

# Computer Algebra for Lattice Path Combinatorics

ALIN BOSTAN

Résumé par BRUNO SALVY

*Séminaire de Combinatoire Philippe Flajolet  
Institut Henri Poincaré, Séance du 6 juin 2013*

## Abstract

Classifying lattice walks in restricted lattices is an important problem in enumerative combinatorics. Recently, computer algebra methods have been used to explore and solve a number of difficult questions related to lattice walks. This talk gave an overview of recent results on structural properties and explicit formulas for generating functions of walks in the quarter plane, with an emphasis on the algorithmic methodology.

## 1 Introduction

Lattice path enumeration is a classical topic of combinatorics, with applications in probability theory (branching processes, gambler's ruin, . . .), queuing theory, discrete mathematics (permutations, words, trees, urns, . . .) and statistical physics (percolation, Ising model, . . .).

The work reported in this talk concentrates on the enumeration of nearest-neighbor walks in the quarter plane  $\mathbb{N}^2$ . A set  $\mathfrak{S}$  of elements of  $\{-1, 0, 1\}^2 \setminus \{(0, 0)\}$  is given and one considers walks starting at  $(0, 0)$ , remaining in  $\mathbb{N}^2$  and using only steps in  $\mathfrak{S}$ . From the enumeration viewpoint, the quantity of interest is  $f_{\mathfrak{S}}(n; i, j)$ , the number of such walks ending at  $(i, j)$  after  $n$  steps, and the corresponding generating series

$$F_{\mathfrak{S}}(t; x, y) = \sum_{n=0}^{\infty} \left( \sum_{i,j=0}^{\infty} f_{\mathfrak{S}}(n; i, j) x^i y^j \right) t^n.$$

Specializing  $x$  and  $y$  to 0 gives the walks returning to the origin (called *excursions*), while specializing them both to 1 gives the total number of walks of length  $n$ . Another meaningful specialization is  $x = 1, y = 0$ , which gives the generating series of walks ending on the horizontal axis (the vertical one is obtained by  $x = 0, y = 1$ ).

The sequence  $f_{\mathfrak{S}}(n; i, j)$  is easily seen to obey a linear recurrence with constant coefficients, but its behavior is much more complicated than for solutions of such recurrences in dimension 1. Instead of a simple linear equation for the generating series, one obtains a functional equation called the *kernel equation* relating  $F_{\mathfrak{S}}(t; x, y)$  and its specializations  $F_{\mathfrak{S}}(t; 0, y)$ ,  $F_{\mathfrak{S}}(t; x, 0)$ ,  $F_{\mathfrak{S}}(t; 0, 0)$ . An exact “formula” for  $f_{\mathfrak{S}}(n; i, j)$  or for  $F_{\mathfrak{S}}(t; x, y)$  is not available in general, but a lot has been discovered in the past few years concerning the nature of these sequences and series. Computer algebra plays an essential role in this study, where a large number of coefficients are used to first produce reasonable conjectures and then rigorous proofs, some of which do not have a purely human version yet.

## 2 Varieties of Power Series

The simplest family of power series consists of polynomials. These occur rarely as generating series of walks: only the polynomial 1 can occur, for trivial walks whose steps all have a negative coordinate. Next in complexity, rational generating series occur when the step set does not have any element with a negative coordinate. For these walks, staying in  $\mathbb{N}^2$  is not a constraint. Rational power series form a subclass of algebraic power series, that are solutions of a polynomial. It has been known for a long time that all walks whose step sets belong to the right half-plane have an algebraic generating series. They are not the only ones. The generating series of the Kreweras walks, with step set  $\{(0, -1), (-1, 0), (1, 1)\}$  has been known to be algebraic for almost 30 years [8, 6]. More recently, the more difficult case of Gessel walks, with step set  $\{(1, 1), (-1, -1), (-1, 0), (1, 0)\}$  has been proved to have an algebraic generating series [2]. Algebraic power series are themselves a special case of D-finite power series, i.e., solutions of linear differential equations with polynomial coefficients. An interesting subclass of D-finite power series is composed of hypergeometric series (where the ratio of two consecutive coefficients is a fixed rational function of the index) and their integrals, that occur very frequently in this application too, see §5 below.

## 3 The group of the walk

To the step set  $\mathfrak{S}$  is associated the characteristic polynomial of the walk

$$\chi_{\mathfrak{S}} = \sum_{(i,j) \in \mathfrak{S}} x^i y^j = \sum_{i=-1}^1 B_i(y) x^i = \sum_{j=-1}^1 A_j(x) y^j.$$

This polynomial is left invariant under the substitutions  $\psi(x, y) = (x, A_{-1}(x)/A_1(x)/y)$  and  $\phi(x, y) = (B_{-1}(y)/B_1(y)/x, y)$ . The *group of the walk* is the group of substitutions generated by  $\psi$  and  $\phi$ . It seems to play an important role in the classification of the possible behaviors of the generating series.

After removing trivial cases and taking symmetry into account, there are 79 inherently different cases of small-step sets [7]. These are now completely classified as follows:

- 23 admit a finite group [13]. Of these
  - 4 are algebraic (Gessel’s and variants of Kreweras’);
  - 19 are D-finite and transcendental [9, 5].
- 56 admit an infinite group [7]. Moreover,
  - 5 singular cases are non-D-finite [14, 12];
  - the other 51 cases are non-D-finite by other arguments [4, 11].

## 4 Computer-aided proofs of algebraicity

The discovery that Gessel’s walks have an algebraic generating series was based on experimental mathematics. It is a very nice story that deserves more than a few lines and we refer to the original article for details [2]. With hindsight it can be obtained by computing a large number of coefficients of the series and then computing a Hermite-Padé approximant for the series and its first powers. Because of the size of these objects, the computation are better performed

modulo primes, concentrating on the specializations at  $x = 0$  or  $y = 0$  and for given values of the remaining variable.

Once polynomials annihilating  $F_{\mathfrak{S}}(t; x, 0)$  and  $F_{\mathfrak{S}}(t; 0, y)$  have been conjectured, they can be proved automatically by showing first that these polynomials has a unique power series solution, next (by resultant computations using fast algorithms) that these power series have to satisfy equations derived from the kernel equation and finally that these derived equations themselves have a unique formal power series solution. This establishes the algebraicity of both generating series. From there, the kernel equation implies that the full generating series  $F_{\mathfrak{S}}(t; x, y)$  is algebraic too, which concludes the proof. The first purely human proof of this result has appeared very recently [3].

The size of the minimal polynomial annihilating the generating function of the Gessel walks is estimated at 30 gigabytes. An interesting extra structure of the generating function has been discovered and proved with computer algebra more recently, resulting in a 5 lines-long formula for the generating series. No human proof of this “simple” formula is known yet.

## 5 Computer-aided proofs of D-finiteness

The same experimental mathematics approach can be used to determine and prove linear differential equations satisfied by the 19 non-algebraic D-finite cases corresponding to finite groups. More precise information on the structure of these series is obtained by starting from the kernel equation. Letting the group act on this equation, an expression of the generating series as the positive part of a rational Laurent series can be obtained [7]. By known properties of D-finite series, this proves D-finiteness and the computer algebra technique of creative telescoping produces the desired linear differential equation. A more thorough analysis of this equation using factorization of linear differential operators leads to the discovery of even more structure: it turns out that all those 19 cases have a generating series that can be expressed using integrals of  ${}_2F_1$  series [1]. A nice example is provided by the “King’s walk”, where all 8 steps are allowed, for which the generating function of all walks (ie the value at  $x = y = 1$ ) is given explicitly as

$$\frac{1}{t} \int_0^t \frac{1}{(1+4x)^3} {}_2F_1 \left( \frac{3}{2} \quad \frac{3}{2} \middle| \frac{16x(1+x)}{(1+4x)^2} \right) dx.$$

Again, no purely human proof is known at this stage.

## 6 Computer-aided proofs of non-D-finiteness

The proof that a generating series does not satisfy a linear differential equation is usually based on exhibiting a property of the series that is not compatible with it being D-finite. Thus, a recent result in probability theory giving access to the asymptotic behavior of the integer sequence  $f_{\mathfrak{S}}(n; 0, 0)$  leads to a proof of non-D-finiteness provided an exponent in the expansion is proved irrational. This is decided algorithmically, using Gröbner basis computations, polynomial factorization and cyclotomy testing [4]. The non-D-finiteness of  $F_{\mathfrak{S}}(t; 0, 0)$  implies that of the more general  $F_{\mathfrak{S}}(t; x, y)$ , for which an alternative proof is known [11]. No human proof is known for the specialization at  $(0,0)$  and the nature of the specialization at  $(1,1)$  is still unknown.

## Conclusion

The classification of the 2D walks with small-step sets has been completed in the past few years thanks in a part to a very efficient toolbox coming from computer algebra. Much remains to be

done. Several of the results or observations still do not have a purely human proof. The exact role played by the group of the walk is still unclear. The non-D-finite nature of  $F_{\mathfrak{S}}(t; 1, 1)$  in the non-singular cases with infinite group is still only a conjecture. Some of these may become clearer with work in progress studying longer 2D steps or higher dimensional walks.

## References

- [1] BOSTAN, A., CHYZAK, F., VAN HOEIJ, M., KAUERS, M., AND PECH, L. Explicit differentially finite generating functions of walks with small steps in the quarter plane. In preparation (2013).
- [2] BOSTAN, A., AND KAUERS, M. The complete generating function for Gessel walks is algebraic. *Proceedings of the American Mathematical Society* 138, 9 (2010), 3063–3078.
- [3] BOSTAN, A., KURKOVA, I., AND RASCHEL, K. A human proof of Gessel’s lattice path conjecture, 2013. Preprint.
- [4] BOSTAN, A., RASCHEL, K., AND SALVY, B. Non-D-finite excursions in the quarter plane. Tech. rep., arXiv, 2013. To appear in the Journal of Combinatorial Theory, Series A.
- [5] BOUSQUET-MÉLOU, M. Counting walks in the quarter plane. In *Mathematics and computer science, II (Versailles, 2002)*, Trends Math. Birkhäuser, Basel, 2002, pp. 49–67.
- [6] BOUSQUET-MÉLOU, M. Algebraic generating functions in enumerative combinatorics and context-free languages. In *STACS 2005: 22nd Annual Symposium on Theoretical Aspects of Computer Science (2005)*, V. Diekert and B. Durand, Eds., vol. 3404 of *Lecture Notes in Computer Science*, Springer-Verlag GmbH, pp. 18–35.
- [7] BOUSQUET-MÉLOU, M., AND MISHNA, M. Walks with small steps in the quarter plane. In *Algorithmic probability and combinatorics*, vol. 520 of *Contemp. Math.* Amer. Math. Soc., Providence, RI, 2010, pp. 1–39.
- [8] GESSEL, I. M. A probabilistic method for lattice path enumeration. *Journal of Statistical Planning and Inference* 14, 1 (1986), 49 – 58.
- [9] GESSEL, I. M., AND ZEILBERGER, D. Random walk in a Weyl chamber. *Proceedings of the American Mathematical Society* 115, 1 (May 1992), 27–31.
- [10] KAUERS, M., KOUTSCHAN, C., AND ZEILBERGER, D. Proof of Ira Gessel’s lattice path conjecture. *Proceedings of the National Academy of Sciences of the USA* 106, 28 (2009), 11502–11505.
- [11] KURKOVA, I., AND RASCHEL, K. On the functions counting walks with small steps in the quarter plane. *Publ. Math. Inst. Hautes Études Sci.* 116, 1 (2012), 69–114.
- [12] MELCZER, S., AND MISHNA, M. Singularity analysis via the iterated kernel method. In *Proceedings of FPSAC’13, Paris, France (2013)*.
- [13] MISHNA, M. Automatic enumeration of regular objects. *Journal of Integer Sequences* 10, 07.5.5 (May 2007), 1–18.
- [14] MISHNA, M., AND RECHNITZER, A. Two non-holonomic lattice walks in the quarter plane. *Theoretical Computer Science* 410, 38-40 (2009), 3616–3630.