$ for successive critical values c_j of h and applying the fundamental lemma of Morse theory proves (i) and (ii).

Being the real part of a complex analytic function, h has critical points of index d only. Ordinary Morse theory tells us that the homology of the pair $(\mathcal{V}^c, \mathcal{V}^{c-\epsilon})$ is a free abelian group generated by $(D_p, \partial D_p)$ where D_p is a d-ball centered at p and p runs through critical points of height c. Thus $\mathcal{C}'$ has a cycle representative that is a sum of arbitrarily small d-balls in $\mathcal{V}^c$ localized to critical points of height c, proving *(iii)*. □

In this way, we fulfill the promise, made at the end of Section 2.1, to find a limiting minimax chain which will no longer lie in $\mathcal{M}$. To each chain $\psi \in \mathcal{V}$ there corresponds a chain $\tilde{\psi} \in \mathcal{M}$ defined to be the product (in some local coordinates) of ψ with a small circle around the origin in the complementary complex one-space to $\mathcal{V}$. The saddle point deformation of ψ corresponds to the quasi-local representation $\tilde{\psi}$ of $\mathcal{C}$ in the proof of Theorem 2.8. The residue identity (3.7) in these coordinates is an obvious consequence of the ordinary residue theorem, the contribution from Morse theory being the existence of a cycle of the form $\tilde{\psi}$ in the homology class of $\mathcal{C}$. Putting together the previous two lemmas directly implies the following result.

THEOREM 3.9. *Suppose $\mathcal{V}$ is smooth and $h = h_{\mathbf{r}_*}$ is proper and Morse. Then there is a set Ξ of critical points for h on $\mathcal{V}$ at some common height c, along with topological $(d-1)$-balls $\mathcal{B}(\mathbf{Z}_*) \subseteq \mathcal{V}$ on which h is maximized at $\mathbf{Z}_*$, such that*

$$a_{\mathbf{r}} = \left(\frac{1}{2\pi \mathrm{i}}\right)^d \sum_{\mathbf{Z}_* \in \Xi} \int_{\mathcal{B}(\mathbf{Z}_*)} \omega$$

up to a difference of $O(e^{(c-\epsilon)|\mathbf{r}|})$ for some $\epsilon > 0$. Asymptotic series expansions for the summands in decreasing fractional powers of $|\mathbf{r}|$ are computable.

REMARKS. This shows that $a_{\mathbf{r}}$ may be estimated up to $O(e^{(c-\epsilon)|\mathbf{r}|})$ by a sum of terms $\sum_{\mathbf{Z}_* \in \Xi} \Phi(\mathbf{Z}_*)$, with Φ as defined by (3.3). This greatly generalizes Theorem 3.1 because it allows the contributing critical points to be non-minimal. When $d-1 \geqslant 2$ the computation of the expansions and even the leading term can be somewhat complicated. It is shown in [**Var77**] how to compute the leading exponent from the Newton diagram at the singularity. When $d-1=1$, the only degenerate possibilities are $h(\mathbf{Z}_* + u) \sim Cu^k$ for some $k \geqslant 3$, which lead to formulae analogous to (3.3) but with the exponent $(1-d)/2$ replaced by $-1/k$, the Hessian determinant replaced by the derivative of order k, and the power of 2π replaced by a value of the Gamma function.

EXAMPLE 3.10. The paper [**DeV10**] considers the generating function

$$F(X,Y) := 2X^2Y \frac{2X^5Y^2 - 3X^3Y + X + 2X^2Y - 1}{X^5Y^2 + 2X^2Y - 2X^3Y + 4Y + X - 2},$$

which is reverse-engineered so that its diagonal counts bi-colored supertrees. When $\mathbf{r}_*$ is the diagonal direction, there are precisely three critical points; none of these is minimal, but it is shown by direct homotopy methods that $\mathbf{Z}_* := (2, 1/8)$ is the unique contributing critical point. Near $\mathbf{Z}_*$, the behavior of h is quartic rather than quadratic, a double degeneracy coming from the merging of three distinct saddles that re-appear in any perturbation. Normally this would produce a factor of $|\mathbf{r}|^{-1/4}$ but the numerator vanishes to degree one at $\mathbf{Z}_*$ leading instead to a factor $n^{-5/4}$

and the asymptotic estimate

$$a_{n,n} \sim \frac{4^n}{8\Gamma(3/4)\, n^{5/4}}\,.$$

Theorem 3.9 and its generalizations as discussed in the subsequent remark reduce the problem of estimating coefficient asymptotics in the smooth case to identification of the contributing set Ξ. The good news is that even without this last step, the estimation problem is solved modulo a choice from among a finite set of possible estimates. The bad news is that we have no general method for determining Ξ. Example 3.10, for example, was handled by *ad hoc* methods. The one case where we appear to have an effective procedure for determining Ξ is in some work in progress on the case $d = 2$. Briefly, in this case one proceeds by computing the Morse complex as a finite multigraph. The generators for $H_1(\mathcal{V})$ given by this cell complex are precisely the one-dimensional saddle point contours descending from each critical point. Resolution of the cycle $\mathcal{C}$ in this basis is done by counting intersections with a dual basis, consisting of ascending contours from each critical point: the steepest ascent contours are replaced by polygonal approximations, whose intersection number with the special cycle $\mathcal{C}$ are computable from the combinatorics of these arcs, each of which connects a critical point to a pole ($x = 0$ or $y = 0$) of the height function.

4. Non-smooth case

Let $\mathbf{Z} \in \mathcal{V} = \{Q = 0\}$. After smooth points, the simplest local geometry $\mathcal{V}$ can have near $\mathbf{Z}$ is to be a union of smooth divisors intersecting transversally; such a critical point is called a *multiple point*. Multiple points consume the bulk of this section because this is the case in which the most is known.

4.1. Multiple points. There are a number of ways to formulate the definition of a multiple point, the most direct and geometric of which is modeled after [**PW04**, Definition 2.1].

Definition 4.1 (multiple point). The point $\mathbf{Z} \in \mathcal{V}$ is a multiple point if and only if there exist analytic functions $v_1, \ldots, v_n$ and ϕ defined on a neighborhood of $(Z_1, \ldots, Z_{d-1})$ in $\mathbb{C}^{d-1}$, and positive integers $k_1, \ldots, k_n$, such that

$$F(\mathbf{Z}') = \frac{\phi(\mathbf{Z}')}{\prod_{j=1}^{n} (1 - Z'_d\, v_j(Z'_1, \ldots, Z'_{d-1}))^{k_j}}\,, \tag{4.1}$$

the equality being one of meromorphic functions in a neighborhood of $\mathbf{Z}$, and such that

(i) $Z_d\, v_j(Z_1, \ldots, Z_d) = 1$ for all $1 \leqslant j \leqslant n$;
(ii) $Z'_d / v_j(Z'_1, \ldots, Z'_{d-1}) = 1$ for some j if and only if $\mathbf{Z}' \in \mathcal{V}$;
(iii) any set of at most d of the vectors $\{\nabla v_j(\mathbf{Z}) : 1 \leqslant j \leqslant n\}$ is linearly independent.

Remarks. (*i*) A smooth point is a multiple point. (*ii*) The reason we allow multiplicities ($k_j \geqslant 2$) but not tangencies ($\nabla v_i(\mathbf{Z}) \parallel \nabla v_j(\mathbf{Z})$) is to ensure genericity of intersections. In particular, $\bigcap_{j=1}^{n} \{\mathcal{V}_j\}$ will be a manifold and $\mathcal{V}$ will have a local product structure.

We let $\mathcal{V}_j := \{Z'_d\, v_j(Z'_1, \ldots, Z'_{d-1}) = 1\}$, so that $\mathcal{V}_1, \ldots, \mathcal{V}_n$ are locally smooth varieties parametrized by $Z'_d = u_j(Z'_1, \ldots, Z'_{d-1}) := 1/v_j(Z'_1, \ldots, Z'_{d-1})$. The geometric formulation is probably the most intuitive, but there is also an algebraic formulation that allows us to compute more effectively whether $\mathbf{Z}$ is a multiple point. To check (i) and (ii), we check whether Q factors completely in the local ring of germs of analytic functions at $\mathbf{Z}$. Supposing this to be true, we may write $Q(\mathbf{Z}') = \prod_{j=1}^n g_j^{k_j}$ where $g_j = Z'_d - u_j(Z'_1, \ldots, Z'_{d-1})$. If $n \geqslant d$, the transversality assumption is equivalent to the assertion that the dimension of $\mathcal{C}[\mathbf{Z}]/I$ as a complex vector space is one, where I is the ideal of the local ring generated by $u_1, \ldots, u_n$. If $n < d$, it is probably easiest to check linear independence directly. Apparatus for doing these computations may be found in [**CLO98**, Chapter 4]. The following example illustrates and underscores that being a multiple point is a local property.

Example 4.2. A generating function is given in [**PW04**, Example 1.3] that counts winning plays in a game with two time parameters and two coins with different biases. The denominator of the generating function is

$$Q := (1 - \frac{1}{3}X - \frac{2}{3}Y)(1 - \frac{2}{3}X - \frac{1}{3}Y)\,.$$

Here, Q is reducible in $\mathbb{C}[X, Y]$ as well as factoring locally, and the pole set $\mathcal{V}$ is the union of the two divisors, in this case two complex lines intersecting at the point $(1, 1)$. The lines intersect transversally, so this point is a multiple point.

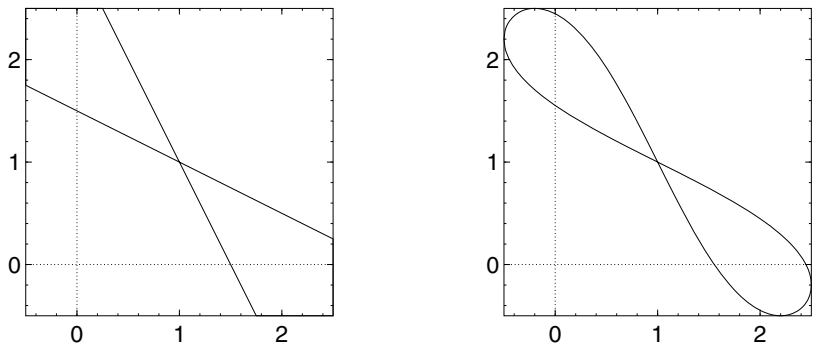


Figure 3. Zero sets of Q and $\tilde{Q}$.

Perturbing the polynomial gives a new polynomial

$$\tilde{Q} := 19 - 20X - 20Y + 5X^2 + 14XY + 5Y^2 - 2X^2Y - 2XY^2 + X^2Y^2\,,$$

the zero set of which, a rational curve, is drawn in [**PW04**, Example 3.2]. The real part of $\mathcal{V}$ looks like a lemniscate (a tilted figure eight) with a double point at $(1, 1)$. Although $\tilde{Q}$ is irreducible in the polynomial ring, it factors in the local ring, and the divisors intersect transversally.

When the denominator of F factors, even locally, a condition on the numerator is needed to ensure we are analyzing the correct denominator. Define the lognormal vectors to the divisors g_j at $\mathbf{Z}_*$ by

$$\mathfrak{n}_j := \left(Z_1 \frac{\partial v_j}{\partial Z_1}, \ldots, Z_{d-1} \frac{\partial v_j}{\partial Z_{d-1}}, Z_d \right) (\mathbf{Z}_*)\,.$$

DEFINITION 4.3 (partial fraction ideal). Let $Q = \phi \cdot \prod_{j=1}^n g_j^{k_j}$ locally near $\mathbf{Z}_*$. Say that a subset $S \subseteq \{1, \ldots, n\}$ is *small* if the vectors $\{\mathfrak{n}_j : j \in S\}$ span a proper subspace $V_S \subseteq \mathbb{R}^d$. Denote by $\mathcal{J} = \mathcal{J}(\mathbf{Z})$ the ideal in the local ring at $\mathbf{Z}$ generated by all products $\prod_{j=1}^n g_j^{\ell_j}$ such that there is a small set S with $\ell_j = k_j$ for $j \notin S$.

Equivalently, $P \in \mathcal{J}$ if and only if P/Q has a partial fraction expansion $\sum_S F_S$ in which the denominator of F_S is a product over $j \in S$ of powers of g_j and S is small. The relevance of this is that $\mathbf{x}_* := \text{ReLog}(\mathbf{Z}_*)$ is not a minimizing point for F_S unless $\mathbf{r}_* \in V_S$. If we let G denote the union over small S of the proper subspaces V_S, we see that if $P \in \mathcal{J}$, then $|\mathbf{Z}_*|^{\mathbf{r}} a_{\mathbf{r}}$ is exponentially small when $|\mathbf{r}| \to \infty$ with $\mathbf{r}_*$ in a compact subset of G^c.

EXAMPLE 4.4. Let $F = P/Q$ and $\tilde{F} = P/\tilde{Q}$ with $P = X - Y$ and $Q, \tilde{Q}$ as in Example 4.2. In the former case, there is a global partial fraction expansion

$$F = \frac{3}{Q_1} - \frac{3}{Q_2} := \frac{3}{1 - (2/3)X - (1/3)Y} - \frac{3}{1 - (2/3)Y - (1/3)X}\,.$$

Clearly, $\overline{\beta}_Q = \max \overline{\beta}_{Q_1} - \overline{\beta}_{Q_2}$, and when $\mathbf{r}_*$ is the diagonal, for instance, we have $\overline{\beta} < 0 = \beta_Q$. In the second case, near $\mathbf{Z}_*$ there is a partial fraction expansion into a sum of meromorphic functions

$$F = \frac{P_1}{g_1} + \frac{P_2}{g_2}\,,$$

and again $\overline{\beta}(\mathbf{r}_*) < 0 = \beta_Q(\mathbf{r}_*)$. It is worth noting that these exceptions are in some sense trivialities, which will be prevented by the hypothesis $P \notin \mathcal{J}$.

4.2. Minimal multiple points and a piecewise polynomial formula when $n \geqslant d$. Suppose that the multiple point $\mathbf{Z}_*$ is minimal, meaning that $\mathbf{x}_* := \text{ReLog}(\mathbf{Z}_*) \in \partial B$. Characterization of the dual cone is reasonably explicit in this case. The hyperplane normal to each lognormal vector $\mathfrak{n}_j$ at $\mathbf{x}_*$ is a support hyperplane to B at $\mathbf{x}_*$, so there is a well defined outward direction (away from B); choosing outward normals and taking the positive gives the cone $\mathbf{N}_*(\mathbf{Z}_*)$. In Theorem 4.7, a formula will be given for the contribution to the Cauchy integral from the critical point $\mathbf{Z}_*$ when $\hat{\mathbf{r}} \to \mathbf{r}_* \in \mathbf{N}_*(\mathbf{Z}_*)$. The contribution is denoted by $\Phi(\mathbf{r})$ in analogy to Theorem 3.1. When $P \notin \mathcal{J}(\mathbf{Z})$, the quantity $|\Phi(\mathbf{r})\mathbf{Z}_*^{-\mathbf{r}}|$ does not decay exponentially, again in analogy with Theorem 3.1. The existence of such a formula implies that Proposition 2.5 holds with "smooth point" replaced by "minimal point". We then have the following analogue of Proposition 2.9.

PROPOSITION 4.5 (combinatorial case for multiple points). *Supposing $a_{\mathbf{r}} \geqslant 0$ for all $\mathbf{r}$, let $\mathbf{r}_*$ be any vector with unique minimizer $\mathbf{x}_*$. If all critical points are multiple points then $\overline{\beta}(\mathbf{r}_*) = \beta_Q(\mathbf{r}_*)$, provided that $P \notin \mathcal{J}(\exp(\mathbf{x}_*))$.*

PROOF. Fix $\mathbf{r}_*$. As in Proposition 2.9, the point $\mathbf{Z}_* := \exp(\mathbf{x}_*)$ is minimal. Locally, $\mathsf{amoeba}(Q)$ is the union of the $\text{ReLog}[\mathcal{V}_j]$, therefore the span of all the lognormal vectors is equal to the dual cone to $\mathbf{K}(\mathbf{r}_*)$, hence contains $\mathbf{r}_*$. It follows from formula (4.3) in Theorem 4.7 below that $\overline{\beta}(\mathbf{r}_*) = \beta_Q(\mathbf{r}_*)$. □

We now turn to asymptotic formulae for the contribution to $a_{\mathbf{r}}$ from a minimal multiple point $\mathbf{Z}_*$. As one might expect from part (*iii*) of the definition of multiple points, there are two somewhat different cases, depending on whether n or d is greater. We consider the case $n \geqslant d$ in somewhat greater detail because the formulae

are nicer. Along with the assumptions on multiple points, the inequality $n \geqslant d$ implies that $\{\mathbf{Z}_*\}$ is a zero-dimensional stratum and that the cone $\mathbf{N}_*$ has nonempty interior. The surprising fact here is that for $\hat{\mathbf{r}} \to \mathbf{r}_* \in \mathbf{N}_*(\mathbf{Z}_*)$, the quantity $a_{\mathbf{r}}$ is nearly piecewise polynomial. This was observed already in [**Pem00**, Theorem 3.1], where the following qualitative description was given.

THEOREM 4.6. *Let $F = P/Q$ and let $\mathbf{Z}_*$ be a multiple point with $n \geqslant d$. There is a finite vector space $\mathcal{W}$ of polynomials and an algebraic set G of positive codimension in $\mathbb{R}^d$ such that*

$$a_{\mathbf{r}} = \mathbf{Z}^{-\mathbf{r}}(\eta(\mathbf{r}) + E(\mathbf{r}))\,,$$

where $\eta \in \mathcal{W}$ depends on P and E decays exponentially in $\mathbf{r}$, uniformly as $\mathbf{r}$ varies over compact connected subsets of $\mathbf{N}_* \setminus G$.

The set G is identified there as the same set $G = \bigcup_S V_S$ discussed subsequently to the definition of the partial fraction ideal. The complement of G may be disconnected, in which case η may be equal to different polynomials on the different components, and is best described as a (piecewise) spline. As long as η is not the zero polynomial, it gives the asymptotic behavior of $a_{\mathbf{r}}$ on that component. As shown in Theorem 4.7 below, $\eta = 0$ if and only if $P \in \mathcal{J}(\mathbf{Z}_*)$, the "if" direction following already from [**Pem00**]. The proof of Theorem 4.6 given in [**Pem00**] is based on finite-dimensionality of the space of shifts of the array $\{a_{\mathbf{r}}\}$ modulo exponentially decaying terms, and is not constructive. This leaves the question of an explicit description of η. This is given by the following result, which combines the published result [**PW04**, Theorem 3.6] with the equivalent but more canonical formula given in [**BP04**]. The formula is visually simplest when $\mathbf{Z}_* = (1, \ldots, 1)$, so this is given first as a special case.

THEOREM 4.7. *Let $F = P/Q$, let B be a component of $\mathsf{amoeba}(Q)^c$, and let $\mathbf{r}_*$ have unique minimal point $\mathbf{x}_*$.*

(i) *Suppose that $\mathsf{mincrit}(\mathbf{r}_*) = \{\mathbf{Z}_*\}$ where $\mathbf{Z}_*$ is a multiple point with local factorization $f := F \circ \exp = \phi / \prod_{j=1}^n g_j^{k_j}$. Let $D := \sum_{j=1}^n k_j$. If $P \notin \mathcal{J}(\mathbf{Z}_*)$ then*

$$a_{\mathbf{r}} \sim \mathbf{Z}_*^{-\mathbf{r}}\, \phi(\mathbf{Z}_*)\, \theta(\mathbf{r})\,, \tag{4.2}$$

where $\theta(\mathbf{r}_*) = \theta_Z(\mathbf{r}_*)$ is the density of the image of Lebesgue measure under the map $\Psi\colon (\mathbb{R}^+)^D \to \mathbf{N}_*$ which takes k_j of the standard basis vectors to the lognormal vector $\mathfrak{n}_j$.

(ii) *If $\mathsf{mincrit}(\mathbf{r}_*)$ is the union of a set Ξ of multiple points with a common value of D and a set of points $\mathbf{Z}$ for which $F(\mathbf{Z}+\mathbf{Z}') = O(|\mathbf{Z}'|^{1-D})$, then*

$$a_{\mathbf{r}} = \sum_{\mathbf{Z}\in\Xi} \mathbf{Z}^{-\mathbf{r}}\, \phi(\mathbf{Z}) \theta_{\mathbf{Z}}(\mathbf{r}) + O(|\mathbf{r}|^{D-1-d})\,, \tag{4.3}$$

where $\theta_{\mathbf{Z}}$ is the polynomial appearing on the right-hand side of (4.2), defined in terms of the local product decomposition at $\mathbf{Z}$.

The formula (4.3) gives an asymptotic estimate for $a_{\mathbf{r}}$ (possibly on a sublattice, in case of periodicity), as long as $P \notin \mathcal{J}(\mathbf{Z})$ for some $\mathbf{Z} \in \Xi$.

SKETCH OF PROOF. Let $\mathcal{B}$ be the collection of sub-multisets of $\{1, \ldots, n\}$ for which the corresponding subset of $\{\mathfrak{n}_j : 1 \leqslant j \leqslant n\}$ is independent and each $\mathfrak{n}_j$ a

number of times $\ell_j \leqslant k_j$. A local partial fraction decomposition allows us to write f as

$$\sum_{S \in \mathcal{B}} \frac{\phi_S}{\prod g_j^{\ell_j}}\,, \tag{4.4}$$

where $\ell_j = 0$ if $j \notin S$. For each summand, integrating by parts $D - d$ times gives

$$\int e^{-\mathbf{r}\cdot\mathbf{z}} \frac{\phi_S(\mathbf{z})}{\prod g_j^{\ell_j - 1}}\,\mathrm{d}\mathbf{z} = \left[\frac{\partial}{\partial r_1}^{\ell_1} \cdots \frac{\partial}{\partial r_d}^{\ell_d}\right] \int e^{-\mathbf{r}\cdot\mathbf{z}} \frac{\phi_S(\mathbf{z})}{\prod_{j\in S} g_j}\,\mathrm{d}\mathbf{z}\,. \tag{4.5}$$

Having reduced to the case $n = d$, a number of methods suffice for us to show that the last integral is asymptotic to $e^{-\mathbf{r}\cdot\mathbf{z}} J^{-1}$ when $\mathbf{r}$ is in the positive hull of $\{\mathfrak{n}_j : j \in S\}$ and J is the determinant of $\{\mathfrak{n}_j : j \in S\}$; see for example [**BP04**, Section 5] or the original proof of Theorem 3.3 in [**PW04**]. Evaluating derivatives shows that (4.5) is asymptotic to

$$e^{-\mathbf{r}\cdot\mathbf{z}} \mathbf{r}^{\boldsymbol{\ell}-\mathbf{1}} J^{-1}\,.$$

Summing over $S \in \mathcal{B}$ gives an expression for (4.4). It is not hard to show by induction that this expression is equal to the spline described on the right-hand side of (4.2).

To prove (4.3), one localizes to neighborhoods of each $\mathbf{Z}_* \in \mathsf{mincrit}$; this is carried out in [**PW04**, Section 4] in the case when $\mathcal{V} \cap \mathrm{ReLog}^{-1}(\mathbf{x}_*)$ contains only critical points; the general localization may be found in [**BP08**, Section 5]. If every $\mathbf{Z}_* \in \mathsf{mincrit}$ is a multiple point, the result now follows by summing (4.2) over $\mathbf{Z}_*$. If there are critical points of another type, then an upper estimate is required on their contribution. This may be found in [**BP08**, Lemma 5.9]. □

Example 4.8 (Example 4.2, continued). Let

$$F = \frac{P}{Q} = \frac{1}{(1 - \frac{1}{3}X - \frac{2}{3}Y)(1 - \frac{2}{3}X - \frac{1}{3}Y)}\,.$$

The variety $\mathcal{V}$ is the union of two complex lines, meeting at the single point $(1, 1)$. This point is a zero-dimensional stratum; each punctured complex line is a stratum as well. The boundary of B is the union of the two smooth curves $e^x + 2e^y = 3$ and $2e^x + e^y = 3$, intersecting at $(0, 0)$ (see Figure 4; cf. Figure 1(b)).

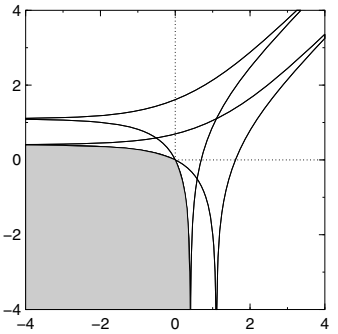


Figure 4. B is the unbounded shaded region.

When $r/s \in [1/2, 2]$, the minimizing point is at the origin. Because $a_{rs} \geqslant 0$, we know that the positive point $\mathbf{Z}_* = (1,1) = \exp(\mathbf{x}_*)$ is in $\mathsf{mincrit}$, and it is easy to verify that $\mathsf{mincrit}(r,s) = \{\mathbf{Z}_*\}$ for $1/2 \leqslant r/s \leqslant 2$. Plugging into Theorem 4.7 and evaluating the Hessian determinant we find that θ is a constant because $n = d$, and that $a_{rs} = 1/9 + o(1)$. When $r/s \notin [1/2, 2]$, we have $\mathbf{x}_* \neq (0,0)$. Here, $\beta_Q < 0$ and a_{rs} decays exponentially; all points other than $(1,1)$ are smooth, so we may use Theorem 3.1 to evaluate $a_{\mathbf{r}}$ asymptotically.

4.3. Non-minimal points and the case $n < d$. Suppose $\mathbf{Z}_*$ is a multiple point and that f is factored locally as $\phi / \prod_{j=1}^n g_j^{k_j}$. The point $\mathbf{Z}_*$ lies in some stratum of $\mathcal{V}$, which, under the transversality assumptions, can be taken to be $\mathcal{V}_0 := \bigcap_{j=1}^n \mathcal{V}_j$. Let δ denote the dimension of $\mathcal{V}_0$. In the next paragraph we justify writing the Cauchy integral (1.2) as an iterated integral

$$\Phi_{\mathbf{Z}_*}(\mathbf{r}) := \left(\frac{1}{2\pi \mathrm{i}}\right)^d \int_B \left[\int_\alpha \mathbf{Z}^{-\mathbf{r}-\mathbf{1}} F(\mathbf{Z})\, \mathrm{d}\mathbf{Z}_\perp\right] \mathrm{d}\mathbf{Z}_\parallel\,, \tag{4.6}$$

where B is a ball of dimension δ in $\mathcal{V}_0$ on which h is strictly maximized at $\mathbf{Z}_*$, and α is some cycle supported on a $(d-\delta)$-dimensional complex plane through $\mathbf{Z}_*$ which is transverse to $\mathcal{V}_0$. This justification pulls in some Morse-theoretic facts which may be skipped on first reading.

Thom's Isotopy Lemma ([**Mat70**], quoted in [**GM88**, Section I.1.4])) asserts that the pair $(\mathbb{C}^d, \mathcal{V})$ is locally a product of $\mathcal{V}_0$ with a pair $(\mathbb{C}^{d-\delta}, X)$, where X is the intersection of $\mathcal{V}$ with a complex space of dimension $d-\delta$ through $\mathbf{Z}_*$ which is transverse to $\mathcal{V}_0$. Recall from the end of Section 2 the decomposition of the chain of integration $\mathcal{C}$ into relative homology classes in $H_d(\mathcal{M}_{\mathrm{loc}}^{\mathbf{Z}_*})$. In general, representability of $\mathcal{C}$ as the sum of these relies on Conjecture 2.11, but in the cases considered in this section (in which $\mathbf{Z}_*$ is minimal or all factors g_j are linear) this is known. In any case, these local homology groups are identified in [**GM88**] as isomorphic to $H_{d-\delta}(\mathbb{C}^{d-\delta}, \mathbb{C}^{d-\delta} \setminus X)$, the latter pair being topologically equivalent to the cone $\mathbb{C}^{d-\delta} \setminus X$. The isomorphism takes a $(d-\delta)$-cycle α to $\alpha \times (B_\delta, \partial B_\delta)$. Here B_δ is a ball of dimension δ about $\mathbf{Z}_*$ in $\mathcal{V}_0$; the tangent space to the complex d-manifold $\mathcal{V}_0$ has a real d-dimensional subspace along which h is maximized at $\mathbf{Z}_*$; and taking B_δ tangent to this assures that $h < h(\mathbf{Z}_*) - \epsilon$ on ∂B_δ.

In general, when $\mathbf{Z}_*$ is not minimal, we have at present no way of knowing for which points $\mathbf{Z}_*$ the integral (4.6) contributes to $\mathcal{C}$ with nonzero coefficient, hence contributing to the estimate for $a_{\mathbf{r}}$. In the special case where $F = P/Q$ and Q is the product of linear factors, a combinatorial method is given in [**BP04**] for writing the chain of integration in the Cauchy integral as a sum of cycles $\alpha \times \beta$, where β is a global cycle that projects to $(B_\delta, \partial B_\delta)$ in $\mathcal{M}^{h(\mathbf{Z}_*)-\epsilon}$. The other well understood case is when $\mathbf{Z}_*$ is minimal. In this case, the local cycle contributes to the Cauchy integral whenever $\mathbf{x}_*$ is a minimizing point for $\mathbf{r}_*$. In either of these two cases, an asymptotic estimate for $a_{\mathbf{r}}$ is obtained by summing (4.6) over those $\mathbf{Z}_*$ known to contribute to the Cauchy integral. The following theorem then rests on evaluation of (4.6). To make sense of the final result, note that $\{\mathfrak{n}_j\}$ span the complex linear space normal to the tangent space to $\mathcal{V}_0$ at $\mathbf{Z}_*$, and that the definition of θ in Theorem 4.7 may be extended by taking it to be the density of $\Psi(\mathrm{d}\mathbf{x})$ with respect to $(d-\delta)$-dimensional Lebesgue measure on the normal space.

THEOREM 4.9. *Let $F = P/Q = \phi \,/\, \prod_j g_j^{\,j}$ near a multiple point $\mathbf{Z}_*$. Let $y_1, \ldots, y_d$ be unitary local coordinates in which $\mathcal{V}_0$ is the set $y_{\delta+1} = \cdots = y_d = 0$, and for $\mathbf{Z} \in \mathcal{V}_0$, define $\Lambda(\mathbf{Z})$ to be the Hessian determinant of the function $h_{\mathbf{r}_*}$ restricted to $\mathcal{V}_0$:*

$$\Lambda(\mathbf{Z}) := \left| \frac{\partial^2 h_{\mathbf{r}_*}}{\partial y_i \partial y_j} \right|_{1 \leqslant i,j \leqslant \delta} .$$

If $\phi(\mathbf{Z}_*)$ and $\Lambda(\mathbf{Z}_*)$ are nonvanishing then the quantity $\Phi(\mathbf{r})$ in (4.6) may be evaluated asymptotically to yield

$$\Phi(\mathbf{r}) \sim (2\pi)^{-\delta/2}\, \Lambda(\mathbf{Z}_*)^{-1/2}\, \mathbf{Z}_*^{-\mathbf{r}}\, \phi(\mathbf{Z}_*)\, \theta(\mathbf{r}) .$$

PROOF. The inner integral is of the form evaluated in Theorem 4.7. At any point $\mathbf{Z} \in \mathcal{V}_0$, the lognormal vectors $\{\mathbf{n}_j\}$ defined there span the complex linear space orthogonal to the tangent space to $\mathcal{V}_0$. Equation 4.2 therefore evaluates the inner integral asymptotically as $\mathbf{Z}_*^{-\mathbf{r}} \phi(\mathbf{Z}_*) \theta(\mathbf{r})$. The function $A(\mathbf{Z}) := \theta(\mathbf{r})\phi(\mathbf{Z})$ does not vanish at $\mathbf{Z}_*$, hence the outer integral is a saddle point integral of the type $\int_{\mathcal{B}} \mathbf{Z}^{-\mathbf{r}} A(\mathbf{Z})\, \mathrm{d}\mathbf{Z}$. Standard integrating techniques (see, e.g., [**PW09**, Theorem 2.3]) then show that the integral is asymptotic to $(2\pi)^{-\delta/2}\, A(\mathbf{r})\, \mathbf{Z}_*^{-\mathbf{r}} \Lambda(\mathbf{Z}_*)^{-1/2}$, as desired. □

4.4. More complicated geometries. When $\mathbf{Z}_*$ is not a smooth point and is not in a stratum which is locally an intersection of normally crossing smooth divisors, residue theory alone does not suffice to estimate $a_{\mathbf{r}}$. The analysis of this case is the most intricate and recent. The rather long preprint [**BP08**] is devoted to a subcase of this in which $\mathcal{V}$ is locally a quadratic, with the same geometry as the cone $\{z^2 - x^2 - y^2 = 0\}$. This section will be limited to a quick sketch of the derivation of results in this case, beginning with an example of such a generating function.

EXAMPLE 4.10. The generating function for *creation rates* in domino tilings of the Aztec Diamond is given by

$$F(x, y, z) = \frac{1}{1 - (x + x^{-1} + y + y^{-1})\dfrac{z}{2} + z^2} . \tag{4.7}$$

The precise meaning of creation rates is not important here; they are differences of the quantities of primary interest, namely the *placement probabilities* that tell the likelihood of the square (i, j) in the order-n diamond being covered by a domino in a particular orientation. (The placement probability generating function is similar to the creation rate generating function but has an extra factor in the denominator.) It is not hard to verify that Q, the denominator of F, has a singularity at $(1, 1, 1)$ which is geometrically a cone. The tangent cone to Q at $(1, 1, 1)$ is $z^2 - (1/2)(x^2 + y^2)$, which is also the tangent cone to the complement of $\mathsf{amoeba}(Q)$ at $(0, 0, 0)$, and its dual is the cone $\mathbf{N}_* := \{(r, s, t) : t \geqslant 0, t^2 \geqslant 2(r^2 + s^2)\}$.

We consider asymptotics whose dominating point is some fixed $\mathbf{Z}_*$ at which $\mathcal{V}$ has nontrivial local geometry. We assume that $\mathbf{r}_* \in \mathbf{N}_*$, the dual to the tangent cone to B at $\mathbf{x}_* := \operatorname{ReLog} \mathbf{Z}_*$ (in the absence of this, $e^{\mathbf{r}\cdot\mathbf{x}} a_{\mathbf{r}}$ will be exponentially small). The Cauchy integral in logarithmic coordinates (2.2) is a generalized Fourier transform of $f = F \circ \exp$, which has a singularity of the same type. What we find in the end is that $a_{\mathbf{r}}$ is estimated by $\hat{f}(\mathbf{r})$, the generalized Fourier transform of f evaluated at $\mathbf{r}$. The Fourier transform of f is estimated asymptotically by the

Fourier transform of its leading homogeneous term, which is the cone $\alpha(x,y,z) := z^2 - (1/2)(x^2+y^2)$. The Fourier transform of a homogeneous quadratic is the dual quadratic $\alpha^*(r,s,t)$ which in this case is given by $t^2 - 2(r^2+s^2)$. Thus we have, for example, the following result.

THEOREM 4.11 ([**BP08**, Theorem 4.2]). *Let $\{a_{\mathbf{r}}\}$ be the coefficients of the series for the Aztec Diamond creation generating function* (4.7)*, with $\mathbf{r} := (r,s,t)$. Then*

$$a_{rst} \sim \frac{4}{\pi}\left(t^2 - 2r^2 - 2s^2\right)^{-1/2},$$

uniformly as $\mathbf{r} \to \infty$ and $\hat{\mathbf{r}}$ varies over compact subsets of the set $\{r^2+s^2 < t^2/2\}$.

SKETCH OF PROOF. Begin with the Cauchy integral (2.2) in logarithmic coordinates. The function $f = F \circ \exp$ is meromorphic. Expand it as p/q and let $q = \alpha \cdot (1+\rho)$, where α is the homogeneous quadratic $z^2 - (1/2)(x^2+y^2)$ and $\rho(\mathbf{x}) = O(|\mathbf{x}|^3)$. We may then expand $1/q$ as a series

$$\frac{1}{q} = \frac{1}{\alpha}\left(1 - \frac{\rho}{\alpha} + \frac{\rho^2}{\alpha^2} + \cdots\right),$$

where $(\rho/\alpha)^j = O(|\mathbf{x}|^j)$ on compact cones avoiding the cone $\alpha = 0$. Expanding the numerator p into monomials, we arrive at a double series

$$f = \sum_{\mathbf{m},j} \frac{\mathbf{x}^{\mathbf{m}}}{\alpha^{1+j}}.$$

Provided the remainders obey the necessary estimate, we may integrate this term by term; the theorem requires only the leading integral and remainder estimate.

The integral

$$\int_{\mathbf{x}+\mathrm{i}\mathbb{R}^d} e^{-\lambda(\mathbf{r}\cdot\mathbf{z})}\frac{1}{\alpha^s}\,\mathrm{d}\mathbf{z}$$

is studied in [**ABG70**]. Although this integral is not convergent, it defines a *distribution* or *generalized function* in the sense of [**GS64**]. As a distribution, this is precisely what M. Riesz [**Rie49**] identified as the generalized function $C(\alpha^*)^{s-d/2}$, where the constant C is the product of values of the Gamma function, slightly misquoted in [**ABG70**, Equation (4.20)]. To get from the generalized function to an estimate for the particular integral (2.2), and to replace the limits of the integral as well by $\mathbf{x}+\mathrm{i}\mathbb{R}^d$, one requires a deformation of $\mathbf{x}+\mathrm{i}\mathbb{R}^d$ to a set where $\mathbf{r}\cdot\mathbf{x}$ is uniformly positive, i.e., $\mathbf{r}\cdot\mathbf{x} \geqslant \epsilon\,|\mathbf{x}|$. This is another Morse-theoretic result, proved in [**ABG70**] for hyperbolic linear differential operators and adapted in [**BP08**, Theorem 5.8]. □

5. Summary, a counterexample, and a conjecture

On the subject of the exponential growth rate, we have seen that $\overline{\beta}(\mathbf{r}_*)$ is at most $\beta_Q(\mathbf{r}_*) := -\mathbf{r}_*\cdot\mathbf{x}_*$ where $\mathbf{x}_*$ is the minimizing point. A list of known reasons why $\overline{\beta}$ (or β) can be strictly less than β_Q is as follows.

(i) Periodicity. For example, if $\beta(\mathbf{r}_*)$ exists for F and is finite, then β will fail to exist for $F(\mathbf{Z}) + F(-\mathbf{Z})$, the coefficients of which are twice those of F when $\mathbf{r}$ is even, and zero when $\mathbf{r}$ is odd.
(ii) mincrit is empty. We saw in Theorem 2.8 that this implies $\overline{\beta} < \beta_Q$.
(iii) mincrit is nonempty, but $P(\mathbf{Z}_*)$ vanishes on mincrit.

(iv) $\mathsf{mincrit}$ is nonempty, but at every $\mathbf{Z}_* \in \mathsf{mincrit}$ on which $P(\mathbf{Z}_*)$ is nonvanishing, the normal cone $\mathbf{N}_*$ fails to contain $\mathbf{r}_*$.

However, if there is no topological obstruction to moving the chain of integration further down, then the minimax height c is less than $-\mathbf{r}_* \cdot \mathbf{x}_*$ and is in fact equal to $-\mathbf{r}_* \cdot \mathbf{x}'$, where $\mathbf{x}' = \operatorname{ReLog} \mathbf{Z}'$ and $\mathbf{Z}'$ is a contributing critical point. If there is a contributing critical point of height c for which $P \notin \mathcal{J}$, then we define $\hat{\beta}_Q(\mathbf{r}) = c$. Otherwise, the partial fraction expansion tells us that the integrand is the sum of pieces whose domains of holomorphy are greater than that of the original integrand. By treating each separately, the chain of integration may be pushed below c to some new minimax height. We continue in this manner until we reach a point $\mathbf{Z}_*$ for which $\mathcal{C}$ has nonzero projection in $H_d(\mathcal{M}_{\text{loc}}^{\mathbf{Z}_*})$ and $P \notin \mathcal{J}(\mathbf{Z}_*)$. In this way, we may define our best guess for the exponential rate as

$$\hat{\beta}_Q(\mathbf{r}_*) := -\mathbf{r}_* \cdot \mathbf{x}_*$$

where $\mathbf{x}_* = \operatorname{ReLog} \mathbf{Z}_*$ for this critical point $\mathbf{Z}_*$ which maximizes height among all contributing critical points with $P \notin \mathcal{J}(\mathbf{Z}_*)$. A reasonable conjecture would seem to be that $\overline{\beta}(\mathbf{r}_*) = \hat{\beta}_Q$, but there is a counterexample (not yet published) as follows. The trivariate generating function $F = P/Q$ for the so-called *fortress* variant of the Aztec Diamond tiling model has a non-smooth point at $(1,1,1)$. The normal cone $\mathbf{N}_*$ dual to $\mathbf{K}(1,1,1)$ may be divided into an inner region A and an outer region B such that for $\mathbf{r}_* \in B$, asymptotics are not exponential, whereas for $\mathbf{r}_* \in A$, asymptotics decay exponentially. For $\mathbf{r}_* \in B$, the magnitude of $a_{\mathbf{r}}$ does not decay exponentially as $|\mathbf{r}| \to \infty$ with $\hat{\mathbf{r}} \to \mathbf{r}_*$. By the contrapositive of Theorem 2.8, since $\mathsf{mincrit}(\mathbf{r}_*) = \{(1,1,1)\}$, we conclude that $\mathbf{r}_* \in \mathbf{N}_*(\mathbf{r}_*)$ for $\mathbf{r}_* \in B$. It follows that this holds for $\mathbf{r}_* \in A$ as well. Exponential decay for $\mathbf{r}_* \in A$ then shoots down the conjecture. We are left with the following weakened version of the conjecture for minimal points, based on the belief that the fortress counterexample is the only kind that can arise and that it must happen on the interior of $\mathbf{N}_*(\mathbf{r}_*)$.

CONJECTURE 5.1 (weak sharpness of $\hat{\beta}_Q$). Suppose $\mathbf{Z}_*$ is a minimal critical point for some $\mathbf{r}_*$. Suppose that $P \notin \mathcal{J}(\mathbf{Z}_*)$, and let $\mathbf{x}_* := \operatorname{ReLog}(\mathbf{Z}_*)$. Then $\overline{\beta}(\mathbf{r}') = -\mathbf{r}' \cdot \mathbf{x}_*$ for $\mathbf{r}'$ in some set A whose convex hull is $\mathbf{N}_*(\mathbf{Z}_*)$.

References

[ABG70] M. Atiyah, R. Bott, and L. Gårding, *Lacunas for hyperbolic differential operators with constant coefficients, I*, Acta Math. **124** (1970), 109–189.

[ABN+01] A. Ambainis, E. Bach, A. Nayak, A. Vishwanath, and J. Watrous, *One-dimensional quantum walks*, pp. 37–49 in *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing*, ACM Press, New York, 2001.

[ADZ93] Y. Aharonov, L. Davidovich, and N. Zagury, *Quantum random walks*, Phys. Rev. A **48** (1993), no. 2, 1687–1690.

[AGZV88] V. I. Arnol'd, S. M. Guseĭn-Zade, and A. N. Varchenko, *Singularities of Differentiable Maps. Vol. II. Monodromy and Asymptotics of Integrals*, Birkhäuser, Boston/Basel, 1988. Translated from the Russian by H. Porteous, translation revised by the authors and J. Montaldi.

[BBBP08] Y. Baryshnikov, W. Brady, A. Bressler, and R. Pemantle, *Two-dimensional quantum random walk*, preprint, 34 pp., 2008. Available as arXiv:0810.5495.

[BP04] Y. Baryshnikov and R. Pemantle, *Convolutions of inverse linear functions via multivariate residues*, preprint, 48 pp., 2004.

[BP08] ———, *Tilings, groves and multiset permutations: Asymptotics of rational generating functions whose pole set is a cone*, preprint, 79 pp., 2008. Available as arXiv:0810.4898.

[CEP96] H. Cohn, N. Elkies, and J. Propp, *Local statistics for random domino tilings of the Aztec diamond*, Duke Math. J. **85** (1996), no. 1, 117–166.

[CLO98] D. A. Cox, J. B. Little, and D. O'Shea, *Using Algebraic Geometry*, no. 185 in Graduate Texts in Mathematics, Springer-Verlag, New York/Berlin, 1998.

[DeV10] T. DeVries, *A case study in bivariate singularity analysis*, pp. 61–81 in *Algorithmic Probability and Combinatorics* (M. E. Lladser, R. S. Maier, M. Mishna, and A. Rechnitzer, eds.), American Mathematical Society, Providence, RI, 2010.

[Ell85] R. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*, vol. 271 of Grundlehren der mathematischen Wissenschaften, Springer-Verlag, New York/Berlin, 1985.

[FO90] Ph. Flajolet and A. Odlyzko, *Singularity analysis of generating functions*, SIAM J. Discrete Math. **3** (1990), no. 2, 216–240.

[FS09] Ph. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge Univ. Press, Cambridge, UK, 2009.

[GKZ94] I. Gel'fand, M. Kapranov, and A. Zelevinsky, *Discriminants, Resultants and Multidimensional Determinants*, Birkhäuser, Boston/Basel, 1994.

[GM88] M. Goresky and R. MacPherson, *Stratified Morse Theory*, no. 14 in Ergebnisse der Mathematik und ihrer Grenzgebiete, Springer-Verlag, New York/Berlin, 1988.

[GS64] I. Gel'fand and G. Shilov, *Generalized Functions, Volume I: Properties and Operations*, Academic Press, New York, 1964.

[Hör90] L. Hörmander, *An Introduction to Complex Analysis in Several Variables*, North-Holland Publishing Co., Amsterdam, 3rd edition, 1990.

[Kem03] J. Kempe, *Quantum random walks: An introductory overview*, Contemp. Phys. **44** (2003), no. 4, 307–327.

[Mat70] J. Mather, *Notes on topological stability*, mimeographed notes, 1970.

[Pem00] R. Pemantle, *Generating functions with high-order poles are nearly polynomial*, pp. 305–321 in *Mathematics and Computer Science (Versailles, 2000)*, Birkhäuser, Boston/Basel, 2000.

[Pem09a] ———, *Lecture notes on analytic combinatorics in several variables*, unpublished manuscript, 210 pp, 2009.

[Pem09b] R. Pemantle, *Uniform local central limits on* $\mathbb{Z}^d$, preprint, 6 pp., 2009.

[PS05] T. K. Petersen and D. Speyer, *An arctic circle theorem for groves*, J. Comb. Theory Ser. A **111** (2005), no. 1, 137–164.

[PW02] R. Pemantle and M. C. Wilson, *Asymptotics of multivariate sequences. I. Smooth points of the singular variety*, J. Combin. Theory Ser. A **97** (2002), no. 1, 129–161.

[PW04] ———, *Asymptotics of multivariate sequences, II. Multiple points of the singular variety*, Combin. Probab. Comput. **13** (2004), no. 4–5, 735–761.

[PW08] ———, *Twenty combinatorial examples of asymptotics derived from multivariate generating functions*, SIAM Review **50** (2008), no. 2, 199–272.

[PW09] ———, *Asymptotic expansions of oscillatory integrals with complex phase*, preprint, 17 pp., 2009.

[Ran86] R. M. Range, *Holomorphic Functions and Integral Representations in Several Complex Variables*, no. 108 in Graduate Texts in Mathematics, Springer-Verlag, New York/Berlin, 1986.

[Rie49] M. Riesz, *L'intégrale de Riemann–Liouville et le problème de Cauchy*, Acta Math. **81** (1949), 1–223.

[RW08] A. Raichev and M. C. Wilson, *A new approach to asymptotics of Maclaurin coefficients of algebraic functions*, preprint, 2008.

[The02] T. Theobald, *Computing amoebas*, Experiment. Math. **11** (2002), no. 4, 513–526.

[Var77] A. N. Varchenko, *Newton polyhedra and estimation of oscillating integrals*, Funct. Anal. Appl. **10** (1976), no. 3, 175–196. Russian original in Funktsional. Anal. i Prilozhen. **10** (1976), no. 3, 13–38.

Department of Mathematics, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104, USA

E-mail address: pemantle@math.upenn.edu

URL: http://www.math.upenn.edu/~pemantle

Contemporary Mathematics
Volume **520**, 2010

Asymptotic expansions of oscillatory integrals with complex phase

Robin Pemantle and Mark C. Wilson

Abstract. We consider saddle point integrals in d variables whose phase functions are neither real nor purely imaginary. Results analogous to those for Laplace (real phase) and Fourier (imaginary phase) integrals hold whenever the phase function is analytic and nondegenerate. These results generalize what is well known for integrals of Laplace and Fourier type. The proofs are via contour shifting in complex d-space. This work is motivated by applications to asymptotic enumeration.

1. Introduction

Integrals of the form

$$\mathcal{I}(\lambda) := \mathcal{I}(\lambda; \phi, A) := \int e^{-\lambda\phi(x)} A(x)\,\mathrm{d}x \tag{1.1}$$

arise in many areas of mathematics. There are many variations. This integral may involve one or more variables; the variables may be real or complex; the integral may be global or taken over a small neighborhood or oddly shaped set; varying degrees of smoothness may be assumed; and varying degrees of degeneracy may be allowed near the critical points of the phase function, ϕ. Often what is sought is a leading order estimate of $\mathcal{I}(\lambda)$ as the positive real parameter λ tends to ∞, or an asymptotic series $\mathcal{I}(\lambda) \sim \sum_n c_n g_n(\lambda)$, where $\{g_n\}$ is a sequence of elementary functions and the expansion is possibly nowhere convergent, but satisfies

$$\mathcal{I}(\lambda) - \sum_{n=0}^{N-1} c_n g_n(\lambda) = O(g_N(\lambda)) \tag{1.2}$$

for all $N \geqslant 1$.

In recent work on the asymptotics of multivariate generating functions [**PW02, PW04, PW08, BBBP08, BP08, BP04**], we have required results of this type in the case where the phase function ϕ and the amplitude function A are analytic functions of several variables. The phase function is typically neither real nor purely

2000 *Mathematics Subject Classification.* Primary 41A60; Secondary 41A63, 05A16.

Key words and phrases. Laplace, Fourier, smooth, analytic, contour methods, Cauchy integral formula, saddle point, stationary phase.

Research supported in part by National Science Foundation grant no. DMS-0603821.

imaginary, but satisfies $\mathrm{Re}\{\phi\} \geqslant 0$. The domain of integration in many of these cases is the product of a manifold with a closed simplex. Although results are known that apply to these cases, the state of the literature is quite poor from the standpoint of ready application. We therefore wrote this note to remedy this situation and to supply what would have been ghost citations in [**PW04**, Lemmas 4.7 and 4.8].

In Section 7, we review the existing literature in greater detail. To explain briefly the relation between the present results and previous work, we first discuss two special cases in which the results are well known, but by substantially different methods.

Integrals of Fourier type. Let $f, A\colon \mathbb{R}^d \to \mathbb{R}$ be smooth (that is, C^∞) functions of d real variables, with A having compact support. Taking $\phi = -\mathrm{i}f$ gives the Fourier-type integral

$$\mathcal{I}(\lambda) = \int e^{\mathrm{i}\lambda f(x)} A(x)\,\mathrm{d}x.$$

The standard method of studying this integral is as follows. If f has no critical points in the support of A, then integration by parts shows $\mathcal{I}(\lambda)$ to be rapidly decreasing: $\mathcal{I}(\lambda) = O(\lambda^{-N})$ for all positive integers, N. Using a partition of unity, the integral may therefore be localized to neighborhoods of the critical points of f. At an isolated critical point ∇f vanishes; if the Hessian matrix is non-degenerate then the Morse lemma produces a smooth change of variables under which $f(x) = S_\pm(x) := \sum_{j=1}^d (\pm)_j x_j^2$. This reduces the general problem to the case $\phi = \mathrm{i}S$. To solve this, expand A in a series. Each term may be explicitly integrated, resulting in an expansion in decreasing powers of λ:

$$\mathcal{I}(\lambda) = \sum_{n \geqslant 0} c_n\, \lambda^{-d/2-n}. \tag{1.3}$$

To see that the resulting series for $\mathcal{I}(\lambda)$ satisfies (1.2), one uses integration by parts again to bound the remainder term. The coefficients $\{c_n\}$ are determined by the derivatives of A and ϕ. A particularly lucid treatment of this may be found in the first two sections of [**Ste93**, Chapter VIII].

Integrals of Laplace type. The other well studied case is the Laplace-type integral, where ϕ is real. Localization of the integral to the minima of ϕ is immediate because the integrand is exponentially small elsewhere. Integrating over balls whose radius has order $\lambda^{-1/2}$ shows that near a quadratically nondegenerate minimum, x_0, the bound $A(x) = O(|x - x_0|^N)$ translates into the bound

$$\mathcal{I}(\lambda) = O(\lambda^{-(d+N)/2}). \tag{1.4}$$

Again one changes variables, expands A into a power series, and integrates term by term, to obtain the series

$$\mathcal{I}(\lambda) = \sum_{n \geqslant 0} c_n\, \lambda^{-(d+n)/2}. \tag{1.5}$$

Applying (1.4) to the integral of the remainder term shows that (1.5) is an asymptotic expansion for $\mathcal{I}(\lambda)$. Classical treatments of integrals of Laplace type may be found in many places [**BH86, Won89**], often accompanied by separate treatments of the Fourier case.

Complex methods. The series (1.3) and (1.5) are formally identical. Further inspection of formulae for c_n in the literature shows these to be nearly the same as well, differing only by constant factors of unit modulus. This points to the possibility of unifying the results and generalizing to arbitrary complex functions. An assumption of analyticity is required; technically, this is stronger than smoothness but in practice one is never satisfied without the other. Assuming analyticity, derivations in the Fourier and Laplace cases may indeed be unified via a hybrid approach. In one variable, this is carried out all the time under various names such as "steepest descent," "stationary phase" or "saddle point." If one writes $\mathcal{I}(\lambda)$ as a complex contour integral, the critical point will be a saddle point in $\mathbb{C}^1$ for the real part of the phase function; the contour may then be re-oriented to pass through the saddle in the direction of steepest descent of $-\phi$, converting the integral into one of Fourier type and explaining why the series are nearly identical. This is carried out, for example, in [**dB81**] or [**BH86**, Chapter 7]. Our aim here is to use similar methods in d dimensions to derive asymptotic expansions for integrals of the form (1.1) such as arise in generating function applications.

The organization of the rest of the paper is as follows. In the next section we give some definitions having to do with stratified spaces. We then state our main result, Theorem 2.3. Section 3 records some easy computations in the case where ϕ is the standard phase function $S(\mathbf{x}) := \sum_{j=1}^{d} x_j^2$. Section 4 handles the general case under the assumption that the real part of ϕ has a strict minimum at the critical point. Theorem 2.3 is proved in Section 5. Section 6 gives an application from [**PW04**] to the estimation of coefficients of a bivariate generating function. Finally, Section 7 discusses the relation to existing literature and further research directions.

2. Notation and statement of results

Stratified spaces. Because of the useful properties ensuing from the definition, we shall use Whitney stratified spaces as our chains of integration. Aside from these useful properties, the details of the definition need not concern us, though for completeness we give a precise definition. Let I be a finite partially ordered set and define an I-decomposition of a topological space Z to be a partition of Z into a disjoint union of sets $\{S_\alpha : \alpha \in I\}$ such that

$$S_\alpha \cap \overline{S_\beta} \neq \varnothing \iff S_\alpha \subseteq \overline{S_\beta} \iff \alpha \leqslant \beta.$$

Definition 2.1 (Whitney stratification). Let Z be a closed subset of $\mathbb{R}^n$. A *Whitney stratification* of Z is an I-decomposition such that

(i) Each S_α is a manifold in $\mathbb{R}^n$.
(ii) If $\alpha < \beta$, if the sequences $\{x_i \in S_\beta\}$ and $\{y_i \in S_\alpha\}$ both converge to $y \in S_\alpha$, if the lines $l_i = \overline{x_i\, y_i}$ converge to a line l and the tangent planes $T_{x_i}(S_\beta)$ converge to a plane T of some dimension, then both l and $T_y(S_\alpha)$ are contained in T.

For example, any manifold is a Whitney stratified space with one stratum; any manifold with boundary is a Whitney stratified space with two strata, one being the interior and one the boundary; a k-simplex is a Whitney stratified space whose strata are all its faces. Whitney stratified spaces are closed under products and the set of products of strata will stratify the product. Every algebraic variety admits a Whitney stratification, although the singular locus filtration may be too coarse to be a Whitney stratification. We remark that our definition supposes a stratified

space always to be embedded in $\mathbb{R}^n$, but even the more general definition in [**GM88**] supposes that the relation between the strata is inherited via embedding in some ambient manifold.

To fix ideas, consider the example of a cylinder $S^1 \times [0,1]$, or more generally, the product of a compact manifold (without boundary) and a simplex. These are precisely the types of domain encountered in applications to asymptotic extraction of coefficients.

Critical points. Associated with the definition of a stratification is the stratified notion of a critical point. Observe that under this definition, any zero-dimensional stratum of $\mathcal{M}$ is a critical point of $\mathcal{M}$.

Definition 2.2 (smooth functions and their critical points). Say that a function $\phi\colon \mathcal{M} \to \mathbb{C}$ on a stratified space $\mathcal{M}$ is *smooth* if it is smooth when restricted to each stratum. A point $p \in \mathcal{M}$ is said to be *critical* for the smooth function ϕ if and only if the restriction $\mathrm{d}\phi_{|S}$ vanishes, where S is the stratum containing p.

Let $\mathcal{M} \subseteq \mathbb{C}^d$ be a real analytic, d-dimensional stratified space. This means that each stratum S is a subset of $\mathbb{C}^d$ and each of the chart maps ψ from a neighborhood of the origin in $\mathbb{R}^k$ to some k-dimensional stratum $S \subseteq \mathbb{C}^d$ is analytic (the coordinate functions are convergent power series) with a nonsingular differential. It follows that ψ may be extended to a holomorphic map on a neighborhood of the origin in $\mathbb{C}^k$, whose range we denote by $S \otimes \mathbb{C}$. Choosing a small enough neighborhood, we may arrange for $S \otimes \mathbb{C}$ to be a complex k-manifold embedded in $\mathbb{C}^d$. We say a function $f\colon \mathcal{M} \to \mathbb{C}$ is analytic if it has a convergent power series expansion in a neighborhood of every point; an analytic function on $\mathcal{M}$ may be extended to a complex analytic one on a neighborhood of $\mathcal{M}$ in $\mathcal{M} \otimes \mathbb{C} := \bigcup_{\alpha \in I} S_\alpha \otimes \mathbb{C}$. Because we are interested in the integrals of d-forms over $\mathcal{M}$, there is no loss of generality in assuming that $\mathcal{M}$ is contained in the closure of its d-dimensional strata, whence $\mathcal{M} \otimes \mathbb{C}$ is a neighborhood of $\mathcal{M}$ in $\mathbb{C}^d$. Real d-manifolds in $\mathbb{C}^d$ are not naturally oriented, so we must assume that an orientation is given for each d-stratum of $\mathcal{M}$, meaning that the chart maps from $\mathbb{R}^d$ to $\mathcal{M}$ must preserve orientations.

Definition. The critical point $p \in \mathcal{M}$ is said to be *quadratically nondegenerate* if p is in a d-dimensional stratum and $\mathcal{H}(p)$ is nonsingular, where

$$\mathcal{H}(p) := \left(\frac{\partial^2 \phi}{\partial x_i \, \partial x_j} \right)_{1 \leqslant i,j \leqslant d}$$

is the Hessian matrix for the function ϕ of Definition 2.2 on a neighborhood of p in $\mathbb{C}^d$.

A point $p \in \mathcal{M}$ is said to be *stationary* if it is critical and $\mathrm{Re}\{\phi(p)\} = 0$.

Results. The main result of the paper is an asymptotic expansion for $\mathcal{I}(\lambda)$ in the case where all critical points are in top-dimensional strata and are quadratically nondegenerate. It may seem surprising, after taking the trouble to work in the generality of stratified spaces, to restrict our results to the simplest case. Indeed, combinatorial applications [**BM93, PW04, BP04**] are known for which the relevant integrals have quadratically nondegenerate critical points in lower-dimensional strata. However, the method in [**PW04**] is to resolve the multiple intersection via a lifting to a simplex, thereby lifting the integral to an integral over the product of a smooth space with a simplex, with the critical points all lifting to top-dimensional

strata. In other words, the generality of a stratified domain of integration is needed, but the critical points are forced into top-dimensional strata. We remark that the taxonomy of degeneracies and lower-dimensional critical points is complicated but reasonably well understood [**Var77, AGZV88, Wat01**].

THEOREM 2.3. *Let $\mathcal{M}$ be a compact, real analytic stratified space of dimension d embedded in $\mathbb{C}^d$. Let A and ϕ be analytic functions on a neighborhood of $\mathcal{M}$ and suppose* $\operatorname{Re}\{\phi\} \geqslant 0$ *on $\mathcal{M}$. Let G be the subset of stationary points of ϕ on $\mathcal{M}$ and assume that G is finite and each stationary point is quadratically nondegenerate. Then the integral $\mathcal{I}(\lambda) := \int_{\mathcal{M}} e^{-\lambda\phi(x)} A(x)\,\mathrm{d}x$ has an asymptotic expansion*

$$\mathcal{I}(\lambda) \sim \sum_{\ell=0}^{\infty} c_\ell \lambda^{-(d+\ell)/2}.$$

If A is nonzero at some point of G then the leading term is given by

$$c_0 = \pi^{d/2} \sum_{x\in G} A(x)\,\omega(x,\lambda)\,(\det \mathcal{H}(x))^{-1/2}\,. \tag{2.1}$$

where $\omega(x,\lambda)$ is the unit complex number $e^{-\lambda\phi(x)}$. The choice of sign of the square root on the right-hand side of (2.1) *is determined by choosing any analytic chart map ψ for a neighborhood of x and defining*

$$(\det \mathcal{H}(x))^{-1/2} := (\det \mathcal{H}(\phi\circ\psi))^{-1/2}\,\det J_\psi,$$

where J_ψ is the Jacobian matrix for ψ at x and the $-1/2$ power of the determinant on the right is the product of the inverses of principal square roots of the eigenvalues.

REMARK. As usual, by $\int_{\mathcal{M}}$ we mean $\int_{\mathcal{C}}$ for some chain $\mathcal{C}$ representing $\mathcal{M}$.

As mentioned above, analyses involving critical points in lower-dimensional strata are not required. There is one exception, which arises in [**PW04**, Lemma 4.7 (ii)]. Here, there is a critical point $p \in \partial S$, where S is a top-dimensional stratum, a neighborhood of p in $\overline{S}$ is diffeomorphic to a half-space, and the gradient of ϕ at p vanishes not only on the $(d-1)$-dimensional stratum containing p but in the d-dimensional half-space neighborhood; quadratic nondegeneracy is assumed in all d dimensions as well.

COROLLARY 2.4. *Assume the hypotheses of Theorem* 2.3 *except that some of the points $p \in G$ are as in the previous paragraph: critical and nondegenerate in a d-dimensional half-space neighborhood. Then the same conclusion holds, except that the summand in* (2.1) *corresponding to such a point p must be multiplied by* $1/2$.

EXAMPLE 2.5. Let $\mathcal{M} = [0,1]\times(\mathbb{R}/2\pi)$ and let $\phi(x,y) = x^2 + 2\mathrm{i}\,x\cos(y)$. There are no critical points in the interior. Every point with $x = 0$ is a critical point of the stratum $\{0\}\times S^1$, but for $(0,y)$ to be a critical point in the half space, the gradient $(2\mathrm{i}\cos y, 0)$ must vanish, so $y = \pm\pi/2$. Half the Hessian is

$$\begin{bmatrix} 1 & \mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix},$$

which is nonsingular. Therefore there are two isolated, quadratically nondegenerate critical points, to which Corollary 2.4 may be applied. Note that the vanishing of ϕ on a part of $\partial\mathcal{M}$ renders many classical results inapplicable, though the specific result may be obtained by hand in a variety of ways.

3. Preliminary results for the standard phase function

For $\mathbf{x} \in \mathbb{R}^d$, let $S(\mathbf{x}) := \sum_{j=1}^d x_j^2$ denote the standard quadratic form. We begin with a couple of results on integrals of Laplace type where the phase function is the standard quadratic and A is a monomial. In one dimension,

PROPOSITION 3.1.

$$\int_{-\infty}^{\infty} x^n \, e^{-\lambda x^2} \, \mathrm{d}x = \frac{n! \, \sqrt{\pi}}{(n/2)! \, 2^n} \, \lambda^{-1/2-n/2}$$

if n is even, while the integral is zero if n is odd.

PROOF. For odd n the result is obvious from the fact that the integrand is an odd function. For $n = 0$ the result is just the standard Gaussian integral. By induction, assume now the result for $n-2$. Integrating by parts to get the second line, we have

$$\begin{aligned}
\int_{-\infty}^{\infty} x^n e^{-\lambda x^2} \, \mathrm{d}x &= \int_{-\infty}^{\infty} \frac{-x^{n-1}}{2\lambda} \left(-2\lambda x \, e^{-\lambda x^2} \, \mathrm{d}x \right) \\
&= \frac{n-1}{2\lambda} \int_{-\infty}^{\infty} x^{n-2} \, e^{-\lambda x^2} \, \mathrm{d}x \\
&= \frac{n-1}{2\lambda} \, \frac{(n-2)! \, \sqrt{\pi}}{(n/2-1)! \, 2^{n-1}} \, \lambda^{1/2-n/2}
\end{aligned}$$

by the induction hypothesis. This is equal to $\lambda^{-1/2-n/2} \sqrt{\pi} \, n! \, / \, ((n/2)! \, 2^n)$, completing the induction. □

COROLLARY 3.2 (monomial integral). *Let $\mathbf{r}$ be any d-vector of nonnegative integers and let $\mathbf{x}^{\mathbf{r}}$ denote the monomial $x_1^{r_1} \cdots x_d^{r_d}$. Then*

$$\int_{\mathbb{R}^d} \mathbf{x}^{\mathbf{r}} e^{-\lambda S(\mathbf{x})} \, \mathrm{d}\mathbf{x} = \beta_{\mathbf{r}} \lambda^{-(d+|\mathbf{r}|)/2},$$

where

$$\beta_{\mathbf{r}} := \pi^{d/2} \prod_{j=1}^{d} \frac{r_j!}{(r_j/2)! \, 2^{r_j}} \tag{3.1}$$

if all the components r_j are even, and zero otherwise.

PROOF. The integral factors into

$$\prod_{j=1}^{d} \left[\int_{-\infty}^{\infty} x_j^{r_j} \, e^{-\lambda r_j^2} \, \mathrm{d}x_j \right]. \qquad \square$$

To integrate term by term over a Taylor series for A, we need the following estimate.

LEMMA 3.3 (big-O estimate). *Let A be any smooth function satisfying $A(\mathbf{x}) = O(|\mathbf{x}|^r)$ at the origin. Then the integral of $A(\mathbf{x}) e^{-\lambda S(\mathbf{x})}$ over any compact set K may be bounded from above as*

$$\int_K A(\mathbf{x}) e^{-\lambda S(\mathbf{x})} \, \mathrm{d}\mathbf{x} = O(\lambda^{-(d+r)/2}).$$

PROOF. Because K is compact and $A(\mathbf{x}) = O(|\mathbf{x}|^r)$ at the origin, it follows that there is some constant c for which $|A(\mathbf{x})| \leqslant c|\mathbf{x}|^r$ on all of K. Let K_0 denote the intersection of K with the ball $|\mathbf{x}| \leqslant \lambda^{-1/2}$, and for $n \geqslant 1$ let K_n denote the intersection of K with the shell $2^{n-1}\lambda^{-1/2} \leqslant |\mathbf{x}| \leqslant 2^n\lambda^{-1/2}$. On K_0 we have

$$|A(\mathbf{x})| \leqslant c\lambda^{-r/2},$$

while trivially

$$\int_{K_0} e^{-\lambda S(\mathbf{x})}\,\mathrm{d}\mathbf{x} \leqslant \int_{K_0} \mathrm{d}\mathbf{x} \leqslant c_d\lambda^{-d/2}.$$

Thus

$$\left|\int_{K_0} A(\mathbf{x})e^{-\lambda S(\mathbf{x})}\,\mathrm{d}\mathbf{x}\right| \leqslant c'\lambda^{-(r+d)/2}.$$

For $n \geqslant 1$, on $K \cap K_n$, we have the upper bounds

$$\begin{aligned} |A(\mathbf{x})| &\leqslant 2^{rn} \quad c\lambda^{-r/2} \\ e^{-\lambda S(\mathbf{x})} &\leqslant \quad e^{-2^{n-1}} \\ \int_{K_n} \mathrm{d}\mathbf{x} &\leqslant \quad 2^{rn}\, c_d\, \lambda^{-d/2}. \end{aligned}$$

Letting $c'' := c \cdot c_d \cdot \sum_{n=1}^{\infty} 2^{2rn}\, e^{-2^{n-1}} < \infty$, we may sum to find that

$$\sum_{n=0}^{\infty} \left|\int_{K_n} A(\mathbf{x})e^{-\lambda S(\mathbf{x})}\,\mathrm{d}\mathbf{x}\right| \leqslant (c' + c'')\lambda^{-(r+d)/2},$$

proving the lemma. □

It is now easy to compute a series for $\mathcal{I}(\lambda)$ in the case where ϕ is the standard quadratic and the integral is over a neighborhood of the origin in $\mathbb{R}^d$.

THEOREM 3.4 (standard phase). *Let $A(\mathbf{x}) = \sum_{\mathbf{r}} a_{\mathbf{r}}\mathbf{x}^{\mathbf{r}}$ be an analytic function defined on a neighborhood $\mathcal{N}$ of the origin in $\mathbb{R}^d$. Let*

$$\mathcal{I}(\lambda) := \int_{\mathcal{N}} A(\mathbf{x})e^{-\lambda S(\mathbf{x})}\,\mathrm{d}\mathbf{x}. \tag{3.2}$$

Then

$$\mathcal{I}(\lambda) \sim \sum_n \left[\sum_{|\mathbf{r}|=n} a_{\mathbf{r}}\beta_{\mathbf{r}}\right] \lambda^{-(n+d)/2}$$

as an asymptotic series expansion in decreasing powers of λ, with $\beta_{\mathbf{r}}$ as in (3.1).

PROOF. Write $A(\mathbf{x})$ as a power series up to degree N plus a remainder term:

$$A(\mathbf{x}) = \left(\sum_{n=0}^{N}\sum_{|\mathbf{r}|=n} a_{\mathbf{r}}\mathbf{x}^{\mathbf{r}}\right) + R(\mathbf{x}),$$

where $R(\mathbf{x}) = O(|\mathbf{x}|^{N+1})$. Using Corollary 3.2 to integrate all the monomial terms and Lemma 3.3 to bound the integral of $R(\mathbf{x})e^{-\lambda S(\mathbf{x})}$ shows that

$$\mathcal{I}(\lambda) = \sum_{n=0}^{N}\sum_{|\mathbf{r}|=n} a_{\mathbf{r}}\beta_{\mathbf{r}}\lambda^{-(d+n)/2} + O(\lambda^{-(d+n)/2-1}),$$

which proves the asymptotic expansion. □

4. The case of a strict minimum

In this section, we continue to integrate over a neighborhood of the origin in $\mathbb{R}^d$, but we generalize to any analytic phase function ϕ with the restriction that the real part of ϕ have a strict minimum at the origin. The assumption of a strict minimum localizes the integral to the origin, so the only tricky aspects are keeping track of the sign (Lemma 4.3) and being rigorous about moving the contour.

THEOREM 4.1. *Let A and ϕ be complex-valued analytic functions on a compact neighborhood $\mathcal{N}$ of the origin in $\mathbb{R}^d$, and suppose the real part of ϕ is nonnegative, vanishing only at the origin. Suppose the Hessian matrix $\mathcal{H}$ of ϕ at the origin is nonsingular. Letting $\mathcal{I}(\lambda) := \int_{\mathcal{N}} A(\mathbf{x}) e^{-\lambda\phi(\mathbf{x})}\,\mathrm{d}\mathbf{x}$, one has an asymptotic expansion*

$$\mathcal{I}(\lambda) \sim \sum_{\ell \geqslant 0} c_\ell \lambda^{d/2-\ell},$$

where

$$c_0 = A(\mathbf{0}) \frac{\pi^{d/2}}{\sqrt{\det \frac{1}{2}\mathcal{H}}} = A(\mathbf{0}) \frac{(2\pi)^{d/2}}{\sqrt{\det \mathcal{H}}}$$

and the choice of sign is defined by taking the product of the principal square roots of the eigenvalues of $\mathcal{H}$.

The proof is essentially a reduction to the case of standard phase. The key is the well known Morse Lemma. The proof given in [**Ste93**, VIII:2.3.2] is for the smooth category and for purely real or imaginary phase but extends without significant change to complex values and the analytic category. For completeness, we include the adapted proof.

LEMMA 4.2 (Morse lemma). *There is a bi-holomorphic change of variables $\mathbf{x} = \psi(\mathbf{y})$ such that $\phi(\psi(\mathbf{y})) = S(\mathbf{y}) := \sum_{j=1}^d y_j^2$. The differential $J_\psi = \mathrm{d}\psi(\mathbf{0})$ will satisfy $(\det J_\psi)^2 = (\det \frac{1}{2}\mathcal{H})^{-1}$.*

PROOF. Addressing the second conclusion first, we recall how the Hessian matrix behaves under a change of variables. If $\psi\colon \mathbb{C}^d \to \mathbb{C}^d$ is bi-holomorphic on a neighborhood of $\mathbf{x}$ and if ϕ has vanishing gradient at $\psi(\mathbf{x})$ and Hessian matrix $\mathcal{H}$ there, then the Hessian matrix $\tilde{\mathcal{H}}$ of $\phi \circ \psi$ at $\mathbf{x}$ is given by

$$\tilde{\mathcal{H}} = J_\psi^T \mathcal{H} J_\psi,$$

where J_ψ is the Jacobian matrix for ψ at $\mathbf{x}$. The standard form S has Hessian matrix equal to twice the identity, hence any function ψ satisfying $\phi \circ \psi = S$ must satisfy

$$2I = J_\psi^T \mathcal{H} J_\psi.$$

Dividing by two and taking determinants yields $|\mathbf{J}_\psi|^2 \det(\frac{1}{2}\mathcal{H}) = 1$, proving the second conclusion.

To prove the change of variables, the first step is to write

$$\phi(\mathbf{x}) = \sum_{j,k=1}^d x_j x_k \phi_{j,k},$$

where the functions $\phi_{j,k} = \phi_{k,j}$ are analytic and satisfy $\phi_{j,k}(\mathbf{0}) = (1/2)\,\mathcal{H}_{j,k}$. It is obvious from a formal power series viewpoint that this can be done because the summand $x_j x_k \phi_{j,k}$ can be any power series with coefficients indexed by the

orthant $\{\mathbf{r} : \mathbf{r} \geqslant \mathbf{e}_j + \mathbf{e}_k\}$. These orthants cover $\{\mathbf{r} : |\mathbf{r}| \geqslant 2\}$, so we may obtain any function ϕ vanishing to order two; matching coefficients on the terms of order precisely two shows that $\phi_{j,k}(\mathbf{0}) = (1/2)\,\mathcal{H}_{j,k}$.

More constructively, we may give a formula for $\phi_{j,k}$. There is plenty of freedom, but a convenient choice is to let $a_{\mathbf{r}}$ denote the coefficient of $\mathbf{x}^{\mathbf{r}}$ in $\phi(\mathbf{x})$ and to take

$$x_k x_k \phi_{j,k}(\mathbf{x}) := \sum_{|\mathbf{r}|\geqslant 2} \frac{r_j(r_k - \delta_{j,k})}{|\mathbf{r}|(|\mathbf{r}|-1)}\, a_{\mathbf{r}} \mathbf{x}^{\mathbf{r}}.$$

For fixed $\mathbf{r}$, it is easy to check that

$$\sum_{1\leqslant j,k\leqslant d} \frac{r_j(r_k - \delta_{j,k})}{|\mathbf{r}|(|\mathbf{r}|-1)} = 1,$$

whence $\phi = \sum x_j x_k \phi_{j,k}$. Alternatively, the following analytic computation from [**Ste93**] verifies that $\phi = \sum_{j,k} x_j x_k \phi_{j,k}$. Any function f vanishing at zero satisfies $f(t) = \int_0^1 (1-s) f'(s)\, ds$, as may be seen by integrating by parts (take $g(s) = -(1-s)$). Fix $\mathbf{x}$ and apply this fact with $f(t) = (\mathrm{d}/\mathrm{d}t)\phi(t\mathbf{x})$ to obtain

$$\phi(\mathbf{x}) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\phi(t\mathbf{x})\,\mathrm{d}t = \int_0^1 (1-t)\frac{\mathrm{d}^2}{\mathrm{d}t^2}\phi(t\mathbf{x})\,\mathrm{d}t.$$

The multivariate chain rule gives

$$\frac{d^2}{\mathrm{d}t^2}\phi(t\mathbf{x}) = \sum_{j,k} x_j x_k \frac{\partial^2 \phi}{\partial x_j \partial x_k}(t\mathbf{x})\,;$$

plug in $\phi = \sum_{\mathbf{r}} a_{\mathbf{r}} \mathbf{x}^{\mathbf{r}}$ and integrate term by term using $\int_0^1 (1-t)t^{n-2}\,\mathrm{d}t = \frac{1}{n(n-1)}$ to see that $\phi = \sum_{j,k} x_j x_k \phi_{j,k}$.

The second step is an induction. Suppose first that $\phi_{j,j}(\mathbf{0}) \neq 0$ for all j. The function $\phi_{1,1}^{-1}$ and a branch of the function $\phi_{1,1}^{1/2}$ are analytic in a neighborhood of the origin. Set

$$y_1 := \phi_{1,1}^{1/2}\left[x_1 + \sum_{k>1} \frac{y_k \phi_{1,k}}{\phi_{1,1}}\right].$$

Expanding, we find that the terms of y_1^2 of total degree at most one in the terms $x_2, \ldots, x_d$ match those of ϕ, and therefore

$$\phi(\mathbf{x}) = y_1^2 + \sum_{j,k\geqslant 2} x_j x_k h_{j,k} \tag{4.1}$$

for some analytic functions $h_{j,k}$ satisfying $h_{j,k}(\mathbf{0}) = (1/2)\,\mathcal{H}_{j,k}$. Similarly, if

$$\phi(\mathbf{x}) = \sum_{j=1}^{r-1} y_j^2 + \sum_{j,k\geqslant r} x_j x_k h_{j,k}$$

then setting

$$y_r := \phi_{r,r}^{1/2}\left[x_r + \sum_{k>r} \frac{y_k \phi_{r,k}}{\phi_{r,r}}\right]$$

gives

$$\phi(\mathbf{x}) = \sum_{j=1}^{r} y_j^2 + \sum_{j,k\geqslant r+1} x_j x_k \tilde{h}_{j,k}$$

for some analytic functions $\tilde{h}_{j,k}$ still satisfying $h_{j,k}(\mathbf{0}) = (1/2)\,\mathcal{H}_{j,k}$. By induction, we arrive at $\phi(\mathbf{x}) = \sum_{j=1}^{d} y_j^2$, finishing the proof of the Morse Lemma in the case where each $\mathcal{H}_{j,j}$ is nonzero.

Finally, if some $\mathcal{H}_{j,k} = 0$, because $\mathcal{H}$ is nonsingular we may always find some unitary map U such that the Hessian $U^T\,\mathcal{H}\,U$ of $\phi \circ U$ has no vanishing diagonal entry. We know there is a ψ_0 such that $(\phi \circ U) \circ \psi_0 = S$, and taking $\psi = U \circ \psi_0$ finishes the proof in this case. □

Proof of Theorem 4.1. The power series allows us to extend ϕ to a neighborhood of the origin in $\mathbb{C}^d$. Under the change of variables ψ from the previous lemma, we see that

$$\begin{aligned}\mathcal{I}(\lambda) &= \int_{\psi^{-1}\mathcal{C}} A \circ \psi(\mathbf{y}) e^{-\lambda S(\mathbf{y})} (\det \mathrm{d}\psi(\mathbf{y}))\,\mathrm{d}\mathbf{y} \\ &= \int_{\psi^{-1}\mathcal{C}} \tilde{A}(\mathbf{y}) e^{-\lambda S(\mathbf{y})}\,\mathrm{d}\mathbf{y}\end{aligned}$$

for some analytic function $\tilde{A}$, where $\mathcal{C}$ is a neighborhood of the origin in $\mathbb{R}^n$. We need to check that we can move the chain $\psi^{-1}\mathcal{C}$ of integration back to the real plane.

Let $h(\mathbf{z}) := \mathrm{Re}\{S(\mathbf{z})\}$. The chain $\mathcal{C}' := \psi^{-1}(\mathcal{C})$ lies in the region $\{\mathbf{z} \in \mathbb{C}^d : h(\mathbf{z}) > 0\}$ except when $\mathbf{z} = 0$, and in particular, $h \geqslant \epsilon > 0$ on $\partial\mathcal{C}'$. Let

$$H(\mathbf{z}, t) := \mathrm{Re}\{\mathbf{z}\} + (1-t)\,\mathrm{i}\,\mathrm{Im}\{\mathbf{z}\}.$$

In other words, H is a homotopy from the identity map to the map π projecting out the imaginary part of the vector $\mathbf{z}$. For any chain σ, the homotopy H induces a chain homotopy, $H(\sigma)$, supported on the image of the support of σ under the homotopy H and satisfying

$$\partial H(\sigma) = \sigma - \pi\sigma + H(\partial\sigma).$$

With $\sigma = \mathcal{C}'$, observing that $h(H(\mathbf{z}, t)) \geqslant h(\mathbf{z})$, we see there is a $(d+1)$-chain $\mathcal{D}$ with

$$\partial\mathcal{D} = \mathcal{C}' - \pi\mathcal{C}' + \mathcal{C}''$$

and $\mathcal{C}''$ supported on $\{h > \epsilon\}$. Stokes' Theorem tells us that for any holomorphic d-form ω,

$$\int_{\partial\mathcal{D}} \omega = \int_{\mathcal{D}} \mathrm{d}\omega = 0,$$

and consequently that

$$\int_{\mathcal{C}'} \omega = \int_{\pi\mathcal{C}'} \omega + \int_{\mathcal{C}''} \omega.$$

When $\omega = \tilde{A} e^{-\lambda S}\,\mathrm{d}\mathbf{y}$, the integral over $\mathcal{C}''$ is $O(e^{-\lambda\epsilon})$, giving

$$\mathcal{I}(\lambda) = \int_{\pi\mathcal{C}'} \tilde{A}(\mathbf{y}) e^{-\lambda S(\mathbf{y})}\,\mathrm{d}\mathbf{y} + O(e^{-\epsilon\lambda}).$$

Up to sign, the chain $\pi\mathcal{C}''$ is a disk in $\mathbb{R}^d$ with the standard orientation plus something supported in $\{h > \epsilon\}$. To see this, note that π maps any real d-manifold in $\mathbb{C}^d$ diffeomorphically to $\mathbb{R}^d$ wherever the tangent space is transverse to the imaginary subspace. The tangent space to the support of $\mathcal{C}'$ at the origin is transverse to the imaginary subspace because $S \geqslant 0$ on $\mathcal{C}'$, whereas the imaginary subspace is precisely the negative d-space of the index-d form S. The tangent space varies

continuously, so in a neighborhood of the origin, π is a diffeomorphism. Observing that $\tilde{A}(\mathbf{0}) = A(\mathbf{0}) \det(\mathrm{d}\psi(\mathbf{0})) = A(\mathbf{0})(\det(1/2)\,\mathcal{H})^{-1/2}$ and using Theorem 3.4 finishes the proof, up to the choice of sign of the square root.

The map $\mathrm{d}\pi \circ \mathrm{d}\psi^{-1}(\mathbf{0})$ maps the standard basis of $\mathbb{R}^d$ to another basis for $\mathbb{R}^d$. Verifying the sign choice is equivalent to showing that this second basis is positively oriented if and only if $\det(\mathrm{d}\psi(\mathbf{0}))$ is the product of the principal square roots of the eigenvalues of $\mathcal{H}$ (it must be either this or its negative). Thus we shall be finished if we apply the following lemma (with $\alpha = \psi^{-1}$). □

LEMMA 4.3. *Let $W \subseteq \mathbb{C}^d$ be the set $\{\mathbf{z} : \mathrm{Re}\{S(\mathbf{z})\} > 0\}$. Pick any $\alpha \in GL_d(\mathbb{C})$ mapping $\mathbb{R}^d$ into $\overline{W}$, and let $M := \alpha^T\alpha$ be the matrix representing the composition $S \circ \alpha$. Let $\pi\colon \mathbb{C}^d \to \mathbb{R}^d$ be projection onto the real part. Then $\pi \circ \alpha$ is orientation preserving on $\mathbb{R}^d$ if and only if $\det \alpha$ is the product of the principal square roots of the eigenvalues of M (rather than the negative of this product).*

PROOF. First suppose $\alpha \in GL_d(\mathbb{R})$. Then M has positive eigenvalues, so the product of their principal square roots is positive. The map π is the identity on $\mathbb{R}^d$ so the statement boils down to saying that α preserves orientation if and only if it has positive determinant, which is true by definition. In the general case, let $\alpha_t := \pi_t \circ \alpha$, where $\pi_t(\mathbf{z}) = \mathrm{Re}\{\mathbf{z}\} + (1-t)\,\mathrm{Im}\{\mathbf{z}\}$. As we saw in the previous proof, $\pi_t(\mathbb{R}^d) \subseteq \overline{W}$ for all $0 \leqslant t \leqslant 1$, whence $M_t := \alpha_t^T\alpha_t$ has eigenvalues with nonnegative real parts. The product of the principal square roots of the eigenvalues is a continuous function on the set of nonsingular matrices with no negative real eigenvalues. The determinant of α_t is a continuous function of t, and we have seen that it agrees with the product of principal square roots of eigenvalues of M_t when $t = 1$ (the real case). So by continuity, this is the correct sign choice for all $0 \leqslant t \leqslant 1$; taking $t = 0$ proves the lemma. □

For later use, we record one easy corollary of Theorem 4.1.

COROLLARY 4.4. *Assume the hypotheses of Theorem* 4.1 *and let $\mathcal{N}'$ be the intersection of $\mathcal{N}$ with a region diffeomorphic to a half-space through the origin. If $A(\mathbf{0}) \neq 0$ then*

$$\mathcal{I}'(\lambda) := \int_{\mathcal{N}'} A(\mathbf{x}) e^{-\lambda\phi(\mathbf{x})}\,\mathrm{d}x \sim \frac{c_0}{2}\,\lambda^{-d/2},$$

where c_0 is the same as in the conclusion of Theorem 4.1.

PROOF. Under the change of variables ψ and the projection π, this region maps to a region $\mathcal{N}''$ diffeomorphic to a half-space with the origin on the boundary. Changing variables by $\mathbf{y} = \lambda^{-1/2}\mathbf{x}$ and writing $\mathcal{N}_\lambda$ for $\lambda^{1/2}\mathcal{N}''$, we have

$$\mathcal{I}'(\lambda) = \lambda^{-d/2} \int_{\mathcal{N}_\lambda} A_\lambda(\mathbf{y}) e^{-S(\mathbf{y})}\,\mathrm{d}\mathbf{y},$$

where $A_\lambda(\mathbf{y}) = (A \circ \psi)(\lambda^{-1/2}\mathbf{y})$. The function A_λ converges to $A(\mathbf{0})$ pointwise and also in $L^2(\mu)$, where μ is the Gaussian measure $e^{-S(\mathbf{x})}\,\mathrm{d}\mathbf{x}$. Also, the regions $\mathcal{N}_\lambda$ converge to a half-space H in the sense that their indicators $\mathbf{1}_{\mathcal{N}_\lambda}$ converge to $\mathbf{1}_H$ in $L^2(\mu)$. Thus $A_\lambda \mathbf{1}_{\mathcal{N}_\lambda}$ converges to $A(\mathbf{0})\mathbf{1}_H$ in $L^1(\mu)$, and unravelling this statement we see that

$$\int_{\mathcal{N}_\lambda} A_\lambda(\mathbf{y}) e^{-S(\mathbf{y})}\,\mathrm{d}\mathbf{y} \to \int_H A(\mathbf{0}) e^{-S(\mathbf{y})}\,\mathrm{d}\mathbf{y}.$$

The last quantity is equal to $c_0/2$, showing that $\lambda^{d/2}\mathcal{I}'(\lambda) \to c_0/2$ and finishing the proof. □

5. Proofs of main results

Theorem 2.3 differs from Theorem 4.1 in several ways. The most important is that the set where $\operatorname{Re}\phi$ vanishes may extend to the boundary of the region of integration. This precludes the use of the easy deformation π, because $\mathcal{C}''$ is no longer supported on $\{h > \epsilon\}$. Consequently, some work is required to construct a suitable deformation. We do so via notions from stratified Morse theory [**GM88**].

Tangent vector fields. If $\mathbf{x}$ is a point of the stratum S of the stratified space $\mathcal{M}$, let $T_{\mathbf{x}}(\mathcal{M})$ denote the tangent space to S at $\mathbf{x}$. Because $\mathcal{M}$ is embedded in $\mathbb{C}^d$, the tangent spaces may all be identified as subspaces of $\mathbb{C}^d$. Thus we have a notion of the tangent bundle $T\mathcal{M}$, a section of which is simply a vector field f on $\mathcal{M} \subseteq \mathbb{C}^d$ such that $f(\mathbf{x}) \in T_{\mathbf{x}}(\mathcal{M})$ for all $\mathbf{x}$. A consequence of the two Whitney conditions is the local product structure of a stratified space: a point $\mathbf{p}$ in a k-dimensional stratum S of a stratified space $\mathcal{M}$ has a neighborhood in which $\mathcal{M}$ is homeomorphic to some product $S \times X$. According to [**GM88**], a proof may be found in mimeographed notes of Mather from 1970; it is based on Thom's Isotopy Lemma which takes up fifty pages of the same mimeographed notes. The next lemma is the only place where we use this (or any) consequence of Whitney stratification.

LEMMA 5.1. *Let f be a smooth section of the tangent bundle to S, that is, $f(s) \in T_s(S)$ for $s \in S$. Then each $s \in S$ has a neighborhood in $\mathcal{M}$ on which f may be extended to a smooth section of the tangent bundle.*

PROOF. Parametrize $\mathcal{M}$ locally by $S \times X$ and extend f by $f(s, x) := f(s)$. □

LEMMA 5.2 (vector field near a non-critical point). *Let $\mathbf{x}$ be a point of the stratum S of the stratified space $\mathcal{M}$, and suppose $\mathbf{x}$ is not critical for the function ϕ. Then there is a vector $\mathbf{v} \in T_{\mathbf{x}}(S \otimes \mathcal{M})$ such that $\operatorname{Re}\{\mathrm{d}\phi(\mathbf{v})\} > 0$ at $\mathbf{x}$. Furthermore, there is a continuous section f of the tangent bundle in a neighborhood $\mathcal{N}$ of $\mathbf{x}$ such that $\operatorname{Re}\{\mathrm{d}\phi(f(\mathbf{y}))\} > 0$ at every $\mathbf{y} \in \mathcal{N}$.*

PROOF. By non-criticality of $\mathbf{x}$, there is a $\mathbf{w} \in T_{\mathbf{x}}(S)$ with $\mathrm{d}\phi(\mathbf{w}) = u \neq 0$ at $\mathbf{x}$. Multiply $\mathbf{w}$ componentwise by $\overline{u}$ to obtain $\mathbf{v}$ with $\operatorname{Re}\{\mathrm{d}\phi(\mathbf{v})\} > 0$ at $\mathbf{x}$. Use any chart map for $S \otimes \mathbb{C}$ near $\mathbf{x}$ to give a locally trivial coordinatization for the tangent bundle and define a section f to be the constant vector $\mathbf{v}$; then $\operatorname{Re}\{\mathrm{d}\phi(f(\mathbf{y}))\} > 0$ on some sufficiently small neighborhood of $\mathbf{x}$ in S. Finally, extend to a neighborhood of $\mathbf{x}$ in $\mathcal{M}$ by Lemma 5.1. □

Although we are working in the analytic category, the chains of integration are topological objects, for which we may use C^∞ methods (in what follows, even C^1 methods will do). In particular, a partition of unity argument enhances the local result above to a global result.

LEMMA 5.3 (global vector field, in the absence of critical points). *Let $\mathcal{M}$ be a compact stratified space and ϕ a smooth function on $\mathcal{M}$ with no critical points. Then there is a global section f of the tangent bundle of $\mathcal{M}$ such that the real part of $\mathrm{d}\phi(f)$ is everywhere positive.*

PROOF. For each point $\mathbf{x} \in \mathcal{M}$, let $f_{\mathbf{x}}$ be a section as in the conclusion of Lemma 5.2, on a neighborhood $U_{\mathbf{x}}$. Cover the compact space $\mathcal{M}$ by finitely many sets $\{U_{\mathbf{x}} : \mathbf{x} \in F\}$, and let $\{\psi_{\mathbf{x}} : \mathbf{x} \in F\}$ be a smooth partition of unity subordinate to this finite cover. Define

$$f(\mathbf{y}) = \sum_{\mathbf{x} \in F} \psi_{\mathbf{x}}(\mathbf{y}) f_{\mathbf{x}}(\mathbf{y}).$$

Then f is smooth; it is a section of the tangent bundle because each tangent space is linearly closed; and the real part of $\mathrm{d}\phi(f(\mathbf{y}))$ is positive because we took a convex combination in which each contribution was nonnegative and at least one was positive. □

Another partition argument gives our final version of this result.

LEMMA 5.4 (global vector field, vanishing only at critical points). *Let $\mathcal{M}$ be a compact stratified space and ϕ a smooth function on $\mathcal{M}$ with finitely many critical points. Then there is a global section f of the tangent bundle of $\mathcal{M}$ such that the real part of $\mathrm{d}\phi(f)$ is nonnegative and vanishes only when $\mathbf{y}$ is a critical point.*

PROOF. Let $\mathcal{M}_\epsilon$ be the compact stratified space resulting in the removal of an ϵ-ball around each critical point of ϕ. Let g_ϵ be a vector field as in the conclusion of Lemma 5.3 with $\mathcal{M}$ replaced by $\mathcal{M}_\epsilon$. Let f_ϵ be the product of g_ϵ with a smooth function that is equal to its maximum of 1 on $M_{2\epsilon}$ and its minimum of 0 on M_ϵ^c. Let $c_n > 0$ be chosen small enough so that the magnitudes of all partial derivatives of $c_n f_{1/n}$ of order up to n are at most 2^{-n}. In the topology of uniform convergence of derivatives of bounded order, the series $\sum_n c_n f_n$ converges to a vector field f with the required properties. □

Proof of Theorem 2.3. Let f be a tangent vector field along which ϕ increases away from critical points, as given by Lemma 5.4. Such a field gives rise to a differential flow, which, informally, is the solution to $\mathrm{d}\mathbf{p}/\mathrm{d}t = f(\mathbf{p})$. To be more formal, let $\mathbf{x}$ be a point in a stratum S of $\mathcal{M}$. Via a chart map in a neighborhood of $\mathbf{x}$, we solve the ODE $\mathrm{d}\Phi(t)/\mathrm{d}t = f(\Phi(t))$ with initial condition $\Phi(0) = \mathbf{x}$, obtaining a trajectory Φ on some interval $[0, \epsilon_{\mathbf{x}}]$ that is supported on S. Doing this simultaneously for all $\mathbf{x} \in \mathcal{M}$ results in a map

$$\Phi \colon \mathcal{M} \times [0, \epsilon] \to \mathbb{C}^d,$$

with $\Phi(\mathbf{x}, t)$ remaining in $S \otimes \mathbb{C}$ when $\mathbf{x}$ is in the stratum S. The map Φ satisfies $\Phi(\mathbf{x}, 0) = \mathbf{x}$ and $(\mathrm{d}/\mathrm{d}t)\Phi(\mathbf{x}, t) = f(\Phi(\mathbf{x}, t))$. The fact that this may be defined up to time ϵ for some $\epsilon > 0$ is a consequence of the fact that the vector field f is bounded and that a small neighborhood of $\mathcal{M}$ in $\mathcal{M} \otimes \mathbb{C}$ is embedded in $\mathbb{C}^d$. Because f is smooth and bounded, for sufficiently small ϵ the map $\mathbf{x} \mapsto \Phi(\mathbf{x}, \epsilon)$ is a homeomorphism.

The flow reduces the real part of ϕ everywhere except at the critical points, which are rest points. Consequently, it defines a homotopy $H(\mathbf{x}, t) := \Phi(\mathbf{x}, \epsilon t)$ between the chain $\mathcal{C}$ representing $\mathcal{M}$ and a chain $\mathcal{C}'$ on which the minima of the real part of ϕ occur precisely on the set G. Recall that H induces a chain homotopy $\mathcal{C}_H$ with $\partial \mathcal{C}_H = \mathcal{C}' - \mathcal{C} + \partial \mathcal{C} \times \sigma$, where σ is a standard 1-simplex. Let ω denote the holomorphic d-form $A(\mathbf{z}) \exp(-\lambda \phi(\mathbf{z}))\, \mathrm{d}\mathbf{z}$. Because ω is a holomorphic d-form

in $\mathbb{C}^d$, we have $\mathrm{d}\omega = 0$. Now, by Stokes' Theorem,

$$\begin{aligned} 0 &= \int_{\mathcal{C}_H} \mathrm{d}\omega \\ &= \int_{\partial \mathcal{C}_H} \omega \\ &= \int_{\mathcal{C}'} \omega - \int_{\mathcal{C}} \omega - \int_{\partial \mathcal{C} \times \sigma} \omega. \end{aligned}$$

The chain $\partial\mathcal{C} \times \sigma$ is supported on a finite union of spaces $S \otimes \mathcal{C}$, where S is a stratum of dimension at most $d-1$. This chain is supported on a finite union of complex manifolds of dimension at most $d-1$; the integral of a holomorphic d-form vanishes over such a chain. Therefore, the last term on the right drops out and we have

$$\int_{\mathcal{C}} \omega = \int_{\mathcal{C}'} \omega.$$

Outside of a neighborhood of G the magnitude of the integrand is exponentially small, so we have shown that there are d-chains $\mathcal{C}_{\mathbf{x}}$ supported on arbitrarily small neighborhoods $\mathcal{N}(\mathbf{x})$ of each $\mathbf{x} \in G$ such that

$$\mathcal{I}(\lambda) - \sum_{\mathbf{x} \in G} \int_{\mathcal{C}_{\mathbf{x}}} \omega \tag{5.1}$$

is exponentially small. To finish that proof, we need only show that each $\int_{\mathcal{C}_{\mathbf{x}}} \omega$ has an asymptotic series in decreasing powers of λ whose leading term, when $A(\mathbf{x}) \neq 0$, is given by

$$c_0(\mathbf{x}) = (2\pi)^{d/2} A(\mathbf{x}) e^{\lambda \phi(\mathbf{x})} (\det \mathcal{H}(\mathbf{x}))^{-1/2}. \tag{5.2}$$

The d-chain $\mathcal{C}_{\mathbf{x}}$ may by parametrized by a map $\psi_{\mathbf{x}} \colon B \to \mathcal{N}(\mathbf{x})$, mapping the origin to $\mathbf{x}$, where B is the open unit ball in $\mathbb{R}^d$. By the chain rule,

$$\int_{\mathcal{C}_{\mathbf{x}}} \omega = \int_B [A \circ \psi](\mathbf{x}) \exp(-\lambda[\phi \circ \psi(\mathbf{x})]) \det \mathrm{d}\psi(\mathbf{x}) \, \mathrm{d}\mathbf{x}.$$

The real part of the analytic phase function $\phi \circ \psi$ has a strict minimum at the origin, so we may apply Theorem 4.1. We obtain an asymptotic expansion whose first term is

$$\left(\frac{2\pi}{\lambda}\right)^{d/2} [A \circ \psi](\mathbf{0}) \det \mathrm{d}\psi(\mathbf{0}) (\det M_{\mathbf{x}})^{-1/2}, \tag{5.3}$$

where $M_{\mathbf{x}}$ is the Hessian matrix of the function $\phi \circ \psi$. The term $[A \circ \psi](\mathbf{0})$ is equal to $A(\mathbf{x})$. The Hessian matrix of $\phi \circ \psi$ at the origin is given by $M_{\mathbf{x}} = \mathrm{d}\psi(\mathbf{0}) \, \mathcal{H}(\mathbf{x}) \, \mathrm{d}\psi(\mathbf{0})$. Thus

$$\det M_{\mathbf{x}} = (\det \mathrm{d}\psi(\mathbf{0}))^2 \det \mathcal{H}(\mathbf{x}),$$

and plugging into (5.3) yields (5.2). □

Proof of Corollary 2.4. Lemma 5.4 does not require the critical points to be in the interior, so the argument leading up to (5.1) is still valid. For those points $\mathbf{x}$ in a $(d-1)$-dimensional stratum, use Corollary 2.4 in place of Theorem 4.1 to obtain (5.2) with an extra factor of $1/2$. □

REMARK. The reason we do not continue with a litany of special geometries (quarter-spaces, octants, and so forth) is that the case of a half-space is somewhat special. The differential of the change of variables at the origin is a nonsingular map, which must send half-spaces to half-spaces, though it will in general alter angles of any smaller cone.

6. Examples

The simplest multidimensional application of our results is a computation from [**PW04**]. The purpose is to estimate coefficients of a class of bivariate generating functions whose denominator is the product of two smooth divisors. We give only a brief summary of how one arrives at (6.1) from a problem involving generating functions; a complete explanation of this can be found in [**PW04**, Section 4]. Note, however, that the mathematics of the integral is not contained in that paper, which instead refers to an earlier draft of this one!

Let v_1, v_2 be distinct analytic functions of z with $v_1(1) = v_2(1) = 1, 0 \neq v_1'(1) \neq v_2'(1) \neq 0$, and such that each $|v_i(z)|$ attains its maximum on $|z| = 1$ only at $z = 1$. For example, the last condition is satisfied by any pair of aperiodic power series with nonnegative coefficients and radius of convergence greater than 1.

Consider the generating function $F(z, w) = 1/H(z, w)$, where

$$H(z, w) = (1 - wv_1(z))\,(1 - wv_2(z))\,.$$

The two branches of the curve $H = 0$ intersect only at $(1, 1)$, and this intersection is transverse. The Maclaurin coefficients of $F(z, w) = \sum_{r,s} a_{rs} z^r w^s$ are given by the Cauchy integral formula

$$a_{rs} = \frac{1}{(2\pi\mathrm{i})^2} \int \frac{\mathrm{d}w\,\mathrm{d}z}{z^{r+1} w^{s+1}\,(1 - wv_1(z))\,(1 - wv_2(z))},$$

where the integral is taken over a product of circles centred at $(0, 0)$ and of sufficiently small radii.

Pushing the contour out to $|z| = 1, |w| = 1 - \varepsilon$ we obtain the same formula, since F is still analytic inside the product of disks bounded by these latter circles. Pushing the w-contour out to $|w| = 1 + \varepsilon$, using the residue formula on the inner integral and observing that the integral over $|w| = 1 + \varepsilon$ is exponentially decaying as $s \to \infty$, we see that

$$a_{rs} \approx \frac{1}{2\pi} \int_{|z|=1} \frac{-R_s(z)}{z^{r+1}}\,\mathrm{d}z,$$

where $\approx$ means that the difference is exponentially decaying as $s \to \infty$ and $R_s(z)$ denotes the sum of residues of $w \mapsto w^{-(s+1)} F(z, w)$ at the roots $w = 1/v_i(z)$, $i \in \{1, 2\}$.

The residue sum $R_s(z)$ can be rewritten in terms of an integral via

$$-R_s(z) = (s + 1) \int_0^1 [(1 - p)v_1 + pv_2]^s\,\mathrm{d}p,$$

and so we have

$$a_{rs} \approx \frac{s + 1}{2\pi} \int_{|z|=1} z^{-(r+1)} \int_0^1 [(1 - p)v_1(z) + pv_2(z)]^s\,\mathrm{d}p\,\mathrm{d}z.$$

In order to cast this into our standard framework, we need to be able to define a branch of the logarithm of $(1 - p)v_1(z) + pv_2(z)$. We do this by localizing on the

circle $|z| = 1$ to a sufficiently small neighbourhood of the point $z = 1$. This is possible since the integrand decays exponentially away from $z = 1$, by hypotheses on the v_i, and we shall show that the integral near $z = 1$ decays only polynomially.

The substitution $z = e^{\mathrm{i}t}$ converts this to an integral

$$a_{rs} \approx \frac{s+1}{2\pi} \int_{\mathcal{N}} \int_0^1 e^{-s\phi(p,t)} A(p,t)\, \mathrm{d}p\, \mathrm{d}t, \tag{6.1}$$

where $\phi(p,t) = \mathrm{i}r\theta/s + \log\left[(1-p)v_1(e^{\mathrm{i}t}) + pv_2(e^{\mathrm{i}t})\right]$, $A(p,t) = 1$, and $\mathcal{N}$ is a closed interval centred at 0. To compute asymptotics in the direction $r/s = \kappa$, for fixed $\kappa > 0$, we can consider ϕ to be independent of r and s.

We now asymptotically evaluate (6.1) using Theorem 2.3. We can rewrite the iterated integral as a single integral over the stratified space $\mathcal{M} = \mathcal{N} \times [0,1]$. The phase ϕ has nonnegative real part and this fits into our framework. There is a single stationary point, at $(p,z) = (1/2, 0)$ (note that $\mathrm{Re}\{\phi\}$ is zero for all $(p,0)$, so Theorem 4.1 does not suffice). This critical point is quadratically nondegenerate, and direct computation using Theorem 2.3 yields

$$a_{rs} = \frac{1}{|v_1'(1) - v_2'(1)|} + O(s^{-1}) \tag{6.2}$$

when $s \to \infty$ with κ fixed. By keeping track of error terms more explicitly, it is easily shown that this approximation is uniform in κ provided κ stays in a compact subset of the open interval formed by $v_1'(1), v_2'(1)$ (it follows from our assumptions that these numbers are positive real—see [**PW04**] for more details). This means that a_{rs} is asymptotically constant in any compact subcone of directions away from the boundary formed by lines of inverse slope $\kappa_i = v_i'(1)$.

This example, and in fact a number of cases in [**PW04**], can also be solved using iterated residues. This is carried out in [**BP04**]. Iterated residues have the advantage of showing that the $O(s^{-1})$ term decays exponentially, but the disadvantage that they do not give any results when κ approaches the boundary. The present methods do give boundary results. Corollary 2.4 shows that a_{rs} converges to one-half the right-hand side of (6.2) when $(r,s) \to \infty$ with $r/s = \kappa_1 + O(1)$, and a small extension yields a Gaussian limit: letting Φ denote the standard normal cumulative distribution function, we have

$$a_{rs} = \frac{\Phi(u)}{|v_1'(1) - v_2'(1)|} + O(s^{-1})$$

when $r, s \to \infty$ with $(r/s - \kappa_1)/s^{1/2} \to u$.

7. Further topics

Higher order terms. We have not emphasized explicit formulae for the higher order terms, giving an equation such as (2.1) only for the leading term in the case where $A(\mathbf{0}) \neq 0$. However, our results establish the validity of existing computations of higher order terms under our more general hypotheses.

To elaborate, we prove Theorem 2.3 by first constructing a change of variables $\mathbf{x} \mapsto \Phi(x, \epsilon)$ homotopic to the identity under which the minimum of $\mathrm{Re}\{\phi\}$ at $\mathbf{0}$ is strict, and then changing variables, again homotopically to the identity, to the standard form. The composition ψ of these two maps is homotopic to the identity

but is far from explicitly given: while the second map is constructed by an explicitly defined Morse function, the first deformation is the solution to a differential equation and is not particularly explicit.

In [**Hör90**], Hörmander derives such an explicit formula (assuming smoothness) for integrals of our type where $\mathcal{M} = \mathbb{R}^d$ and A has compact support. The formula is indeed rewritten and used in [**RW08**] to compute higher order terms for generating function applications, in which more restrictive hypotheses preclude the vanishing of $\mathrm{Re}\{\phi\}$ on a curve reaching the boundary of the chain of integration. Their methods, while not covering the cases of interest here, do have the virtue of dealing with the change of variables ψ only through the equation $S = \phi \circ \psi$. In particular, the derivatives of ψ arising in the computation of the new amplitude function $(A \circ \psi) \det \mathrm{d}\psi$ can be computed by implicitly differentiating the equation $S = \phi \circ \psi$. Having found at least one such ψ homotopic to the identity, we are now free to replicate the computations of [**RW08**] under our more general hypotheses, as follows.

In the case of standard phase, the coefficient of $\lambda^{-(n+d)/2}$ is given (provided that all r_i are even) by

$$\sum_{|\mathbf{r}|=n} a_{\mathbf{r}} \beta_{\mathbf{r}},$$

where $a_{\mathbf{r}}$ is the Maclaurin coefficient of A corresponding to the monomial $\mathbf{r}$ and $\beta_{\mathbf{r}}$ is the constant defined in Corollary 3.2. Note that n must be even for this coefficient to be nonzero, so we write $n = 2k$. The differential operator $\partial_1^{r_1} \cdots \partial_d^{r_d}$ when applied to A and evaluated at $\mathbf{0}$ yields precisely $\prod_i r_i! a_{\mathbf{r}}$. Thus the operator

$$\sum_{|\mathbf{r}|=k} \frac{\partial_1^{2r_1} \cdots \partial_d^{2r_d}}{4^k \, r_1! \cdots r_d!}$$

applied to A and evaluated at zero yields the coefficient we seek.

After the Morse lemma is applied using the change of variables $S = \phi \circ \psi$, we need to apply the displayed operator to the new amplitude $(A \circ \psi) \det \mathrm{d}\psi$. The resulting expression evaluated at $\mathbf{x}$ can be computed directly via the rules of Leibniz and Faà di Bruno. Evaluating at $\mathbf{x}$ simplifies some terms, and, as mentioned above, derivatives of $(A \circ \psi) \det \mathrm{d}\psi$ may be computed without explicitly specifying ψ.

As a relatively simple example, consider the case $k = 1$ and $d = 1$. The differential operator reduces to $\frac{1}{4}\partial^2$, where ∂ denotes differentiation with respect to the variable x. Applying this to $(A \circ \psi) \det \mathrm{d}\psi$ yields (with superscripts denoting the order of derivatives and arguments suppressed)

$$\frac{1}{4}\left(A^{(2)}\left(\psi^{(1)}\right)^3 + 3A^{(1)}\psi^{(1)}\psi^{(2)} + A^{(0)}\psi^{(3)}\right).$$

The defining equation $S = \phi \circ \psi$ can be differentiated to yield the system

$$\begin{aligned}
2x &= \phi^{(1)}\psi^{(1)}, \\
2 &= \phi^{(2)}\left[\psi^{(1)}\right]^2 + \phi^{(1)}\psi^{(2)}, \\
0 &= \phi^{(3)}\left[\psi^{(1)}\right]^3 + 3\phi^{(2)}\psi^{(1)}\psi^{(2)} + \phi^{(1)}\psi^{(3)}, \\
0 &= \phi^{(4)}\left[\psi^{(1)}\right]^4 + 6\phi^{(3)}\left[\psi^{(1)}\right]^2\psi^{(2)} + 4\phi^{(2)}\psi^{(1)}\psi^{(3)} + 3\phi^{(2)}\left[\psi^{(2)}\right]^2 + \phi^{(1)}\psi^{(4)}.
\end{aligned}$$

Evaluating these at the point in question, we see that the terms with highest derivatives of ψ vanish in each equation. The system is triangular and can be solved explicitly to obtain

$$\psi^{(1)} = \sqrt{\frac{2}{\phi^{(2)}}},$$
$$\psi^{(2)} = \frac{-2\phi^{(3)}}{3\left[\phi^{(2)}\right]^2},$$
$$\psi^{(3)} = \frac{\left[5\phi^{(3)}\right]^2 - 3\phi^{(2)}\phi^{(4)}}{3\sqrt{2}\left[\phi^{(2)}\right]^{7/2}}.$$

Putting these together with the expression for the derivative of $(A\circ\psi)\det \mathrm{d}\psi$ above yields an expression for the $\lambda^{-3/2}$ term in the integral that is a rational function with denominator $\left[\phi^{(2)}\right]^{7/2}$, and with numerator equal to a polynomial in the derivatives of A up to order 2, and of ϕ up to order 4. In summary, the results of this paper show that the computational apparatus and formulae for higher order terms given in [**RW08**] hold in the case of complex phase functions integrated over stratified spaces.

Relation to existing literature. As stated at the outset, the aim of the present article has been to outline results that will be useful to the combinatorics community concerning asymptotics of multidimensional integrals with complex phase. Existing treatments may be classified as belonging to one of two types.

Those of the first type are easily accessible but not sufficiently general. These include standard references such as [**BH86, Won89**], which may be found in engineering libraries as well as mathematics libraries. Also in this class are [**Bre94**] and [**dB81**]. These treatments are self-contained, rigorous, and assume little knowledge beyond standard real and complex analysis. These sources treat real phases and purely imaginary phases but do not treat complex phases and do not use inherently complex methods. The treatment by Stein [**Ste93**], which was written as background for the study of differential operators, is similarly limited (purely imaginary phase, C^∞ methods). The book by Paris and Kaminski [**PK01**] contains great detail on Laplace integrals in dimensions up to three, but has very little on non-real phases. These are the sources typically cited by combinatorialists. When multidimensional contour methods are required, combinatorialists will either reduce somehow to a one-dimensional case, e.g. [**BFSS00**], or work out the results from scratch, e.g., [**Wor04**]. Perhaps the closest we have found to an off-the-shelf usable source is [**Hör90**]. This work is self-contained, considers general complex phases, and is conducive to explicit computation. However, the methods are strictly C^∞, and the amplitude function is assumed to be compactly supported in $\mathbb{R}^d$.

References of the second type are sufficiently high-powered to do everything we need. However, from our point of view they suffer from a number of drawbacks. Most noticeable is the significant overhead required to make use of these sources. The complex algebraic geometry contained in the work of Pham [**Pha85**] (extending ideas of [**Ler50**]) and Malgrange [**Mal74**], and later [**Vas77, Fed77**] and the Arnol'd school [**Var77, AGZV88**], is well beyond the comfort zone of most of the

combinatorial community. Furthermore, the results we need, while undeniably implicit in these works, are not always explicitly stated, or easy to find and identify. In some cases, proofs are absent as well [**AGZV88, Fed89**].

It should be noted that the approach taken in these treatments is in some sense the right approach. It is the forerunner to the constructions of [**GM88**], taking the viewpoint of the Thom–Mather theory. We draw heavily on these ideas in the way we frame and prove our results. The Morse-theoretic deformations we use to prove the lemmas in Section 5 lead, for example, to Pham's "thimbles of Lefschetz." Had we been able to find the exact result we needed in any of these sources, we would have greatly preferred to cite it than to reprove it. However, each of these sources comes with hypotheses that make sense in the context it was written in, but which are too restrictive for our application. These include restrictions on the form of the phase function, the nature of the domain of integration, and specifically the behavior of the integrand on the boundary of the domain. The work of Howls [**BH91, DH02**] deserves mention because it is more accessible than the other works mentioned in the last two paragraphs. Howls' aim is close to ours: algorithmic understanding of the asymptotics of integrals. These papers emphasize hyperasymptotic theory (asymptotic terms beyond the leading exponential order), which involves non-local classification of the chain of integration and requires added geometric assumptions on the domain of integration.

Acknowledgements. We are indebted to three referees for educating us on the existing literature and for insightful comments on the exposition and the proofs of Theorem 2.3 and Lemma 5.4.

References

[AGZV88] V. I. Arnol'd, S. M. Guseĭn-Zade, and A. N. Varchenko, *Monodromy and Asymptotics of Integrals*, vol. II of *Singularities of Differentiable Maps*, Birkhäuser, Boston, MA, 1988. Translated from the Russian by Hugh Porteous, translation revised by the authors and James Montaldi.

[BBBP08] Y. Baryshnikov, W. Brady, A. Bressler, and R. Pemantle, *Two-dimensional quantum random walk*, preprint, 34 pp., 2008. Available as arXiv:0810.5495.

[BFSS00] C. Banderier, Ph. Flajolet, G. Schaeffer, and M. Soria, *Planar maps and Airy phenomena*, pp. 388–402 in *Automata, Languages and Programming (Geneva, 2000)*, no. 1853 in Lecture Notes in Computer Science, Springer-Verlag, New York/Berlin, 2000.

[BH86] N. Bleistein and R. A. Handelsman, *Asymptotic Expansions of Integrals*, Dover, New York, 2nd edition, 1986.

[BH91] M. Berry and C. Howls, Hyperasymptotics for integrals with saddles, Proc. Roy. Soc. London Ser. A **434** (1991), 657–675.

[BM93] A. Bertozzi and J. McKenna, *Multidimensional residues, generating functions, and their application to queueing networks*, SIAM Rev. **35** (1993), 239–268.

[BP04] Y. Baryshnikov and R. Pemantle, *Convolutions of inverse linear functions via multivariate residues*, preprint, 48 pp., 2004.

[BP08] Y. Baryshnikov and R. Pemantle, *Asymptotics of quadratic sequences, III. Quadratic points*, preprint, 82 pp., 2008. Available as arXiv:0810.4898.

[Bre94] Karl Breitung. *Asymptotic Approximations for Probability Integrals*, vol. 1592 of *Lecture Notes in Mathematics*, Springer-Verlag, New York/Berlin, 1994.

[dB81] N. de Bruijn, *Asymptotic Methods in Analysis*, Dover, New York, 3rd (corrected reprint) edition, 1981.

[DH02] E. Delabaere and C. Howls, *Global asymptotics for multiple integrals with boundaries*, Duke Math. J. **112** (2002), 199–264.

[Fed77] M. V. Fedoryuk, *Saddle Point Method* (in Russian), Nauka, Moscow, 1977.

[Fed89] M. V. Fedoryuk, *Asymptotic methods in analysis*, pp. 83–191 in *Integral Representations and Asymptotic Methods*, vol. 13 of The Encyclopedia of Mathematical Sciences, Springer-Verlag, New York/Berlin, 1989.

[GM88] M. Goresky and R. MacPherson, *Stratified Morse Theory*, no. 14 in Ergebnisse der Mathematik und ihrer Grenzgebiete (3), Springer-Verlag, New York/Berlin, 1988.

[Hör90] Lars Hörmander. *An Introduction to Complex Analysis in Several Variables*, 3rd edition, North-Holland, Amsterdam, 1990.

[Ler50] J. Leray. *Le calcul différentiel et intégral sur un variété analytique complexe*, Bull. Soc. Math. France **87** (1950), 81–180.

[Mal74] B. Malgrange, *Intégrales asymptotiques et monodromie*, Ann. Sci. École Norm. Sup. (4) **7** (1974), 405–430.

[Pha85] Fr. Pham, *La descente des cols par les onglets de Lefschetz, avec vues sur Gauss–Manin*, Asterisque **130** (1985), 11–47.

[PK01] R. Paris and D. Kaminski, *Asymptotics and Mellin–Barnes Integrals*, vol. 85 of The Encyclopedia of Mathematics and its Applications, Cambridge Univ. Press, Cambridge, UK, 2001.

[PW02] R. Pemantle and M. C. Wilson, *Asymptotics of multivariate sequences, I. Smooth points of the singular variety*, J. Combin. Theory Ser. A **97** (2002), 129–161.

[PW04] R. Pemantle and M. C. Wilson, *Asymptotics of multivariate sequences, II. Multiple points of the singular variety*, Combin. Probab. Comput. **13** (2004), 735–761.

[PW08] R. Pemantle and M. C. Wilson, *Twenty combinatorial examples of asymptotics derived from multivariate generating functions*, SIAM Rev. **50** (2008), 199–272.

[RW08] A. Raichev and M. C. Wilson, *Asymptotics of coefficients of multivariate generating functions: Improvements for smooth points*, Electron. J. Combin. **15** (2008), no. 1, research paper 89, 17 pp.

[Ste93] Elias M. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, vol. III of Monographs in Harmonic Analysis, Princeton Univ. Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy.

[Var77] A. N. Varchenko. *Newton polyhedra and estimation of oscillating integrals*, Funct. Anal. Appl. **10** (1977), 175–196.

[Vas77] V. Vasiliev. *Asymptotic exponential integrals, Newton's diagram, and the classification of minimal points*, Funct. Anal. Appl. **11** (1977), 163–172.

[Wat01] S. Watanabe. *Algebraic analysis for nonidentifiable learning machines*, Neural Computation **13** (2001), 899–933.

[Won89] R. Wong, *Asymptotic Approximations of Integrals*, Academic Press, Boston, MA, 1989.

[Wor04] N. Wormald, *Tournaments with many Hamilton cycles*, preprint, 2004.

Department of Mathematics, University of Pennsylvania, 209 S. 33rd Street, Philadelphia, PA 19104, USA

E-mail address: pemantle@math.upenn.edu

URL: http://www.math.upenn.edu/~pemantle

Department of Computer Science, University of Auckland, Private Bag 92019, Auckland, New Zealand

E-mail address: mcw@cs.auckland.ac.nz

URL: http://www.cs.auckland.ac.nz/~mcw/blog

This volume contains the proceedings of the AMS Special Sessions on Algorithmic Probability and Combinatorics held at DePaul University on October 5–6, 2007 and at the University of British Columbia on October 4–5, 2008.

This volume collects cutting-edge research and expository on algorithmic probability and combinatorics. It includes contributions by well-established experts and younger researchers who use generating functions, algebraic and probabilistic methods as well as asymptotic analysis on a daily basis. Walks in the quarter-plane and random walks (quantum, rotor and self-avoiding), permutation tableaux, and random permutations are considered. In addition, articles in the volume present a variety of saddle-point and geometric methods for the asymptotic analysis of the coefficients of single- and multi-variable generating functions associated with combinatorial objects and discrete random structures. The volume should appeal to pure and applied mathematicians, as well as mathematical physicists; in particular, anyone interested in computational aspects of probability, combinatorics and enumeration. Furthermore, the expository or partly expository papers included in this volume should serve as an entry point to this literature not only to experts in other areas, but also to graduate students.

ISBN 978-0-8218-4783-1

9 780821 847831

CONM/520