AUTOMATIC CLASSIFICATION OF RESTRICTED LATTICE WALKS

ALIN BOSTAN* AND MANUEL KAUERS†

Abstract. We propose an *experimental mathematics approach* leading to the computer-driven *discovery* of various structural properties of general counting functions coming from enumeration of walks.

1. Introduction

There is a strange phenomenon about the generating functions that count lattice walks restricted to the quarter plane: depending on the choice of the set $\mathfrak{S} \subseteq \{\swarrow, \leftarrow, \nwarrow, \uparrow, \nearrow, \rightarrow, \searrow, \downarrow\}$ of admissible steps, the generating function may be rational, algebraic [but not rational], D-finite [but not algebraic], or not even D-finite. Much progress was made recently on understanding why this is so, and only a couple of days ago, Bousquet-Melou and Mishna [8] have announced a classification of all the 256 possible step sets into algebraic, transcendental D-finite, and non-D-finite cases, together with proofs for the algebraic and D-finite cases and strong evidence supporting the conjectured non-D-finiteness of the others.

As usual, a power series $S(t) \in \mathbb{Q}[[t]]$ is called *algebraic* if there exists a bivariate polynomial $P(T, t) \in \mathbb{Q}[T, t]$ such that $P(S(t), t) = 0$, and transcendental otherwise. Also as usual, a power series $S(t)$ is called D-finite if it satisfies a linear differential equation with polynomial coefficients. (Every algebraic power series is D-finite, but not vice versa.) At first glance, it might seem easy to prove that a power series is algebraic or D-finite: just come up with an appropriate equation, and then verify that the series satisfies this equation. But as far as lattice walks are concerned, most proofs given so far are *indirect* in the sense that they only imply that an equation *exists,* while the explicit equation remains unknown. This is probably so because the equations appearing in this context are often too big to be dealt with by hand.

Nevertheless, it is interesting to know the equations explicitly, because they provide a standard canonical representation for a series, from which lots of further information can be extracted in a straightforward manner. By applying a well-known technique from computer algebra (in modern fashion, cf. Section 2), we have systematically searched for differential equations and algebraic equations that the series counting the number of walks in the quarter plane satisfy. These are given in Section 3. We have also made a first step towards classifying walks in $\mathbb{Z}^3$ confined to the first octant (cf. Section 4) by considering all step sets $\mathfrak{S}$ with up to five elements, and performed a systematic search for equations of the corresponding series. More than 2000 hours of computation time have been spent in order to analyze about 3500 different sequences.

We do not provide proofs that the equations we found are indeed correct, but the computational evidence in favor of our equations is striking. We have no doubt

* Partially supported by the French National Agency for Research (ANR Project "Gecko") and the Microsoft Research-INRIA Joint Centre.

† Partially supported by the Austrian Science Foundation (FWF) grants P19462-N18 and P20162-N18.

that all the equations we found are correct. In principle, it would be possible to supplement the "automatically guessed" equations by computer proofs in a systematic fashion, using techniques that have recently been applied to isolated step sets [22, 21, 5]. But we found the computational cost for performing these computations more than one hundred times too high.

2. Methodology

To study generating functions for lattice walks, we follow a classical scheme in experimental mathematics. It is based on the following steps: (S1) computation of high order expansions of generating power series; (S2) guessing differential and/or algebraic equations satisfied by those power series; (S3) empirical certification of the guessed equations (sieving by inspection of their analytic, algebraic and arithmetic properties); (S4) rigourous proof, based on (exact) polynomial computations.

In the sequel, we only explain Steps (S1), (S2) and (S3). A full description of Step (S4) is given in [5]. By way of illustration, we choose an example requiring computations with human-sized outputs, namely the classical case, initially considered by Kreweras [24, 6, 7], of walks in the quarter plane restricted to the step set $\mathfrak{S} = \{\leftarrow, \nearrow, \downarrow\}$.

2.1. **Basic Definitions and Facts.** We focus on 2D and 3D lattice walks. The 2D walks that we consider are confined to the quarter plane $\mathbb{N}^2$, they join the origin of $\mathbb{N}^2$ to an arbitrary point $(i, j) \in \mathbb{N}^2$, and are restricted to a fixed step set $\mathfrak{S} \subseteq \{\swarrow, \leftarrow, \nwarrow, \uparrow, \nearrow, \rightarrow, \searrow, \downarrow\}$. If $f(n; i, j)$ denotes the number of such walks of length n (i.e., using n steps chosen from $\mathfrak{S}$), the sequence $f(n; i, j)$ satisfies the multivariate recurrence with constant coefficients

$$f(n+1; i, j) = \sum_{(h,k)\in\mathfrak{S}} f(n; i-h, j-k) \quad \text{for} \quad n, i, j \geq 0. \tag{1}$$

Together with the appropriate boundary conditions

$$f(0; 0, 0) = 1 \quad \text{and} \quad f(n; i, j) = 0 \quad \text{if} \quad i < 0 \quad \text{or} \quad j < 0 \quad \text{or} \quad n < 0,$$

the recurrence relation (1) uniquely determines the sequence $f(n; i, j)$. As is customary in combinatorics, let

$$F(t; x, y) = \sum_{n\geq 0} \Big(\sum_{i,j\geq 0} f(n; i, j) x^i y^j \Big) t^n$$

be the trivariate generating power series of the sequence $f(n; i, j)$. As $f(n; i, j) = 0$ as soon as $i > n$ or $j > n$, the inner sum is actually finite, and so we may regard $F(t; x, y)$ as a formal power series in t with polynomial coefficients in $\mathbb{Q}[x, y]$.

Specializing $F(t; x, y)$ to selected values of x and y leads to various combinatorial interpretations. Setting $x = y = 1$ yields the power series $F(t; 1, 1)$ whose coefficients count the total number of walks with prescribed number of steps (and arbitrary endpoint); the choice $x = y = 0$ gives the series $F(t; 0, 0)$ whose coefficients count the number of walks returning to the origin; setting $x = 1$, $y = 0$ yields the power series $F(t; 1, 0)$ whose coefficients count the number of walks ending somewhere on the vertical axis, etc.

By [9, Th. 7], multivariate sequences that satisfy recurrences with constant coefficients have moderate growth, and thus their generating series are analytic in the complex plane. The next theorem refines this result in our context.

Theorem 1. *The following inequality holds*

$$f(n; i, j) \leq |\mathfrak{S}|^n \quad \textit{for all} \quad (i, j, n) \in \mathbb{N}^3. \tag{2}$$

In particular, the power series $F(t;0,0), F(t;1,0), F(t;0,1)$ *and* $F(t;1,1)$ *are convergent in* $\mathbb{C}[[t]]$ *at* $t=0$ *and their radius of convergence is at most* $1/|\mathfrak{S}|$.

Proof. We proceed by induction on n. By the boundary conditions, the inequality (2) is obviously true for $n=0$. Now, the recurrence (1) implies that if (2) holds for n, then it also holds for $n+1$:

$$f(n+1;i,j) = \sum_{(h,k)\in\mathfrak{S}} f(n;i-h,j-k) \leq \sum_{(h,k)\in\mathfrak{S}} |\mathfrak{S}|^n = |\mathfrak{S}|^{n+1}.$$

This readily implies that the radius of convergence of $F(t;0,0)$ at least $1/|\mathfrak{S}|$. Similarly we have the estimate

$$[t^n]F(t;1,1) = \sum_{i=0}^{n}\sum_{j=0}^{n} f(n;i,j) \leq \sum_{i=0}^{n}\sum_{j=0}^{n} |\mathfrak{S}|^n = (n+1)^2|\mathfrak{S}|^n$$

which, by the Hadamard rule, implies that $F(t;1,0)$, $F(t;0,1)$ and $F(t;1,1)$ have also radii of convergence that are at least $1/|\mathfrak{S}|$. □

2.1.1. *Generating series of walks are G-functions.* A power series $S(t) = \sum_{n\geq 0} a_n t^n$ in $\mathbb{Q}[[t]]$ is called a G-function[1] if (a) it is D-finite; (b) its radius of convergence in $\mathbb{C}[[t]]$ is positive; (c) there exists a constant $C>0$ such that for all $n \in \mathbb{N}$, the common denominator of $a_0, \ldots, a_n$ is bounded by C^n.

Examples of G-functions are the series $\log(1-t)$ and $(1-t)^\alpha$ for $\alpha \in \mathbb{Q}$. More generally, the Gauss hypergeometric series ${}_2F_1(\alpha,\beta,\gamma;t)$ with rational parameters α,β,γ, is also a G-series. A celebrated theorem of Eisenstein assures that any algebraic power series must be a G-function (if S is algebraic, there exists an integer $C \in \mathbb{N}$ such that $a_n C^{n+1}$ is an integer for all n.) The fact that G-functions arise frequently in combinatorics was recently pointed out by Garoufalidis [18].

Theorem 2. *Let* $S(t)$ *be one of the power series* $F(t;0,0), F(t;1,0), F(t;0,1)$ *and* $F(t;1,1)$. *If* S *is D-finite, then* S *is a* G*-series. In particular, its minimal order homogeneous linear differential equation is Fuchsian and it has only rational exponents. Moreover, the coefficient sequence of* $S(t)$ *is asymptotically equal to a term of the form* $\kappa\rho^n n^\alpha (\log n)^\beta$ *for some constants* $\kappa \in \mathbb{R}$, $\alpha \in \mathbb{Q}$, $\rho \in \bar{\mathbb{Q}}$, *and* $\beta \in \mathbb{N}$.

Proof. The conditions (a) are (c) in the definition are clearly satisfied. The only non-trivial point is the fact that the series S has a positive radius of convergence in $\mathbb{C}$. This follows from Theorem 1. The last assertion follows from a combination of results of Katz and Honda [20, 19], and Chudnovsky [12]; one of the intermediate steps is the fact that the minimal order operator of a G-series must be *globally nilpotent*, see Section 2.4.4 below for the definition and an algorithmic use. By a Theorem of Katz and Honda, the global nilpotence of a differential operator implies that all of its singular points are regular singular points with rational exponents. See also [1, 15, 11] for more details on this topic. The minimal equation for $S(t)$ being Fuchsian with only rational exponents implies the claim on the asymptotics of its coefficients [17]. □

For 3D lattice walks, the definitions are analogous. The trivariate power series $F(t;x,y)$ is simply replaced by the generating series $G(t;x,y,z) \in \mathbb{Q}[x,y,z][[t]]$ of the sequence $g(n;i,j,k)$ that counts walks in $\mathbb{N}^3$ starting at $(0,0,0)$ and ending at $(i,j,k) \in \mathbb{N}^3$. Note that the appropriate versions of Theorems 1 and 2 hold; in particular, the generating series of total walks $G(t;1,1,1)$ is a G-series.

[1]The usual definition is more general, the coefficients of S can be taken in an arbitrary algebraic number field. For our purposes it is sufficient and convenient to restrict to rational coefficients.

2.2. **Computing large series expansions.** The recurrence (1) can be used to determine the value of $f(n;i,j)$ for specific integers $n,i,j\in\mathbb{N}$. Theorem 1 implies that $f(n;i,j)$ is a positive integer whose bit size is at most $O(n)$. If $N\in\mathbb{N}$, the values $f(n;i,j)$ for $0\le n,i,j\le N$ can thus be computed altogether by a straightforward algorithm that uses $O(N^3)$ arithmetic operations and $\tilde{O}(N^4)$ bit operations. (We assume that two integers of bit-size N can be multiplied in $\tilde{O}(N)$ bit operations; here, the softO notation $\tilde{O}(\ )$ hides logarithmic factors.) The memory storage requirement is proportional to N^3. The same is also true for the truncated power series $F_N=F(t;x,y) \bmod t^N$. For our experiments in 2D, we have chosen $N=1000$. With this choice, the computation of the $f(n;i,j)$ is the step which consumes by far the most computation time in our calculations.[2]

Example 1. *The Kreweras walks satisfy the recurrence*

$$f(n+1,i,j)=f(n,i+1,j)+f(n,i,j+1)+f(n,i-1,j-1) \quad \textit{for} \quad n,i,j\ge 0,$$

which allows the computation of the first terms of the series $F(t;x,y)$

$$\begin{aligned}F(t;x,y)&=1+xyt+(x^2y^2+y+x)t^2+(x^3y^3+2xy^2+2x^2y+2)t^3\\&+(x^4y^4+3x^2y^3+3x^3y^2+2y^2+6xy+2x^2)t^4\\&+(x^5y^5+4x^3y^4+4x^4y^3+5xy^3+12x^2y^2+5x^3y+8y+8x)t^5+\cdots\end{aligned}$$

and also the first terms of the generating series $F_{1,1}(t)$ *for the total number of Kreweras walks*

$$\begin{aligned}F_{1,1}(t)\ &=1+t+3t^2+7t^3+17t^4+47t^5+125t^6+333t^7+939t^8+2597t^9+\\&7183t^{10}+20505t^{11}+57859t^{12}+163201t^{13}+469795t^{14}+O(t^{15}).\end{aligned}$$

In the 3D case, the values $g(n;i,j,k)$ for $0\le n,i,j,k\le N$ can be computed in $O(N^4)$ arithmetic operations, $\tilde{O}(N^5)$ bit operations and $O(N^4)$ memory space. In practice, we found that computing $G \bmod t^N$ with $N=400$ is feasible.

2.3. **Guessing.** Once the first terms of a power series are determined, our approach is to search systematically for candidates of linear differential equations or of algebraic equations which the series may possibly satisfy. This technique is classical in computer algebra and mathematical physics, see for example [10, 28, 25]. Differential and algebraic guessing procedures are available in some computer algebra systems like `Maple` and `Mathematica`.

2.3.1. *Differential guessing.* If the first N terms of a power series $S\in\mathbb{Q}[[t]]$ are available, one can search for a differential equation satisfied by S at precision N, that is, for an element $\mathcal{L}$ in the Weyl algebra $\mathbb{Q}[t]\langle D_t\rangle$ of differential operators in the derivation $D_t=\frac{d}{dt}$ with polynomial coefficients in t, such that

$$\mathcal{L}(S)=c_r(t)S^{(r)}(t)+\cdots+c_1(t)S'(t)+c_0(t)S(t)=0 \bmod t^N. \tag{3}$$

Here, the coefficients $c_i(t)\in\mathbb{Q}[t]$ are not simultaneously zero, and their degrees are bounded by a prescribed integer d. By a simple linear algebra argument, if d and r are chosen such that $(d+1)(r+1)>N$, such a differential equation always exists. On the other side, if d,r and N are such that $(d+1)(r+1)\ll N$, the equation (3) translates into a highly over-determined linear system, so it has no reason to possess a non-trivial solution.

The idea is that if the given power series $S(t)$ happens to be D-finite, then for a sufficiently large N, a differential equation of type (3) (thus satisfied a priori only at

[2]We have carried out our computations on various different machines whose main memory ranges from 8 Gb to 32 Gb and which are equipped with (multiple) processors all running at about 3Ghz.

precision N) will provide a differential equation which is really satisfied by $S(t)$ in $\mathbb{Q}[[t]]$ (i.e., at precision infinity). In other words, the D-finiteness of a power series can be (conjecturally) recognized using a finite amount of information.

Given the values d, r, N, and the first N terms of the series S, a candidate differential equation of type (3) for S can be computed by Gaussian elimination in $O(N^3)$ arithmetic operations and $\tilde{O}(N^4)$ bit operations. Actually, a modular approach is preferred to a direct Gaussian elimination over $\mathbb{Q}$. Precisely, the linear algebra step is performed modulo several primes p, and the results (differential operators modulo p) are recombined over $\mathbb{Q}$ via rational reconstruction based on an effective version of the Chinese remainder theorem.

If no differential equation is found, this definitely rules out the possibility that a differential equation of order r and degree d exists. This does not, however, imply that the series at hand is not D-finite. It may still be that the series satisfies a differential equation or higher order than r or an equation with polynomial coefficients exceeding r.

Asymptotically more efficient algorithms exist, based on fast Hermite-Padé approximation [3] of the vector of (truncated) power series $[S, S', \ldots, S^{(r)}]$; they have arithmetic complexity quadratic or even softly-linear in N. Such sophisticated algorithms were not needed to obtain the results of this paper, but they have provided crucial help in the treatment of examples of critical sizes (e.g. guessing with higher values of d, r, N or/and over a parametric base field like $\mathbb{Q}(x)$ instead of $\mathbb{Q}$) needed for the proof step (S4), see [5].

Example 2 (continued)**.** *$N = 100$ terms of the generating series $F_{1,1}(t) = F(t; 1, 1)$ of the total Kreweras walks are sufficient to conjecture that $F_{1,1}(t)$ is D-finite, since it verifies the differential equation $\mathcal{L}_{1,1}(F_{1,1}) = 0 \bmod t^N$, where*

$$
\begin{aligned}
\mathcal{L}_{1,1} = {} & 4t^2(t+1)(3t-4)(3t-1)^3(9t^2+3t+1)D_t^4 \\
& + 2t(3t-1)^2(2916t^5 - 1296t^4 - 3564t^3 - 477t^2 - 93t + 52)D_t^3 \\
& + 3(3t-1)(29808t^6 - 26244t^5 - 28440t^4 + 2754t^3 + 431t^2 + 448t - 40)D_t^2 \\
& + 6(68040t^6 - 88452t^5 - 37206t^4 + 16758t^3 + 954t^2 + 253t - 126)D_t \\
& + 18(6480t^5 - 8856t^4 - 3078t^3 + 714t^2 + 211t + 2).
\end{aligned}
$$

Thus, with high probability, $F_{1,1}$ verifies the differential equation $\mathcal{L}_{1,1}(F_{1,1}) = 0$.

Sometimes (e.g. see Section 2.4.4 below) we are willing to guess the minimal-order differential equation $\mathcal{L}_{\min}(S) = 0$ satisfied by the given generating power series. Most of the time, the choice (d, r) of the target degree and order does not lead to this minimal operator $\mathcal{L}_{\min}$. Worse, it may even happen that the number of initial terms N is not large enough to recover $\mathcal{L}_{\min}$, while these N terms suffice to guess non-minimal order operators. The explanation of why such a situation occurs systematically was given in full detail in [4], for the case of differential equations satisfied by algebraic functions. A good heuristic in this case is to compute several non-minimal operators and to take their right greatest common divisor; generically, the result is exactly $\mathcal{L}_{\min}$.

As a final general remark, let us point out that a power series satisfies a linear differential equation if and only if its coefficients satisfy a linear recurrence equation with polynomial coefficients. A recurrence equation can be computed either from a differential equation, or it can be guessed from scratch by proceeding analogously as described above for differential equations.

Example 3 (continued). *$N = 100$ terms of the series $S(t) = F_{1,1}(t)$ suffice to guess that they satisfy the order-6 recurrence*

$$\begin{aligned}
&- 972(n+1)(n+2)(n+4)(7n+41)u_n \\
&- 324(n+2)(7n^3+90n^2+382n+545)u_{n+1} \\
&+ 108(n+3)(35n^3+443n^2+1787n+2309)u_{n+2} \\
&- 18(n+4)(28n^3+311n^2+897n+304)u_{n+3} \\
&+ 3(28n^4+626n^3+5123n^2+18281n+24070)u_{n+4} \\
&- (n+6)(140n^3+2402n^2+13687n+25843)u_{n+5} \\
&+ 2(n+6)(n+7)(2n+13)(7n+34)u_{n+6} = 0.
\end{aligned}$$

2.3.2. *Algebraic guessing.* If the first N terms of a power series $S \in \mathbb{Q}[[t]]$ are available, one can also search for an algebraic equation satisfied by S at precision N, that is, for a bivariate polynomial $P(T, t)$ in $\mathbb{Q}[T, t]$ such that

$$P(S) = c_r(t)S(t)^r + \cdots + c_0(t)S(t) = 0 \bmod t^N. \tag{4}$$

A similar discussion shows that candidates algebraic equations of type (4) for S can be "guessed" by performing either Gaussian elimination or Hermite-Padé approximation on the vector $[S, S^2, \ldots, S^r]$, followed by a gcd computation in $\mathbb{Q}[T, t]$.

Example 4 (continued). *$N = 100$ terms of the series $S(t) = F_{1,1}(t)$ counting the total Kreweras walks suffice to guess that $F_{1,1}$ is very probably algebraic, namely solution of the bivariate polynomial*

$$\begin{aligned}
P_{1,1}(T,t) \; &= t^5(3t-1)^3T^6 + 6t^4(3t-1)^3T^5 + t^3(3t-1)(135t^2-78t+14)T^4 \\
&\quad + 4t^2(3t-1)(45t^2-18t+4)T^3 + t(3t-1)(135t^2-26t+9)T^2 \\
&\quad + 2(3t-1)(27t^2-2t+1)T + 43t^2 + t + 2.
\end{aligned}$$

2.4. **Empirical certification of guesses.** Once discovered a differential equation (3) or an algebraic equation (4) that the power series $S(t)$ seem to satisfy, we inspect several properties of these equations, in order to provide more convincing evidence that they are correct. These properties have various natures: some are computational features (moderate bit sizes), other are algebraic, analytic and even arithmetic properties. We check them systematically on all the candidates; if they are verified, as in the Kreweras example, this offers striking evidence that the guessed equations are not artefacts.

2.4.1. *Size sieve: Reasonable bit size.* The differential equation for $S(t)$ has typically much lower bit size than a differential equation produced by the same guessing procedure applied to the same order, degree and precision, but to an arbitrary series having coefficients of bit-size comparable to that of $S(t)$. A similar observation holds for the algebraic equation (4).

Example 5 (continued). *If we perturb the coefficients of $S(t) = F(t; 1, 1)$ by just adding 1 to each of its coefficients, then the differential guessing procedures at order $r = 4$, degree $d = 9$ and precision $N = 100$ will either give no result (overdetermined system approach) or produce fake candidates (the Hermite-Padé approach) with polynomial coefficients in t, whose coefficients in $\mathbb{Q}$ have numerators and denominators of about 500 decimal digits each, instead of 4 digits for $\mathcal{L}_{1,1}$.*

2.4.2. *Algebraic sieve: High order series matching.* The equations (3) and (4) were obtained starting from N coefficients of the power series $S(t)$. They are therefore satisfied a priori only modulo t^N. We compute more terms of $S(t)$, say $2N$, and check whether the same equations still hold modulo t^{2N}. If this is the case, chances increase that the guessed equations even hold at infinite precision.

2.4.3. *Analytic sieve: singularity analysis.* By Theorem 2, the minimal order operator for power series like $S(t) = F(t; 0, 0)$ and $S(t) = F(t; 1, 1)$ must have only regular singularities (including the point at infinity) and their exponents must be rational numbers.

Example 6 (continued)**.** *The differential operator $\mathcal{L}_{1,1}$ is Fuchsian. Indeed, a (fully automated) local singularity analysis shows that the set of its singular points*

$$\left\{-1, 0, \infty, \tfrac{1}{3}, \tfrac{4}{3}, -\tfrac{1}{6}(1 \pm i\sqrt{3})\right\}$$

is formed uniquely of regular singularities. Moreover, the indicial polynomials are, respectively: $t(t-1)(t-2)(2t-1), t(t-1)(2t+1)(t+1), (t-5)(t-1)(t-2)(t-4), (t+1)t(4t-1)(4t+1), t(t-1)(t-2)(t-4)$, and $t(t-2)(2t-3)(t-1)$.

2.4.4. *Arithmetic sieve: G-series and global nilpotence.* Last, but not least, we check an arithmetic property of the guessed differential equations, by exploiting the fact that those expected to arise in our combinatorial context are very special.

Indeed, by a Theorem due to the Chudnovsky brothers [12], the minimal order differential operator $\mathcal{L} \in \mathbb{Q}\langle t, D_t \rangle$ killing a G-series enjoys a remarkable arithmetic property: $\mathcal{L}$ is *globally nilpotent*. By definition, this means that for almost every prime number p (i.e., for all with finitely many exceptions), there exist an integer μ such that the remainder of the Euclidian (right) division of $D_t^{p\mu}$ by $\mathcal{L}$ is congruent to zero modulo p [19, 14].

From a computational view-point, a fine feature is that the nilpotence modulo p is checkable. If r denotes the order of $\mathcal{L}$, let M be the *p-curvature matrix of* $\mathcal{L}$, defined as the $r \times r$ matrix with entries in $\mathbb{Q}(t)$, whose (i, j) entry is the coefficient of t^{j-1} of the remainder of the Euclidian (right) division of D_t^{p+i-1} by $\mathcal{L}$. Then, $\mathcal{L}$ is globally nilpotent if, for almost all primes p, the matrix M is nilpotent modulo p [14, 29].

In combination with Theorem 2, this yields a fast algorithmic filter: as soon as we guess a candidate differential equation satisfied by a G-series (e.g. by $F(t; 1, 1)$), we check whether its p-curvature is nilpotent, say modulo the first 50 primes for which the reduced operator $\mathcal{L} \bmod p$ is well-defined. If the p-curvature matrix of $\mathcal{L}$ is nilpotent modulo p for all those primes p, then the guessed equation is, with very high probability, the correct one.

We push even further this arithmetic sieving. A famous conjecture, attributed to Grothendieck, says that the differential equation $\mathcal{L}(S) = 0$ possesses a basis of *algebraic solutions* (over $\mathbb{Q}(x)$) if and only if its p-curvature *is zero* for almost all primes p. Even if the conjecture is, for the moment, fully proved only for order one operators and partially in the other cases [11], we freely use it as an oracle to detect whether a guessed differential equation has a basis of algebraic solutions. For instance, the computation of the p-curvature of an order 11 differential operator with polynomial coefficients of degree 96 in t, was one of the key points in our discovery [5] the that fully (trivariate) generating function for Gessel walks is algebraic.

Example 7 (continued)**.** *The 5-curvature matrix $M(t)$ of the differential operator $\mathcal{L}_{1,1}$ has the form $\frac{1}{d(t)}\tilde{M}(t)$, where $d(t) = (3t-1)^7 t^6 (t+1)^5 (9t^2+3t+1)^5 (3t-4)$ and $\tilde{M}(t)$ is a 4×4 polynomial matrix whose entries are polynomials in $\mathbb{Q}[t]$ of degree at most 27. The characteristic polynomial of M reads*

$$\begin{aligned}\chi_M(T, t) \;&= \frac{1}{D(t)} \cdot \Bigg(T^4 + \frac{3 \cdot 5}{2^5}\, N_3(t)\, t^5\, (3t-1)^{10}\, T^3 + \\ &+ \frac{3^3 \cdot 5}{2^{10}}\, N_2(t)\, (3t-1)^5\, T^2 + \frac{3^5 \cdot 5^2 \cdot 7}{2^7}\, N_1(t)\, T + \frac{3^9 \cdot 5^3 \cdot 7^2}{2^3}\, N_0(t) \Bigg),\end{aligned}$$

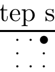
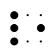
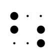
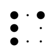
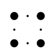
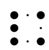
where $D(t) = (3t-1)^{15}t^{10}(t+1)^5(9t^2+3t+1)^5(3t-4)$ *and where* N_0, N_1, N_2, N_3 *are irreducible polynomials in* $\mathbb{Z}[t]$*, of degree, respectively,* $21, 26, 26, 21$ *and with coefficients having at most 20 decimal digits.*

The polynomial χ_M *obviously equals* T^4 *modulo* $p = 5$*, so the* 5*-curvature of* $\mathcal{L}_{1,1}$ *is nilpotent (but not zero*[3]*) modulo* 5*. In fact, for all the primes* $7 \leq p < 100$*, the p-curvature matrix of* $\mathcal{L}_{1,1}$ *is also nilpotent modulo p; it is even zero modulo p. Under the assumption that Grothendieck's conjecture is true, this indicates that* $\mathcal{L}_{1,1}$ *admits a basis of algebraic solutions, and so provides independent evidence that also* $S(t) = F(t;1,1)$ *is algebraic.*

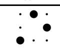
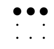
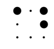
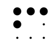
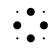
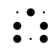
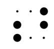
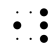
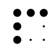
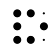
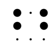
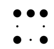
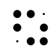
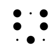
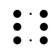
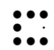
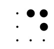
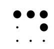
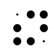
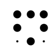
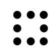
3. Empirical Results in 2D

3.1. **The Classification.** The table below contains the results of our computations for walks in the plane. We consider the total number of walks only, i.e., the generating function $F(t;1,1)$. Because of symmetries, the 256 possible step sets give rise to 92 different sequences only. Among those, we found that the following ones are D-finite, based on an inspection of the first $N = 1000$ terms.

The second column contains a diagram depicting one step set whose sequence starts as indicated in the first column (bold dots indicate arrow tips, e.g., the step set in the 8th row is Kreweras'). In the remaining columns, we give the size of equations we found for the sequence: differential equation of the power series, recurrence equation of the sequence itself, and, if applicable, algebraic equation of the series. For example, for Kreweras (row 8), there is a differential equation of order 4 with polynomial coefficients of degree 9 (the operator $\mathcal{L}_{1,1}$ given in Section 2.3.1), a recurrence equation of order 6 with polynomial coefficients of degree 4 (also in Section 2.3.1), and the power series satisfies a polynomial $P_{1,1}(T,t)$ of degree 8 with respect to T and degree 6 with respect to t (given in Section 2.3.2).

Sequence	Step set	Equation sizes		
1, 1, 1, 1, 1, 1, 1, 1		1, 1	1, 0	1, 1
1, 1, 2, 3, 6, 10, 20, 35		2, 3	2, 1	2, 2
1, 1, 2, 3, 8, 15, 39, 77		4, 16	6, 8	–
1, 1, 2, 4, 9, 21, 51, 127		2, 3	2, 1	2, 2
1, 1, 3, 5, 15, 29, 87, 181		3, 4	3, 1	2, 2
1, 1, 3, 5, 17, 34, 121, 265		4, 15	5, 8	–
1, 1, 3, 6, 20, 50, 175, 490		3, 5	2, 3	–
1, 1, 3, 7, 17, 47, 125, 333		4, 9	6, 4	8, 6
1, 1, 3, 7, 19, 49, 139, 379		5, 16	7, 10	–
1, 1, 4, 7, 28, 58, 232, 523		3, 4	3, 1	2, 2
1, 1, 4, 9, 36, 100, 400, 1225		3, 5	2, 3	–
1, 1, 5, 13, 61, 199, 939, 3389		5, 24	9, 18	–
1, 2, 4, 8, 16, 32, 64, 128		1, 1	1, 0	1, 1

[3]Modulo 5, the 5-curvature matrix $M(t)$ has T^2 as minimal polynomial.

Sequence	Step set	Equation sizes		
1, 2, 4, 10, 26, 66, 178, 488		4, 12	7, 4	8, 6
1, 2, 5, 13, 35, 96, 267, 750		2, 3	2, 1	2, 2
1, 2, 6, 16, 48, 136, 408, 1184		3, 4	3, 1	2, 2
1, 2, 6, 18, 58, 190, 638, 2170		3, 4	3, 1	2, 2
1, 2, 6, 18, 60, 200, 700, 2450		3, 4	2, 2	–
1, 2, 6, 21, 76, 290, 1148, 4627		5, 19	7, 11	–
1, 2, 7, 21, 78, 260, 988, 3458		3, 5	2, 3	9, 8
1, 2, 7, 23, 84, 301, 1127, 4186		5, 15	6, 10	–
1, 2, 7, 23, 85, 314, 1207, 4682		3, 4	3, 1	2, 2
1, 2, 7, 26, 105, 444, 1944, 8728		5, 18	7, 11	–
1, 2, 8, 24, 96, 320, 1280, 4480		2, 3	2, 1	2, 2
1, 2, 8, 29, 129, 535, 2467, 10844		5, 24	9, 18	–
1, 2, 8, 32, 144, 672, 3264, 16256		2, 3	2, 1	2, 2
1, 2, 9, 34, 151, 659, 2999, 13714		5, 20	8, 15	–
1, 2, 10, 39, 210, 960, 5340, 26250		3, 8	4, 5	–
1, 2, 11, 49, 277, 1479, 8679, 49974		5, 24	9, 18	–
1, 3, 9, 27, 81, 243, 729, 2187		1, 1	1, 0	1, 1
1, 3, 11, 41, 157, 607, 2367, 9277		3, 4	3, 1	2, 2
1, 3, 13, 55, 249, 1131, 5253, 24543		2, 3	2, 1	2, 2
1, 3, 14, 67, 342, 1790, 9580, 52035		2, 3	2, 1	4, 4
1, 3, 15, 74, 392, 2116, 11652, 64967		5, 18	7, 14	–
1, 3, 16, 86, 509, 3065, 19088, 120401		5, 24	9, 18	–
1, 3, 18, 105, 684, 4550, 31340, 219555		3, 6	3, 4	–

3.2. **Observations.**

3.2.1. Our classification matches the results of Bousquet-Melou and Mishna [8]: for every sequences they prove D-finite our software found a recurrence and a differential equation, and whenever a series is algebraic indeed, our software recognized it. Moreoever, we found no recurrence or differential equation for any step set conjectured non-D-finite by Bousquet-Melou and Mishna. This strengthens the evidence in favor of the conjectured non-D-finiteness of these cases.

3.2.2. All but two of the the minimal polynomials of the algebraic series share the property that they define a curve of genus 0. As a consequence, there exists a rational parametrization in all these cases. For example, for the Kreweras step set

, the minimal polynomial $P_{1,1}$ given in Section 2.3.2 defines a curve which is parameterized by

$$T = \frac{(u^2+24u+151)a(u)}{(u+9)(u^2+24u+147)}, \quad t = \frac{2}{a(u)},$$

where $a(u) = \big(u^6+66u^5+1827u^4+27180u^3+229431u^2+1042866u+1995717\big)/(u+11)(u^2+22u+125)^2$. The two algebraic series which do not admit a rational parametrization belong to the step sets  (reverse Kreweras) and  (Gessel's). Their genus is 1.

Another feature of the series which we found to be algebraic is that they all admit closed forms in terms of radical expressions. For example, for the Kreweras step set, we find that $F(t;1,1)$ is equal to

$$-\frac{1}{t} + \sqrt{\frac{\left(i-\sqrt{3}\right)\left(216t^3+1\right)\left(t-3t^2\right)^2 - 2it(36t^2-15t+1)a(t) + \left(i+\sqrt{3}\right)a(t)^2}{6it^3(3t-1)^3a(t)}}$$

where $i = \sqrt{-1}$ and

$$a(t) = \sqrt[3]{24\sqrt{3t^9(3t-1)^9\left(9t^2+3t+1\right)^3} - t^3(3t-1)^3\left(5832t^6+540t^3-1\right)}.$$

Both features are remarkable because among all algebraic power series, those which are rationally parameterizable or expressible in terms of radicals form a set of measure zero.

3.2.3. Also the transcendental D-finite series appear to have some remarkable features. Being D-finite, these series are annihilated by some linear differential operator

$$\mathcal{L} = c_0(t) + c_1(t)D_t + \cdots + c_r(t)D_t^r \in \mathbb{Q}(t)\langle D_t\rangle.$$

According to `Maple`, all the operators can be factorized into a product of one irreducible operator of order 2 and zero or more operators of order 1. As all the operators are globally nilpotent, so are all their factors [14, 15].

We can therefore expect that every solution of these factors can be written as a product of an algebraic function (a fractional power of a rational function) with a hypergeometric ${}_2F_1$ composed with a rational function (such an expression is called a *weak pullback*). Indeed, Dwork [14] has conjectured that any globally nilpotent second order differential equation has either algebraic solutions or is a weak pullback of the Gauss hypergeometric differential equation; this conjecture was disproved by Krammer [23] and recently by Dettweiler and Reiter [13], the counter-examples given in these papers require involved tools in algebraic geometry (systems associated to periods of Shimura curves, ...)

We are therefore in a win-win situation: either all order 2 operators appearing as factors of our operators are indeed weak-pullbacks of ${}_2F_1$ functions, or there is a simple combinatorial counter-example to Dwork's conjecture.

Let us illustrate this at one of the most simple examples, the step set . We find here the differential operator

$$4(32t^2-12t-1) + 4(8t-1)(20t^2-3t-1)D_t \\ + t(4t-1)(112t^2-5)D_t^2 + t^2(4t-1)^2(4t+1)D_t^3$$

which Maple factors into

$$\big(2(192t^3-56t^2-6t+1) + 4(24t^2)(4t-1)tD_t + (4t-1)^2(4t+1)t^2D_t^2\big)\big(1/t+D_t\big)$$

Again using to Maple, this factorization gives rise to the representation

$$F(t;1,1) = -\frac{1}{4t} + \Big(1+\frac{1}{4t}\Big){}_2F_1\left(\begin{matrix}1/2 & 1/2\\ & 1\end{matrix}\,\middle|\, 16t^2\right).$$

(Incidentally, this solution can also be expressed in terms of elliptic functions.) We believe that all the transcendental D-finite generating functions for any step admit a representation as (a nested integral of) such an expression. The solvers of Maple and Mathematica, however, are able to discover such a representation only in the most simple cases. (Note that at present, no complete algorithm is known that is capable of finding general pullback representations.)

3.2.4. Recall that all the coefficient sequencess grow like $\kappa n^\alpha \rho^n \log(n)^\beta$ for some constants $\kappa, \rho, \alpha, \beta$. From the differential equation or the recurrence equation, we can determine ρ, α, and β exactly as roots of characteristic polynomials and indicial equations, respectively. (See [30, 17] on how this is done.) We find that $\beta = 0$ in all cases. Knowing the recurrence, we can also compute easily tens of thousands of sequence terms. With the help of convergence acceleration techniques applied to so many terms, it is possible to determine the remaining constant κ to an accuracy of thirty digits or more. With that many digits, it makes sense to search systematically for potential exact expressions of these constants using Plouffe's inverter [27] and/or algorithms like LLL and PSLQ [2]. We could find "closed form" expressions for all these constants.

By Theorem 1, the numbers ρ are bounded by the cardinality of the step set $\mathfrak{S}$. It turns out that $\rho = |\mathfrak{S}|$ unless the vector sum of the elements of the step set points outside the first quadrant. In these cases, ρ is an algebraic numbers of degree 2 (e.g., $\rho = 1 + 2\sqrt{2}$ for the step set ). For α, we find only negative numbers. Note that α being a negative integer implies that the corresponding series is transcendental [16].

All the constants κ have the form

$$u\rho^{e_0}\phi_1^{e_1}\phi_2^{e_2}\cdots\phi_r^{e_r}$$

where the ϕ_i are usually small integers, the e_i are rational numbers, and u is $1/\pi$ in the case of transcendental sequences or $1/\Gamma(\alpha+1)$ in the case of algebraic sequences. There are some cases where the ϕ_i are not integers. For the step set , we have $\rho = 1 + 2\sqrt{2}$ and $\phi_1 = 1 + \sqrt{2}$. For the step set , we have $\rho = 1 + 2\sqrt{3}$ and $\phi_1 = 1 + \sqrt{3}$. Very strange is only the step set , where we found $\rho = 2 + 2\sqrt{6}$ and

$$\phi_1 = \sqrt{17693 + 7223\sqrt{6}}.$$

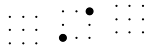
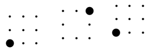
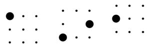
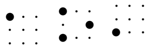
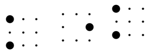
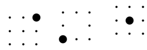
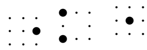
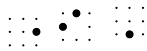
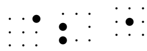
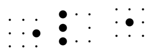
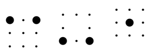
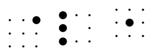
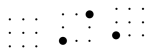
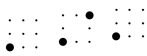
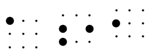
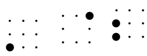
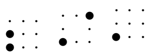
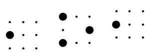
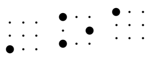
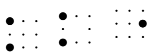
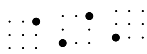
This number may look like a guessing artefact at first glance, but we trust in its correctness, because the number of correct digits exceeds by far the number of correct digits to be expected from an artefact. It is possible that the above number can be written as fractional powers of "simpler" numbers, but we did not succeed in finding such a rewriting.

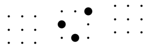
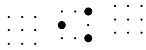
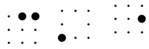
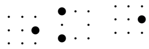
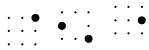
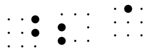
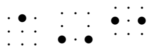
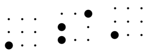
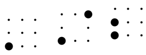
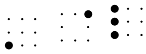
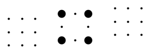
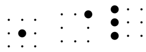
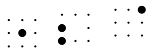
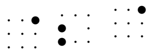
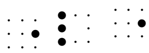
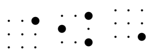
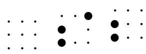
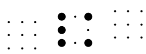
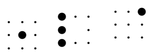
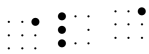
4. Empirical Results in 3D

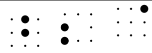
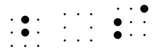
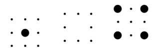
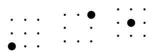
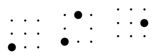
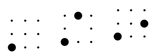
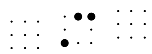
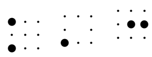
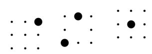
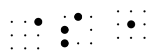
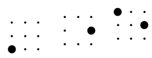
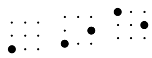
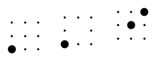
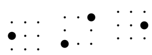
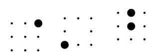
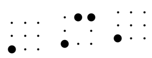
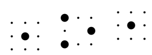
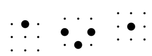
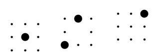
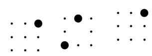
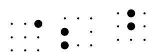
4.1. **The Classification.** The table below contains the results of our computations for walks in three dimensions confined to the first octand with step sets of up to five elements. Again, we restrict the attention to the total number of walks.

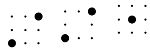
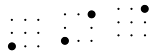
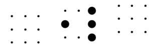
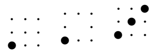
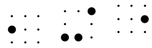
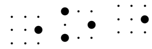
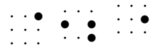
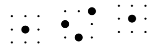
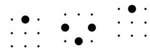
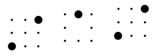
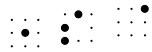
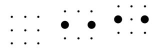
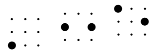
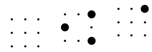
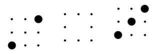
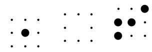
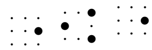
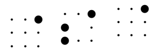
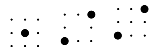
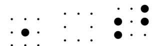
A priori, there are 83682 such step sets, and they give rise to 3334 different sequences. Among those, we found that the following 134 sequences are D-finite, based on an inspection of the first $N = 400$ terms.

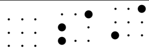
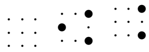
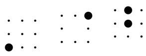
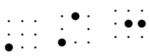
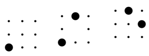
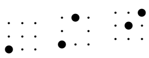
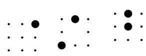
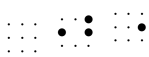
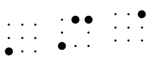
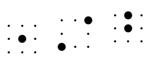
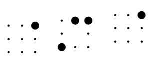
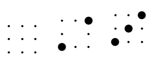
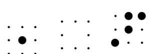
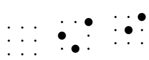
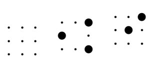
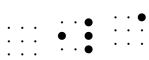
The numbers in the table have the same meaning as in the table in Section 3.1. The three dimensional step set is depicted in three separate slices: first the arrows tops of the forms $(x, y, -1)$, then $(x, y, 0)$, then $(x, y, 1)$. For example, the step set depicted in row four is $\{(-1, 1, -1),\ (-1, -1, 0),\ (1, 0, 0),\ (-1, 0, 1)\}$.

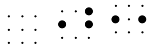
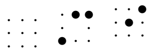
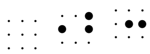
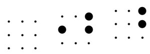
Sequence	Step set	Equation sizes		
1, 1, 1, 1, 1, 1, 1, 1		1, 1	1, 0	1, 1
1, 1, 2, 3, 6, 10, 20, 35		2, 3	2, 1	2, 2
1, 1, 2, 3, 8, 15, 39, 77		4, 16	6, 8	–
1, 1, 2, 3, 8, 15, 44, 91		4, 15	5, 8	–
1, 1, 2, 3, 10, 20, 58, 119		5, 28	9, 17	–
1, 1, 2, 3, 10, 20, 63, 133		6, 27	7, 18	–
1, 1, 2, 3, 12, 25, 77, 161		5, 28	9, 17	–
1, 1, 2, 3, 12, 25, 87, 189		6, 25	7, 17	–
1, 1, 2, 4, 9, 21, 51, 127		2, 3	2, 1	2, 2
1, 1, 2, 4, 9, 21, 56, 148		10, 69	12, 57	–
1, 1, 2, 4, 10, 25, 70, 196		3, 4	2, 2	–
1, 1, 2, 4, 11, 31, 91, 267		3, 6	4, 3	–
1, 1, 2, 5, 12, 32, 92, 261		5, 16	6, 10	–
1, 1, 2, 5, 12, 32, 97, 282		10, 71	14, 57	–
1, 1, 2, 5, 14, 42, 137, 464		9, 53	11, 43	–
1, 1, 2, 6, 15, 43, 143, 437		5, 17	7, 10	–
1, 1, 3, 5, 15, 29, 87, 181		3, 4	3, 1	2, 2
1, 1, 3, 5, 17, 34, 121, 265		4, 15	5, 8	–
1, 1, 3, 5, 17, 34, 126, 279		4, 11	4, 6	–
1, 1, 3, 5, 19, 39, 145, 321		5, 28	9, 17	–
1, 1, 3, 5, 19, 39, 155, 349		5, 28	9, 17	–
1, 1, 3, 5, 19, 39, 155, 349		6, 25	7, 16	–
1, 1, 3, 5, 21, 44, 179, 405		6, 28	7, 19	–
1, 1, 3, 5, 23, 49, 203, 461		4, 16	6, 8	–
1, 1, 3, 6, 20, 50, 175, 490		3, 5	2, 3	–

Sequence	Step set	Equation sizes		
1, 1, 3, 7, 17, 47, 125, 333		4, 9	6, 4	8, 6
1, 1, 3, 7, 19, 49, 139, 379		5, 16	7, 10	–
1, 1, 3, 7, 21, 61, 191, 603		2, 3	2, 1	2, 2
1, 1, 3, 7, 23, 64, 223, 687		8, 31	8, 25	–
1, 1, 3, 7, 23, 71, 246, 848		3, 6	4, 3	–
1, 1, 3, 8, 25, 77, 257, 853		5, 18	6, 12	–
1, 1, 3, 9, 27, 93, 335, 1193		5, 17	7, 10	–
1, 1, 3, 9, 28, 100, 365, 1365		4, 11	5, 5	–
1, 1, 4, 7, 28, 58, 232, 523		3, 4	3, 1	2, 2
1, 1, 4, 7, 30, 63, 281, 649		4, 11	4, 6	–
1, 1, 4, 7, 32, 68, 320, 747		5, 28	9, 17	–
1, 1, 4, 7, 34, 73, 349, 817		5, 28	9, 17	–
1, 1, 4, 9, 36, 100, 400, 1225		3, 5	2, 3	–
1, 1, 4, 10, 37, 121, 451, 1639		2, 3	2, 1	2, 2
1, 1, 4, 11, 32, 110, 360, 1163		7, 30	10, 22	17, 12
1, 1, 4, 11, 36, 114, 392, 1319		6, 29	10, 21	–
1, 1, 4, 11, 40, 118, 456, 1507		5, 14	6, 10	–
1, 1, 4, 11, 44, 133, 585, 2067		10, 75	14, 62	–
1, 1, 4, 11, 45, 166, 690, 2855		9, 61	11, 53	–
1, 1, 4, 12, 44, 160, 635, 2520		6, 24	7, 16	–
1, 1, 5, 9, 45, 97, 485, 1145		3, 4	3, 1	2, 2
1, 1, 5, 13, 61, 199, 939, 3389		5, 24	9, 18	–
1, 1, 5, 15, 51, 199, 755, 2789		7, 31	11, 22	17, 12
1, 1, 5, 15, 57, 205, 809, 3119		6, 29	10, 21	–
1, 1, 5, 15, 69, 217, 1061, 3923		5, 16	7, 11	–

Sequence	Step set	Equation sizes		
1, 1, 5, 17, 53, 233, 909, 3361		4, 12	6, 5	9, 6
1, 1, 5, 17, 69, 249, 1117, 4529		6, 29	10, 21	–
1, 1, 5, 17, 71, 289, 1269, 5529		7, 29	8, 22	–
1, 2, 4, 8, 16, 32, 64, 128		1, 1	1, 0	1, 1
1, 2, 4, 10, 26, 66, 178, 488		4, 12	7, 4	8, 6
1, 2, 4, 12, 36, 100, 324, 1052		4, 16	8, 6	9, 6
1, 2, 4, 14, 46, 134, 502, 1820		4, 16	8, 6	9, 6
1, 2, 5, 13, 35, 96, 267, 750		2, 3	2, 1	2, 2
1, 2, 5, 13, 37, 111, 346, 1100		4, 15	5, 8	–
1, 2, 5, 13, 39, 126, 425, 1450		5, 28	9, 17	–
1, 2, 5, 14, 42, 132, 429, 1430		2, 2	1, 1	2, 2
1, 2, 5, 15, 49, 168, 601, 2222		5, 17	7, 10	–
1, 2, 6, 16, 48, 136, 408, 1184		3, 4	3, 1	2, 2
1, 2, 6, 16, 56, 176, 632, 2080		4, 13	5, 7	–
1, 2, 6, 16, 64, 216, 856, 2976		5, 28	9, 17	–
1, 2, 6, 18, 58, 190, 638, 2170		3, 4	3, 1	2, 2
1, 2, 6, 18, 60, 200, 700, 2450		3, 4	2, 2	–
1, 2, 6, 18, 60, 204, 716, 2560		5, 16	6, 10	–
1, 2, 6, 18, 60, 205, 732, 2667		4, 16	6, 8	–
1, 2, 6, 18, 62, 215, 809, 3045		9, 52	12, 39	–
1, 2, 6, 18, 62, 220, 816, 3066		5, 28	9, 17	–
1, 2, 6, 19, 67, 246, 947, 3746		3, 6	3, 3	–
1, 2, 6, 20, 68, 238, 854, 3106		4, 10	5, 5	8, 6
1, 2, 6, 20, 70, 250, 910, 3362		5, 15	6, 10	–
1, 2, 6, 20, 72, 272, 1064, 4272		2, 3	2, 1	2, 2

Sequence	Step set	Equation sizes		
1, 2, 6, 21, 76, 290, 1148, 4627		5, 19	7, 11	–
1, 2, 7, 21, 78, 260, 988, 3458		3, 5	2, 3	9, 8
1, 2, 7, 23, 84, 301, 1127, 4186		5, 15	6, 10	–
1, 2, 7, 23, 85, 314, 1207, 4682		3, 4	3, 1	2, 2
1, 2, 7, 23, 89, 334, 1361, 5438		4, 12	6, 8	–
1, 2, 7, 23, 94, 366, 1572, 6510		10, 75	14, 62	–
1, 2, 7, 24, 94, 370, 1537, 6440		4, 11	4, 6	–
1, 2, 7, 25, 99, 402, 1687, 7242		8, 42	12, 33	–
1, 2, 7, 25, 101, 414, 1773, 7680		9, 61	14, 48	–
1, 2, 7, 25, 101, 416, 1787, 7792		3, 6	3, 3	–
1, 2, 7, 27, 105, 426, 1787, 7590		7, 31	11, 22	17, 12
1, 2, 7, 27, 109, 450, 1903, 8194		6, 29	10, 21	–
1, 2, 7, 27, 113, 474, 2051, 9054		5, 15	7, 10	–
1, 2, 8, 24, 96, 320, 1280, 4480		2, 3	2, 1	2, 2
1, 2, 8, 24, 104, 360, 1624, 6048		4, 16	6, 8	–
1, 2, 8, 24, 112, 400, 1888, 7168		5, 28	9, 17	–
1, 2, 8, 28, 104, 400, 1536, 5936		4, 12	6, 5	9, 6
1, 2, 8, 28, 108, 408, 1600, 6208		6, 29	10, 21	–
1, 2, 8, 28, 120, 496, 2192, 9696		5, 17	7, 10	–
1, 2, 8, 29, 129, 535, 2467, 10844		5, 24	9, 18	–
1, 2, 8, 30, 124, 512, 2200, 9470		6, 29	10, 21	–
1, 2, 8, 30, 126, 530, 2330, 10290		6, 19	7, 13	–
1, 2, 9, 33, 153, 636, 3007, 13250		5, 16	7, 10	–
1, 2, 9, 34, 151, 659, 2999, 13714		5, 20	8, 15	–
1, 2, 9, 35, 155, 677, 3095, 14118		7, 29	8, 23	–

Sequence	Step set	Equation sizes		
1, 2, 10, 32, 160, 584, 2920, 11360		3, 4	3, 1	2, 2
1, 2, 10, 40, 176, 808, 3720, 17152		7, 31	11, 22	17, 12
1, 2, 10, 40, 184, 824, 3864, 17984		5, 16	7, 11	–
1, 2, 10, 40, 192, 840, 4136, 19072		6, 29	10, 21	–
1, 3, 9, 27, 81, 243, 729, 2187		1, 1	1, 0	1, 1
1, 3, 9, 29, 99, 347, 1237, 4479		4, 10	5, 4	8, 6
1, 3, 9, 31, 111, 397, 1465, 5479		7, 48	13, 33	21, 12
1, 3, 9, 31, 117, 451, 1777, 7199		4, 16	8, 6	9, 6
1, 3, 9, 35, 141, 551, 2329, 9995		7, 51	15, 34	21, 12
1, 3, 10, 35, 126, 462, 1716, 6435		2, 2	1, 1	2, 2
1, 3, 10, 37, 144, 586, 2454, 10491		9, 48	14, 38	–
1, 3, 11, 41, 157, 607, 2367, 9277		3, 4	3, 1	2, 2
1, 3, 11, 41, 165, 687, 2951, 12861		4, 15	6, 7	–
1, 3, 11, 43, 175, 731, 3111, 13427		3, 4	3, 1	2, 2
1, 3, 11, 43, 177, 751, 3263, 14421		3, 5	3, 2	–
1, 3, 11, 43, 177, 755, 3303, 14727		5, 17	7, 10	–
1, 3, 12, 45, 180, 702, 2808, 11097		3, 4	3, 1	2, 2
1, 3, 12, 49, 213, 941, 4256, 19461		3, 6	3, 3	9, 8
1, 3, 12, 49, 214, 946, 4304, 19727		4, 12	6, 8	–
1, 3, 12, 51, 227, 1032, 4771, 22303		5, 16	7, 11	–
1, 3, 13, 55, 249, 1131, 5253, 24543		2, 3	2, 1	2, 2
1, 3, 13, 55, 251, 1141, 5335, 24963		4, 9	4, 6	–
1, 3, 13, 59, 273, 1291, 6189, 29891		4, 12	6, 6	9, 6
1, 3, 13, 59, 277, 1319, 6361, 30919		6, 29	10, 21	–
1, 3, 14, 63, 302, 1440, 7020, 34175		6, 29	10, 21	–

Sequence	Step set	Equation sizes		
1, 3, 15, 63, 315, 1431, 7155, 33615		3, 4	3, 1	2, 2
1, 4, 16, 64, 256, 1024, 4096, 16384		1, 1	1, 0	1, 1
1, 4, 16, 68, 302, 1368, 6286, 29252		7, 50	15, 21	21, 12
1, 4, 16, 72, 336, 1568, 7488, 36160		4, 16	8, 5	9, 6
1, 4, 17, 75, 339, 1558, 7247, 34016		2, 3	2, 1	2, 2
1, 4, 18, 84, 400, 1928, 9368, 45776		3, 4	3, 1	2, 2
1, 4, 19, 91, 445, 2188, 10819, 53644		3, 4	3, 1	2, 2
1, 4, 20, 96, 480, 2368, 11840, 58880		3, 4	3, 1	2, 2
1, 5, 25, 125, 625, 3125, 15625, 78125		1, 1	1, 0	1, 1

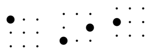
4.2. **Observations.**

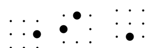
4.2.1. Not for all the above sequences, it is equally surprising that they appear to be D-finite. Intuitively, one would expect that step sets with many "regularities" (e.g., symmetries) tend to be D-finite, and that step sets with few regularities tend to be non-D-finite. This intuition is not totally wrong, but there are some quite surprising cases, too.

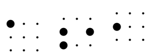
Inspired by criteria known for the 2D case [8], we considered the following criteria:

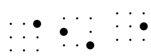
(A) The step set is symmetric about (at least) one of the axes
(P) The step set is symmetric about (at least) one plane spanned by two axes
(H) The step set is such that there exists a plane such that all elements of the step set point into the same direction of the plane.

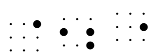
Among the 134 D-finite step sets we found, only the following five do not satisfy any of these three criteria:

 $1, 1, 2, 3, 8, 15, 44, 91, 286, 633\ldots$

 $1, 1, 2, 4, 10, 25, 70, 196, 588, 1764\ldots$

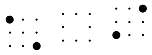
 $1, 1, 3, 5, 17, 34, 126, 279, 1095, 2588\ldots$

 $1, 1, 3, 8, 25, 77, 257, 853, 2946, 10178\ldots$

 $1, 2, 7, 24, 94, 370, 1537, 6440, 27736, 120438\ldots$

Indeed, we cound not find any simple geometric condition that would cover these cases as well. Note, however, that the corresponding equations are relatively small.

Conversely, every step set simultaneously satisfying (A) and (P) is D-finite, and so is every step set with (A) and (H). But there are 37 step sets with (P) and (H) which are, as far as we can tell, not D-finite. We were also not able to find an equation for the step set

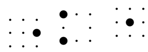
 $1, 1, 4, 7, 28, 70, 280, 787, 3148, 9526, 38104 \dots$

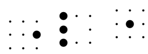
which is symmetric about all three axes, not even with 800 terms instead of 400. Also the step set

 $1, 1, 4, 13, 40, 136, 496, 1753, 6256, 22912, 85216 \dots$

which enjoy a rotational symmetry about the middle line of the first octant, and which may be viewed as a three dimensional analogue of Kreweras's step set, appears to be non-D-finite, even when 800 terms are taken into account.

For walks in the quarter plane, D-finiteness is preserved under reversing arrows, i.e., the generating function for a step set $\mathfrak{S}$ is D-finite if and only if the generating function for the step set $\mathfrak{S}'$ is, when $\mathfrak{S}'$ is obtained from $\mathfrak{S}$ by reversing all arrows [26, 8]. Our computations do not suggest that this criterion also applies in 3D. The table in Section 4.1 contains 42 step sets for whose counterpart we were not able to find an equation. Among those,

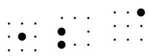
 $1, 1, 2, 4, 9, 21, 56, 148, 421 \dots$

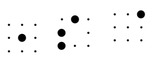
 $1, 1, 2, 5, 12, 32, 97, 282, 870 \dots$

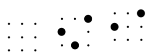
are the only ones which satisfy criterion (A) or (P). Note that the equations we found in these cases are rather large, so that chances are that their reverse does satisfy an equation but it is too big for us to find. The other step sets satisfy (H), but have no symmetry. Among those, there are sets which satisfy quite small equations, for example

 $1, 4, 17, 75, 339, 1558, 7247, 34016 \dots$

4.2.2. As in the 2D case, it turns out that most of the minimal polynomials of the algebraic series define curves of genus 0, which therefore can be rationally parameterized. There are twelve cases of genus 1, these are elliptic curves. Some of them turn out to be equivalent. For example, the cases

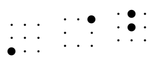
 $1, 1, 4, 11, 32, 110, 360, 1163, \dots$

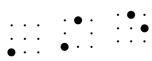
 $1, 2, 7, 27, 105, 426, 1787, 7590, \dots$

 $1, 2, 10, 40, 176, 808, 3720, 17152$

all have 1728 as j-invariant.

Most interestingly, there are also three step sets whose genus is 5 (!). These are:

 $1, 3, 9, 31, 111, 397, 1465, 5479$

 $1, 3, 9, 35, 141, 551, 2329, 9995$

 $1, 4, 16, 68, 302, 1368, 6286, 29252.$

The existence of such step sets was not at all expected.

For the transcendental series, we could observe the same phenomenon as in 2D: all the operators factor as a product of a single irreducible operator of order two and zero or more operators of order one. We therefore expect again that all these series admit a representation as a hypergeometric pullback, although we cannot justify this by a concrete example.

Also concerning asymptotics, similar remarks apply as in 2D. All series grow like $\kappa n^\alpha \rho^n$ for some constants κ, α, ρ, where ρ is an integer or an algebraic number of degree 2 and α is a negative number. We have not gone through the laborious task of determining the constants κ.

References

1. Yves André, *G-functions and geometry*, Aspects of Mathematics, E13, Friedr. Vieweg & Sohn, Braunschweig, 1989.
2. D. H. Bailey, J. M. Borwein, N. J. Calkin, R. Girgensohn, D. R. Luke, and V. H. Moll, *Integer relation detection*, Experimental Mathematics in Action, AK Peters, 2007, pp. 29–31.
3. Bernhard Beckermann and George Labahn, *A uniform approach for the fast computation of matrix-type Padé approximants*, SIAM J. Matrix Anal. Appl. **15** (1994), no. 3, 804–823.
4. Alin Bostan, Frédéric Chyzak, Grégoire Lecerf, Bruno Salvy, and Éric Schost, *Differential equations for algebraic functions*, ISSAC'07: Proceedings of the 2007 international symposium on Symbolic and algebraic computation (C. W. Brown, ed.), ACM Press, 2007, pp. 25–32.
5. Alin Bostan and Manuel Kauers, *The complete generating function for Gessel walks is algebraic*, In preparation, 2008.
6. Mireille Bousquet-Mélou, *Counting walks in the quarter plane*, Mathematics and computer science, II (Versailles, 2002), Trends Math., Birkhäuser, Basel, 2002, pp. 49–67.
7. ———, *Walks in the quarter plane: Kreweras' algebraic model*, Ann. Appl. Probab. **15** (2005), no. 2, 1451–1491.
8. Mireille Bousquet-Mélou and Marni Mishna, *Walks with small steps in the quarter plane*, Available at http://arxiv.org/abs/0810.4387, 2008.
9. Mireille Bousquet-Mélou and Marko Petkovšek, *Linear recurrences with constant coefficients: the multivariate case*, Discrete Math. **225** (2000), no. 1-3, 51–75, Formal power series and algebraic combinatorics (Toronto, ON, 1998).
10. R. Brak and A. J. Guttmann, *Algebraic approximants: a new method of series analysis*, J. Phys. A **23** (1990), no. 24, L1331–L1337.
11. A. Chambert-Loir, *Théorèmes d'algébricité en géométrie diophantienne (d'après J.-B. Bost, Y. André, D. & G. Chudnovsky)*, Séminaire Bourbaki **282** (2002), no. 886, 175–209.
12. D. V. Chudnovsky and G. V. Chudnovsky, *Applications of Padé approximations to Diophantine inequalities in values of G-functions*, Number theory (New York, 1983–84), Lecture Notes in Math., vol. 1135, Springer, Berlin, 1985, pp. 9–51.
13. Michael Dettweiler and Stefan Reiter, *On globally nilpotent differential equations*, Available at http://arxiv.org/pdf/math/0605383, 2006.
14. Bernard Dwork, *Differential operators with nilpotent p-curvature*, Amer. J. Math. **112** (1990), no. 5, 749–786.
15. Bernard Dwork, Giovanni Gerotto, and Francis J. Sullivan, *An introduction to G-functions*, Annals of Mathematics Studies, vol. 133, Princeton University Press, Princeton, NJ, 1994.
16. Philippe Flajolet, *Analytic models and ambiguity of context-free languages*, Theoretical Computer Science **49** (1987), no. 2–3, 283–309.
17. Philippe Flajolet and Robert Sedgewick, *Analytic combinatorics*, Cambridge University Press, 2008, in press.
18. Stavros Garoufalidis, *G-functions and multisum versus holonomic sequences*, Advances in Mathematics (in press).
19. Taira Honda, *Algebraic differential equations*, Symposia Mathematica, Vol. XXIV (Sympos., INDAM, Rome, 1979), Academic Press, London, 1981, pp. 169–204.
20. Nicholas M. Katz, *Nilpotent connections and the monodromy theorem: Applications of a result of Turrittin*, Inst. Hautes Études Sci. Publ. Math. (1970), no. 39, 175–232.
21. Manuel Kauers and Doron Zeilberger, *Proof of Ira Gessel's lattice path conjecture*, Tech. Report 2008-08, SFB F13, 2008.
22. ———, *The quasi-holonomic ansatz and restricted lattice walks*, J. Difference Equ. Appl. **14** (2008), no. 10-11, 1119–1126.
23. Daan Krammer, *An example of an arithmetic Fuchsian group*, J. Reine Angew. Math. **473** (1996), 69–85.

24. Germain Kreweras, *Sur une classe de problèmes de dénombrement liés au treillis des partitions des entiers*, Cahiers du B.U.R.O. **6** (1965), 5–105.
25. Christian Mallinger, *Algorithmic manipulations and transformations of univariate holonomic functions and sequences*, Master's thesis, J. Kepler University, Linz, August 1996.
26. Marni Mishna, *Classifying lattice walks restricted to the quarter plane*, J. Combin. Theory Ser. A (in press), Preliminary version published in Proceedings of FPSAC07, 2007, available at http://arxiv.org/pdf/math/0611651.
27. Simon Plouffe, *Plouffe's inverter*, http://pi.lacim.uqam.ca/.
28. Bruno Salvy and Paul Zimmermann, *Gfun: a Maple package for the manipulation of generating and holonomic functions in one variable*, ACM Transactions on Mathematical Software **20** (1994), no. 2, 163–177.
29. Hans Schmitt, *Operators with nilpotent p-curvature*, Proc. Amer. Math. Soc. **119** (1993), no. 3, 701–710.
30. Jet Wimp and Doron Zeilberger, *Resurrecting the asymptotics of linear recurrences*, Journal of Mathematical Analysis and Applications **111** (1985), 162–176.

Algorithms Project, INRIA Paris-Rocquencourt, 78153 Le Chesnay, France
E-mail address: `Alin.Bostan@inria.fr`

Research Institute for Symbolic Computation, J. Kepler University Linz, Austria
E-mail address: `mkauers@risc.uni-linz.ac.at`