The Irrational Chess Knight

Marko Petkovšek

Department of Mathematics
University of Ljubljana
Jadranska 19, 1000 Ljubljana, Slovenia
E-mail: Marko.Petkovsek@fmf.uni-lj.si
Fax: +386 61 217-281
Phone: +386 61 1766-680

Summary

While in the univariate case solutions of linear recurrences with constant coefficients have rational generating functions, in the multivariate case the situation is much more interesting: even though initial conditions have rational generating functions, the corresponding solutions can have generating functions which are algebraic but not rational, and perhaps even non-algebraic.

We start by an existence and uniqueness theorem for partial recurrences of the form

$$a_{\boldsymbol{n}} = \Phi(a_{\boldsymbol{n}+\boldsymbol{h}_1}, a_{\boldsymbol{n}+\boldsymbol{h}_2}, \ldots, a_{\boldsymbol{n}+\boldsymbol{h}_k}), \quad \text{for } \boldsymbol{n} \geq \boldsymbol{s},$$

where the values of $a_{\boldsymbol{n}}$ for $\boldsymbol{n} \not\geq \boldsymbol{s}$ are given explicitly. In particular, we show that the lattice points in the first orthant can be enumerated in such a way that for all $\boldsymbol{n}$ the points $\boldsymbol{n} + \boldsymbol{h}_i$ precede $\boldsymbol{n}$ in this enumeration, if and only if the convex hull of the set $H = \{\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_k\}$ does not intersect the first orthant. This condition on H which ensures existence and uniqueness of solution is assumed to be satisfied in the sequel.

For linear partial recurrences with constant coefficients we show that when initial conditions grow at most exponentially, the same is true of the solution, which is consequently analytic in a neighbourhood of the origin. Next we consider the algebraic nature of the generating function of the solution of such recurrences, and define the *apex* of H as the componentwise maximum of the points in $H \cup \{\boldsymbol{0}\}$. When the initial conditions have rational generating functions and the apex of H is $\boldsymbol{0}$, the generating function of the solution is rational and is given by an explicit formula. When the initial conditions have algebraic generating functions and the apex of H has at most one positive coordinate, the generating function of the solution is algebraic and can be found by solving an algebraic equation and a system of linear equations. We give several applications of this procedure to enumeration of various lattice paths with algebraic generating functions such as generalized Dyck paths and paths consisting of nonnegative steps and staying below a certain line.

Finally, when the apex has more than one positive coordinate, and the initial conditions have rational generating functions, the generating function of the solution need not be rational, which we demonstrate on the problem of the chess knight with restricted moves. We also conjecture that in this case the generating function is not even algebraic.

Résumé

Alors que dans le cas d'une seule variable les séries génératrices des solutions des équations de récurrences linéaires à coefficients constants sont rationnelles, dans le cas de plusieurs variables ces séries peuvent être algébriques non-rationnelles, ou pas même algébriques.

Nous commençons par donner un théorème d'existence et d'unicité pour des récurrences partielles ayant la forme

$$a_{\boldsymbol{n}} = \Phi(a_{\boldsymbol{n}+\boldsymbol{h}_1}, a_{\boldsymbol{n}+\boldsymbol{h}_2}, \ldots, a_{\boldsymbol{n}+\boldsymbol{h}_k}), \quad \text{pour } \boldsymbol{n} \geq \boldsymbol{s},$$

où les valeurs $a_{\boldsymbol{n}}$ pour $\boldsymbol{n} \not\geq \boldsymbol{s}$ sont connues. Nous montrons qu'il existe une énumération des points entiers naturels, telle que pour tout $\boldsymbol{n}$ les points $\boldsymbol{n} + \boldsymbol{h}_i$ précèdent $\boldsymbol{n}$ dans cette énumération si et seulement si l'intersection de l'enveloppe convexe de $H = \{\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_k\}$ avec le premier orthant est vide. Nous supposons que cette condition pour H qu'implique l'existence et l'unicité de la solution est satisfaite par toutes les récurrences que nous considérons.

Pour des récurrences partielles linéaires à coefficients constants nous montrons que la solution est analytique dans un voisinage de l'origine si les conditions initiales sont bornées par une fonction exponentielle de $\boldsymbol{n}$. Nous appelons *l'apex* de H le maximum des points de $H \cup \{\boldsymbol{0}\}$ calculé par coordonnées. Si les séries génératrices des conditions initiales sont rationnelles et l'apex de H est $\boldsymbol{0}$, la série génératrice de la solution est rationnelle et peut être calculée par une formule explicite. Si les séries génératrices des conditions initiales sont algébriques et l'apex de H a une coordonnée positive au plus, la série génératrice de la solution est algébrique. Pour la trouver on doit résoudre une équation algébrique et un système d'équations linéaires.

Enfin, si l'apex de H a plus d'une coordonnée positive et les séries génératrices des conditions initiales sont rationnelles, la série génératrice de la solution n'est pas necéssairement rationnelle. Nous présentons un exemple de ce type provenant du "problème de cavalier". Nous conjecturons que dans ce cas-là, la série génératrice n'est pas même algébrique.

1 An existence and uniqueness theorem

Throughout the paper, we use $\mathbb{N}$ to denote the set of nonnegative integers. We write $\boldsymbol{u} = (u_1, u_2, \ldots, u_d)$ for d-tuples of numbers or indeterminates, and $\boldsymbol{u} \geq \boldsymbol{v}$ when $u_i \geq v_i$ for $1 \leq i \leq d$.

Let A be a nonempty set. We consider d-dimensional partial recurrence equations of the form

$$a_{\boldsymbol{n}} = \Phi(a_{\boldsymbol{n}+\boldsymbol{h}_1}, a_{\boldsymbol{n}+\boldsymbol{h}_2}, \ldots, a_{\boldsymbol{n}+\boldsymbol{h}_k}), \quad \text{for } \boldsymbol{n} \geq \boldsymbol{s}, \tag{1}$$

where $a : \mathbb{N}^d \to A$ is the unknown d-dimensional sequence (*d-sequence* for short) of elements of A, $\Phi : A^k \to A$ is a given function, $H = \{\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_k\} \subseteq \mathbb{Z}^d$ is the set of *shifts*, and $\boldsymbol{s} \in \mathbb{N}^d$ is the *starting point* satisfying $\boldsymbol{s} + H \subseteq \mathbb{N}^d$. We assume that the initial conditions are of the form

$$a_{\boldsymbol{n}} = \varphi(\boldsymbol{n}), \quad \text{for } \boldsymbol{n} \geq \boldsymbol{0}, \boldsymbol{n} \not\geq \boldsymbol{s}, \tag{2}$$

where $\varphi : \{\boldsymbol{n} \in \mathbb{N}^d;\ \boldsymbol{n} \not\geq \boldsymbol{s}\} \to A$ is a given function.

We think of the $\boldsymbol{h}_i$ as having mostly negative coordinates, and of the point $\boldsymbol{n}$ as depending on the points $\boldsymbol{n} + \boldsymbol{h}_1, \boldsymbol{n} + \boldsymbol{h}_2, \ldots, \boldsymbol{n} + \boldsymbol{h}_k$ as far as the value of $a_{\boldsymbol{n}}$ is concerned.

The objective of this section is to characterize the sets H for which there is a well-ordering of $\mathbb{N}^d$ of order type ω such that the points $\boldsymbol{n} + \boldsymbol{h}_1, \boldsymbol{n} + \boldsymbol{h}_2, \ldots, \boldsymbol{n} + \boldsymbol{h}_k$ precede $\boldsymbol{n}$ in this ordering. Then there exists a unique solution $a_{\boldsymbol{n}}$ of (1), (2), and for any $\boldsymbol{n} \in \mathbb{N}^d$ it is possible to compute the value of $a_{\boldsymbol{n}}$ directly from (1), (2) in a finite number of steps.

Definition 1 *For $H \subseteq \mathbb{Z}^d$ and $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{N}^d$, let*

$$\boldsymbol{p} \prec_H \boldsymbol{q} \quad \text{iff} \quad \boldsymbol{p} \ \in \ \boldsymbol{q} + H \ \subseteq \ \mathbb{N}^d. \tag{3}$$

The transitive closure $\prec_H^+$ of $\prec_H$ in $\mathbb{N}^d$ is the dependency relation *corresponding to H. When $\boldsymbol{p} \prec_H^+ \boldsymbol{q}$ we say that $\boldsymbol{q}$* depends on *$\boldsymbol{p}$.* □

For a set $H \subseteq \mathbb{R}^d$ we denote by $\operatorname{conv} H$ its convex hull, and by $\operatorname{i-cone} H$ its *integer cone*:

$$\operatorname{i-cone} H = \{\boldsymbol{x} \in \mathbb{R}^d;\ \boldsymbol{x} = \sum_{i=0}^{k} \lambda_i \, \boldsymbol{h}_i,\ \lambda_i \in \mathbb{N},\ \boldsymbol{h}_i \in H\}.$$

The following theorem is proved in [7, Sec. 3.3, Cor. 2]:

Theorem 1 *Let $H \subseteq \mathbb{Z}^d$ be a finite set, and $\prec_H^+$ the corresponding dependency relation. Then the following are equivalent:*

(i) $\prec_H^+$ *is asymmetric and has no infinite descending chain in* $\mathbb{N}^d$,

(ii) $\{\boldsymbol{x} \in \mathbb{R}^d;\ \boldsymbol{x} \geq \boldsymbol{0}\} \ \cap\ \text{i-cone}\, H \ = \ \emptyset$,

(iii) $\{\boldsymbol{x} \in \mathbb{R}^d;\ \boldsymbol{x} \geq \boldsymbol{0}\} \ \cap\ \text{conv}\, H \ = \ \emptyset$,

(iv) *there exists an* $\boldsymbol{a} \in \mathbb{R}^d$, $\boldsymbol{a} > \boldsymbol{0}$, *such that* $\boldsymbol{a} \cdot \boldsymbol{h} < 0$ *for all* $\boldsymbol{h} \in H$,

(v) *there exists an* $\boldsymbol{a} \in \mathbb{N}^d$, $\boldsymbol{a} > \boldsymbol{0}$, *such that* $\boldsymbol{a} \cdot \boldsymbol{h} < 0$ *for all* $\boldsymbol{h} \in H$,

(vi) $\prec_H^+$ *can be extended to a well-ordering of* $\mathbb{N}^d$ *of order type* ω. $\square$

Now it is easy to state and prove the existence and uniqueness theorem for recurrences of the form (1), (2).

Theorem 2 *Let* $H \subseteq \mathbb{Z}^d \setminus \{\boldsymbol{0}\}$ *be a nonempty set such that* $\{\boldsymbol{x} \in \mathbb{R}^d;\ \boldsymbol{x} \geq \boldsymbol{0}\} \ \cap\ \text{conv}\, H = \emptyset$. *Then there exists a unique d-sequence* $a : \mathbb{N}^d \to A$ *which satisfies* (1), (2).

Proof: Write $H = \{\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_k\}$. To prove existence, note that by Theorem 1 there exists a well-ordering $<_H$ of $\mathbb{N}^d$ of order type ω which extends $\prec_H^+$. Let $\boldsymbol{p} : \mathbb{N} \to \mathbb{N}^d$ be a bijection which satisfies $i < j \Leftrightarrow \boldsymbol{p}_i <_H \boldsymbol{p}_j$. Such a bijection exists because $<_H$ has order type ω. Define a function $f : \mathbb{N} \to A$ recursively by $f(0) = \varphi(\boldsymbol{p}_0)$ and

$$f(i) = \begin{cases} \Phi(f(\boldsymbol{p}^{-1}(\boldsymbol{p}_i + \boldsymbol{h}_1)), f(\boldsymbol{p}^{-1}(\boldsymbol{p}_i + \boldsymbol{h}_2)), \ldots, f(\boldsymbol{p}^{-1}(\boldsymbol{p}_i + \boldsymbol{h}_k)))\,, & \text{if } \boldsymbol{p}_i \geq \boldsymbol{s}, \\ \varphi(\boldsymbol{p}_i)\,, & \text{otherwise,} \end{cases}$$

for $i > 0$. As H is nonempty and $\boldsymbol{0} \notin H$, we have $\boldsymbol{p}_0 \not\geq \boldsymbol{s}$. Because $\boldsymbol{p}_i + \boldsymbol{h}_j \prec_H^+ \boldsymbol{p}_i$ it follows that $\boldsymbol{p}^{-1}(\boldsymbol{p}_i + \boldsymbol{h}_j) < i$ so that $f(i)$ is defined in terms of values of f at smaller arguments when $\boldsymbol{p}_i \geq \boldsymbol{s}$. We conclude that f is well defined. Obviously $a_{\boldsymbol{n}} = f(\boldsymbol{p}^{-1}(\boldsymbol{n}))$ satisfies (1) and (2). Uniqueness of this solution follows easily by induction on the well-founded set $(\mathbb{N}^d, \prec_H^+)$. $\square$

This theorem generalizes the result of [10].

2 Partial recurrences with constant coefficients

In the rest of the paper, we limit our attention to recurrences of the form

$$a_{\boldsymbol{n}} = \sum_{\boldsymbol{h} \in H} c_{\boldsymbol{h}} a_{\boldsymbol{n}+\boldsymbol{h}}\,, \qquad \text{for } \boldsymbol{n} \geq \boldsymbol{s}\,, \tag{4}$$

where the set of values A is a field of characteristic zero, and $c_{\boldsymbol{h}}$ are given nonzero constants from A.

Theorem 3 *Take* $A = \mathbb{C}$, *and let* $H \subseteq \mathbb{Z}^d$ *be a finite set such that* $\{\boldsymbol{x} \in \mathbb{R}^d;\ \boldsymbol{x} \geq \boldsymbol{0}\} \ \cap\ \text{conv}\, H = \emptyset$. *Let a be the unique solution of* (2), (4). *If there are constants* $m > 0$ *and* $\boldsymbol{u} \in \mathbb{R}^d$ *such that* $|\varphi(\boldsymbol{n})| \leq m^{\boldsymbol{u}\cdot\boldsymbol{n}}$ *for all* $\boldsymbol{n} \not\geq \boldsymbol{s}$, *then the generating function of a*

$$F(x_1, x_2, \ldots, x_d) = \sum_{\boldsymbol{n} \in \mathbb{N}^d} a_{\boldsymbol{n}} x_1^{n_1} x_2^{n_2} \cdots x_d^{n_d}$$

is analytic in a neighborhood of the origin.

Proof: Existence and uniqueness of a follow from Theorem 2. By Theorem 1(iv), there is $\boldsymbol{v} \in \mathbb{R}^d$ such that $\boldsymbol{v} > \boldsymbol{0}$, and $\boldsymbol{v} \cdot \boldsymbol{h} < 0$ for all $\boldsymbol{h} \in H$. Since H is finite there exists an $\varepsilon > 0$ such that $\boldsymbol{v} \cdot \boldsymbol{h} \leq -\varepsilon$ for all $\boldsymbol{h} \in H$. Let

$$M \ = \ \max\left\{1,\ \left(\sum_{\boldsymbol{h} \in H} |c_{\boldsymbol{h}}|\right)^{\frac{1}{\varepsilon}},\ \max_{1 \leq i \leq d} m^{\frac{u_i}{v_i}}\right\}.$$

We now prove that $|a_{\boldsymbol{n}}| \leq M^{\boldsymbol{v}\cdot\boldsymbol{n}}$ for all $\boldsymbol{n} \in \mathbb{N}^d$, using induction on the well-founded set $(\mathbb{N}^d, \prec_H^+)$.

If $\boldsymbol{n} \ngeq \boldsymbol{s}$ then

$$\begin{aligned} |a_{\boldsymbol{n}}| &= |\varphi(\boldsymbol{n})| \le m^{\boldsymbol{u}\cdot\boldsymbol{n}} = m^{u_1 n_1} m^{u_2 n_2} \cdots m^{u_d n_d} \\ &= \left(m^{\frac{u_1}{v_1}}\right)^{v_1 n_1} \left(m^{\frac{u_2}{v_2}}\right)^{v_2 n_2} \cdots \left(m^{\frac{u_d}{v_d}}\right)^{v_d n_d} \le M^{\boldsymbol{v}\cdot\boldsymbol{n}} . \end{aligned}$$

Otherwise we assume inductively that $|a_{\boldsymbol{n}+\boldsymbol{h}}| \le M^{\boldsymbol{v}\cdot(\boldsymbol{n}+\boldsymbol{h})}$ for all $\boldsymbol{h} \in H$. Then

$$\begin{aligned} \left| \sum_{\boldsymbol{h}\in H} c_{\boldsymbol{h}} a_{\boldsymbol{n}+\boldsymbol{h}} \right| &\le \sum_{\boldsymbol{h}\in H} |c_{\boldsymbol{h}}| M^{\boldsymbol{v}\cdot(\boldsymbol{n}+\boldsymbol{h})} \le \sum_{\boldsymbol{h}\in H} |c_{\boldsymbol{h}}| M^{\boldsymbol{v}\cdot\boldsymbol{n}-\varepsilon} \\ &= M^{\boldsymbol{v}\cdot\boldsymbol{n}-\varepsilon} \sum_{\boldsymbol{h}\in H} |c_{\boldsymbol{h}}| < M^{\boldsymbol{v}\cdot\boldsymbol{n}} , \end{aligned}$$

proving the claim. It follows that $F(x_1, x_2, \ldots, x_d)$ converges when $|x_i| < 1/M^{v_i}$. $\square$

In the case of constant coefficients any term $a_{\boldsymbol{n}+\boldsymbol{h}}$ with a non-zero coefficient $c_{\boldsymbol{h}}$ can be expressed explicitly from (4). As it turns out, there is always at least one "good" term.

Theorem 4 *Let $G \subseteq \mathbb{Z}^d$ be a nonempty finite set. Then there exists a point $\boldsymbol{g}_0 \in G$ such that the set $H = \{\boldsymbol{g} - \boldsymbol{g}_0;\ \boldsymbol{g} \in G, \boldsymbol{g} \ne \boldsymbol{g}_0\}$ satisfies the equivalent conditions of Theorem 1.*

Proof: Let $\boldsymbol{g}_0$ be the last point in G with respect to the lexicographic ordering of $\mathbb{Z}^d$. Then it can be shown that $\prec_H^+$ is asymmetric and has no infinite descending chains. $\square$

3 Recurrences with algebraic generating functions

Definition 2 *Let $H \subseteq \mathbb{N}^d$ be a finite set. The* apex *of H is the point $\boldsymbol{p} = (p_1, p_2, \ldots, p_d) \in \mathbb{N}^d$ defined by*

$$p_i = \max\{h_i;\ \boldsymbol{h} \in H \cup \{\boldsymbol{0}\}\} \quad (i = 1, 2, \ldots, d).$$

In dimension $d = 2$, the apex of H is the upper right corner of the smallest rectangle (with its sides parallel to the axes) enclosing the set $H \cup \{\boldsymbol{0}\}$.

Theorem 5 *Let $H \subseteq \mathbb{Z}^d$ be a finite set such that $\{\boldsymbol{x} \in \mathbb{R}^d;\ \boldsymbol{x} \ge \boldsymbol{0}\} \cap \operatorname{conv} H = \emptyset$. Let a be the unique solution of (2), (4), and $F(\boldsymbol{x}) = \sum_{\boldsymbol{n}\in\mathbb{N}^d} a_{\boldsymbol{n}} \boldsymbol{x}^{\boldsymbol{n}}$ its generating function.*

(i) *If the apex of H is $\boldsymbol{0}$ and all the initial sections*

$$f_{j_1,\ldots,j_m}^{(i_1,\ldots,i_m)}(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{n}\in\mathbb{N}^d \\ n_{i_1}=j_1,\ldots,n_{i_m}=j_m}} a_{\boldsymbol{n}} \boldsymbol{x}^{\boldsymbol{n}}$$

$$(1 \le m < d, \quad 1 \le i_1 < \cdots < i_m \le d, \quad 0 \le j_r < s_{i_r} \text{ for } 1 \le r \le m) \tag{5}$$

are rational power series, then the generating function $F(\boldsymbol{x})$ is rational.

(ii) *If the apex of H has one positive coordinate and all the initial sections* (5) *are algebraic power series, then the generating function $F(\boldsymbol{x})$ is algebraic.*

Proof: From (4) we obtain

$$\sum_{\boldsymbol{n}\ge\boldsymbol{s}} a_{\boldsymbol{n}} \boldsymbol{x}^{\boldsymbol{n}} = \sum_{\boldsymbol{h}\in H} c_{\boldsymbol{h}} \sum_{\boldsymbol{n}\ge\boldsymbol{s}} a_{\boldsymbol{n}+\boldsymbol{h}} \boldsymbol{x}^{\boldsymbol{n}}$$

which can be rewritten as

$$\left(1 - \sum_{\boldsymbol{h}\in H} c_{\boldsymbol{h}} \boldsymbol{x}^{-\boldsymbol{h}}\right) F(\boldsymbol{x}) = \sum_{\boldsymbol{n}\ngeq\boldsymbol{s}} a_{\boldsymbol{n}} \boldsymbol{x}^{\boldsymbol{n}} - \sum_{\boldsymbol{h}\in H} c_{\boldsymbol{h}} \boldsymbol{x}^{-\boldsymbol{h}} \sum_{\boldsymbol{n}\ngeq\boldsymbol{s}+\boldsymbol{h}} a_{\boldsymbol{n}} \boldsymbol{x}^{\boldsymbol{n}}. \tag{6}$$

(i) If the apex of H is $\mathbf{0}$ then $\boldsymbol{h} \leq \mathbf{0}$ for all $\boldsymbol{h} \in H$. Therefore the terms $\sum_{\boldsymbol{n} \not\geq \boldsymbol{s}} a_{\boldsymbol{n}} \boldsymbol{x}^{\boldsymbol{n}}$ and $\sum_{\boldsymbol{n} \not\geq \boldsymbol{s}+\boldsymbol{h}} a_{\boldsymbol{n}} \boldsymbol{x}^{\boldsymbol{n}}$ are finite sums of initial sections and hence rational by assumption. Then

$$F(\boldsymbol{x}) = \frac{\sum_{\boldsymbol{n} \not\geq \boldsymbol{s}} a_{\boldsymbol{n}} \boldsymbol{x}^{\boldsymbol{n}} - \sum_{\boldsymbol{h} \in H} c_{\boldsymbol{h}} \boldsymbol{x}^{-\boldsymbol{h}} \sum_{\boldsymbol{n} \not\geq \boldsymbol{s}+\boldsymbol{h}} a_{\boldsymbol{n}} \boldsymbol{x}^{\boldsymbol{n}}}{1 - \sum_{\boldsymbol{h} \in H} c_{\boldsymbol{h}} \boldsymbol{x}^{-\boldsymbol{h}}} \tag{7}$$

is rational, too.

(ii) Let $\boldsymbol{p}$ be the apex of H. Wlg. assume that $p_1 > 0$ while the remaining coordinates of $\boldsymbol{p}$ are zero. Then $h_i \leq 0$ for all $i \geq 2$ and $\boldsymbol{h} \in H$, so the right side of (6) is an affine combination, with coefficients which are algebraic power series in $x_1, x_2, \ldots, x_d$, of the p_1 sections $f_{s_1}^{(1)}, f_{s_1+1}^{(1)}, \ldots, f_{s_1+p_1-1}^{(1)}$. Note that these sections are not given by the initial conditions (2). From the definition of sections it follows that we can write

$$f_t^{(1)}(x_1, x_2, \ldots, x_d) = x_1^t \, g_{t-s_1+1}(x_2, \ldots, x_d) \qquad (s_1 \leq t \leq s_1 + p_1 - 1)$$

where $g_1, g_2, \ldots, g_{p_1}$ are unknown power series. Denote

$$P(x_1, x_2, \ldots, x_d) \;=\; \left(1 - \sum_{\boldsymbol{h} \in H} c_{\boldsymbol{h}} \boldsymbol{x}^{-\boldsymbol{h}}\right) \boldsymbol{x}^{\boldsymbol{p}}.$$

From the definition of apex it follows that $\boldsymbol{p} \geq \mathbf{0}$ and $\boldsymbol{p} \geq \boldsymbol{h}$ for all $\boldsymbol{h} \in H$, hence $P(x_1, x_2, \ldots, x_d)$ is a polynomial. Now (6) can be rewritten as

$$P(x_1, x_2, \ldots, x_d) F(\boldsymbol{x}) \;=\; r_0(x_1, x_2, \ldots, x_d) + \sum_{i=1}^{p_1} r_i(x_1, x_2, \ldots, x_d) g_i(x_2, \ldots, x_d) \tag{8}$$

where r_i are algebraic series. By setting $x_2, \ldots, x_d$ to zero, we find that

$$P(x_1, 0, \ldots, 0) \;=\; \left(x_1^{p_1} - \sum_{\substack{\boldsymbol{h} \in H \\ h_2 = \cdots = h_d = 0}} c_{\boldsymbol{h}} x_1^{p_1 - h_1} \right).$$

Because $\{\boldsymbol{x} \in \mathbb{R}^d;\ \boldsymbol{x} \geq \mathbf{0}\} \cap \operatorname{conv} H = \emptyset$, we have $h_1 < 0$ for all $\boldsymbol{h} \in H$ which have $h_2 = \cdots = h_d = 0$. Therefore $p_1 - h_1 > p_1$ for all such $\boldsymbol{h}$, which means that $x_1 = 0$ is a root of $P(x_1, 0, \ldots, 0)$ of multiplicity p_1. Hence there exist p_1 Puiseux series $\xi_1(x_2, \ldots, x_d), \ldots, \xi_{p_1}(x_2, \ldots, x_d)$ which satisfy

$$P(\xi_j(x_2, \ldots, x_d), x_2, \ldots, x_d) = 0 \qquad (1 \leq j \leq p_1)$$

and which pass through the origin (i.e., $\xi_j(0, \ldots, 0) = 0$ for $1 \leq j \leq p_1$). If all ξ_j are different, substituting $\xi_1, \xi_2, \ldots, \xi_{p_1}$ for x_1 in (8) gives a linear system of p_1 linear equations with algebraic coefficients

$$\sum_{i=1}^{p_1} r_i(\xi_j(x_2, \ldots, x_d), x_2, \ldots, x_d) g_i(x_2, \ldots, x_d) + r_0(\xi_j(x_2, \ldots, x_d), x_2, \ldots, x_d) = 0 \quad (1 \leq j \leq p_1)$$

satisfied by the p_1 unknown series $g_1, \ldots, g_{p_1}$. If $P(x_1, x_2, \ldots, x_x)$ considered as a polynomial of x_1 has multiple roots, the number of equations will be less than p_1, but the missing equations can be obtained by differentiating (8) wrt. x_1 an appropriate number of times before substitution. The resulting system is in both cases equivalent to the original recurrence relation (4) and initial conditions (2) satisfied by the coefficients of $F(\boldsymbol{x})$ and is therefore uniquely solvable. It follows that $g_1, \ldots, g_{p_1}$ are algebraic, and hence so is

$$F(\boldsymbol{x}) \;=\; \frac{r_0(x_1, x_2, \ldots, x_d) + \sum_{i=1}^{p_1} r_i(x_1, x_2, \ldots, x_d) g_i(x_2, \ldots, x_d)}{P(x_1, x_2, \ldots, x_d)},$$

as claimed. □

Example 1 In the problem of *generalized Dyck paths* [3, 4] we are given a finite set of steps $\mathcal{S} = \{(r_1, s_1), \ldots, (r_m, s_m)\}$ where $r_i, s_i \in \mathbb{Z}$ and $r_i > 0$. We are interested in the number d_n of paths from $(0,0)$ to $(n,0)$ using only steps from $\mathcal{S}$ and staying within the first quadrant.

Denote by $d_{i,k}$ the number of such paths ending at (i,k) (rather than at $(n,0)$), and let $r = \max_{1\le i\le m} r_i$, $s = \max_{1\le i\le m} s_i$. Attach $r-1$ columns of zeros to the left of the array d, and s rows of zeros below the array d. Call the resulting array a. Then $d_{i,k} = a_{i+r-1,k+s}$ for $i,k \ge 0$, and a satisfies

$$\begin{aligned} a_{i,k} &= a_{i-r_1,k-s_1} + \cdots + a_{i-r_m,k-s_m} \quad (i \ge r, k \ge s), && (9)\\ a_{i,k} &= \delta_{i,r-1}\delta_{k,s} \quad (i < r \text{ or } k < s). && (10)\end{aligned}$$

This is a problem of the type (2), (4) where $\boldsymbol{s} = (r,s)$, $H = \{(-r_1,-s_1), \ldots, (-r_m,-s_m)\}$, and $c_{\boldsymbol{h}} = 1$ for all $\boldsymbol{h} \in H$. As all r_i are positive, $\{\boldsymbol{x} \in \mathbb{R}^d;\ \boldsymbol{x} \ge \boldsymbol{0}\} \cap \operatorname{conv} H = \emptyset$, hence there exists a unique solution of (9), (10). Let $t = -\min\{s_1, \ldots, s_m\}$. We distinguish two cases.

a) If $t \le 0$ then the apex of H is $(0,0)$, so we can use formula (7) to obtain the rational generating functions

$$F(x,y) = \sum_{i,k=0}^{\infty} a_{i,k} x^i y^k = \frac{x^{r-1}y^s}{1 - x^{r_1}y^{s_1} - \cdots - x^{r_m}y^{s_m}},$$

$$G(x,y) = \sum_{i,k=0}^{\infty} d_{i,k} x^i y^k = \frac{F(x,y)}{x^{r-1}y^s} = \frac{1}{1 - x^{r_1}y^{s_1} - \cdots - x^{r_m}y^{s_m}},$$

$$g(x) = \sum_{n=0}^{\infty} d_n x^n = G(x,0) = \left(1 - \sum_{\substack{1\le i\le m\\ s_i=0}} x^{r_i}\right)^{-1}.$$

b) If $t > 0$ then the apex of H is $(0,t)$ and the corresponding generating functions are algebraic. Here

$$P(x,y) = y^t - x^{r_1}y^{s_1+t} - \cdots - x^{r_m}y^{s_m+t}$$

and Eqn. (6) multiplied by y^{t-s} yields

$$\frac{P(x,y)}{y^s} F(x,y) = x^{r-1}y^t - \sum_{\substack{1\le i\le m\\ s_i<0}} x^{r_i}y^{s_i+t} \sum_{k=s}^{s-1-s_i} y^{k-s} f_k(x) \tag{11}$$

where $f_k(x) = \sum_{i=0}^{\infty} a_{i,k}x^i$. Let $y = \xi_1(x), \ldots, \xi_t(x)$ be those Puiseux series solutions of $P(x,y) = 0$ which pass through the origin. According to the procedure described in the proof of Theorem 5 (ii), we should now substitute $\xi_1(x), \ldots, \xi_t(x)$ for y in (11) and solve the resulting linear system for the unknown $f_s(x), \ldots, f_{s+t-1}(x)$. However, in this particular case we can immediately see what $F(x,y)$ is. This is because the right-hand side of (11) is a polynomial in y of degree t with leading coefficient x^{r-1}, and we know all its zeros: they are $\xi_1(x), \ldots, \xi_t(x)$ (counted with their respective multiplicities), because they are zeros of the left-hand side. Therefore the right-hand side of (11) is $x^{r-1}(y - \xi_1(x)) \cdots (y - \xi_t(x))$ and we have finally

$$F(x,y) = \sum_{i,k=0}^{\infty} a_{i,k}x^i y^k = \frac{x^{r-1}y^s(y-\xi_1(x))\cdots(y-\xi_t(x))}{P(x,y)},$$

$$G(x,y) = \sum_{i,k=0}^{\infty} d_{i,k}x^i y^k = \frac{F(x,y)}{x^{r-1}y^s} = \frac{(y-\xi_1(x))\cdots(y-\xi_t(x))}{P(x,y)},$$

$$g(x) = \sum_{n=0}^{\infty} d_n x^n = G(x,0) = \frac{(-1)^{t+1}\xi_1(x)\cdots\xi_t(x)}{P(x,0)} = (-1)^{t+1}\xi_1(x)\cdots\xi_t(x)\left(\sum_{\substack{1\le i\le m\\ s_i=-t}} x^{r_i}\right)^{-1}.$$

In the special case $\mathcal{S} = \{(s,r),(r,s),(s,-r),(r,-s)\}$ where $r > s > 0$ we have

$$g(x) = \frac{(-1)^{r+1}\xi_1(x)\cdots\xi_r(x)}{x^s}$$

where $y = \xi_1(x),\ldots,\xi_r(x)$ are those solutions of $y^r - x^s y^{2r} - x^r y^{r+s} - x^s - x^r y^{r-s} = 0$ which pass through the origin.

In the same way we could count generalized Dyck paths with coloured steps – then the corresponding coefficient $c_{\boldsymbol{h}}$ would equal the number of colours allowed for this step.

This approach generalizes without significant change to higher-dimensional paths, provided that the steps have positive coordinates in all but perhaps one (fixed) dimension. In all these cases, the generating functions are algebraic.

Example 2 Consider the problem of counting lattice paths with steps $(1,0)$ and $(0,1)$ which start at the origin and stay on or below the line $y = (m-1)x$. Using the linear transformation $(i,k) \mapsto ((m-1)i - k, k)$ we obtain the equivalent problem of counting lattice paths with steps $(m-1,0)$ and $(-1,1)$ which start at the origin and stay within the first quadrant. Let $d_{i,k}$ denote the number of such paths which end at (i,k). Attaching $m-1$ columns of zeros to the left of the array d and one row of zeros below the array d, and changing the element at $(0,1)$ to 1, we obtain a new array a which satisfies $a_{i,k} = d_{i-m+1,k-1}$ and

$$a_{i,k} = a_{i-m+1,k} + a_{i+1,k-1} \quad (i \geq m-1, k \geq 1), \tag{12}$$

$$a_{i,k} = \delta_{i,0}\delta_{k,1} \quad (i < m-1 \text{ or } k < 1). \tag{13}$$

Here $\boldsymbol{s} = (m-1,1)$, $H = \{(-(m-1),0),(1,-1)\}$, the apex is at $(1,0)$, and $P(x,y) = x - x^m - y$. From (6) we obtain

$$P(x,y)F(x,y) = xy - y^2 - yx^{m-1}f(y) \tag{14}$$

where $f(y) = \sum_{k=0}^{\infty} a_{m-1,k}$ is unknown. Let $x = \xi(y)$ be that solution of $P(x,y) = 0$ which passes through the origin. Then, by substituting $\xi(y)$ for x in (14), we find

$$f(y) = \frac{\xi(y) - y}{\xi(y)^{m-1}} = \xi(y)$$

and

$$F(x,y) = \sum_{i,k=0}^{\infty} a_{i,k}x^i y^k = y\frac{x - y - x^{m-1}\xi(y)}{x - y - x^m},$$

$$G(x,y) = \sum_{i,k=0}^{\infty} d_{i,k}x^i y^k = \frac{F(x,y) - y}{x^{m-1}y} = \frac{x - \xi(y)}{x - y - x^m},$$

$$g(y) = \sum_{k=0}^{\infty} d_{0,k}y^k = G(0,y) = \frac{\xi(y)}{y}.$$

The algebraic equation satisfied by $g(y)$ is easily obtained from $P(\xi(y),y) = 0$, and is $g - y^{m-1}g - 1 = 0$.

Only every $(m-1)$-st coefficient of g is nonzero. Omitting the zero coefficients we have

$$h(y) = \sum_{k=0}^{\infty} d_{0,(m-1)k}y^k = g(y^{1/(m-1)})$$

which satisfies $h - yh^m - 1 = 0$.

In a similar way, we could count the paths which start at the origin and stay below the line $y = (m-1)x$ (cf. [1]). Instead of the steps $(1,0)$ and $(0,1)$ we could take any set of steps with nonnegative components, and still obtain algebraic generating functions. Generalization to higher dimensions is also possible.

Additional examples of this kind can be found in [2, Exer. 2.2.1.–4, 11] (the ballot numbers) and in [6] (the number of elements of the free modular lattice generated by the poset $1+1+n$).

$i \setminus k$	0	1	2	3	4	5	6	7	8	9	10
0	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	2	2	3	3	5	6	8	12	18
3	1	1	2	2	4	5	7	11	17	23	40
4	1	1	3	4	6	10	16	22	39	62	91
5	1	1	3	5	10	14	27	44	67	123	208
6	1	1	5	7	16	27	44	83	145	225	432
7	1	1	6	11	22	44	83	134	268	476	767
8	1	1	8	17	39	67	145	268	450	908	1656
9	1	1	12	23	62	123	225	476	908	1534	3161
10	1	1	18	40	91	208	432	767	1656	3161	5422

Figure 1: Solution of (15), (16) for $i, k \leq 10$.

4 The problem of the knight

When the apex of the set of shifts H has more than one positive coordinate, we suspect that the generating function of the solution of (2), (4) need not be algebraic even though all the initial sections (5) are. In this section we present an example of this type and prove that the generating function is not rational.

Example 3 On an otherwise empty chessboard which is infinite upwards and to the right there is a knight occupying the square (i, k). If the knight is only allowed to move either 2 left and 1 up, or 2 down and 1 right, in how many different ways could it reach the border of the chessboard? The border consists of the first two rows and the first two columns, and once the knight reaches it, it is not allowed to move anymore.

Let $a_{i,k}$ be the answer to this lattice-path problem. Then, obviously,

$$a_{i,k} = a_{i-2,k+1} + a_{i+1,k-2} \qquad (i, k \geq 2), \tag{15}$$

$$a_{i,0} = a_{i,1} = a_{0,k} = a_{1,k} = 1 \qquad (i, k \geq 0). \tag{16}$$

Here $H = \{(-2, 1), (1, -2)\}$ and $\{\boldsymbol{x} \in \mathbb{R}^d;\ \boldsymbol{x} \geq \boldsymbol{0}\} \cap \operatorname{conv} H = \emptyset$, so there exists a unique solution a of (15), (16), part of which is shown in Fig. 1. The apex of H is $(1, 1)$, so Theorem 5 does not apply. In fact, this is the simplest example in which the apex of H has two positive coordinates and H is symmetric w.r.t. the line $i = k$. By induction on $i + k$ one can show that $a_{i,k} = a_{k,i}$ and hence $F(x, y) = F(y, x)$. One can also show that $1 \leq a_{ik} \leq 2^{i+k}$, therefore the power series

$$F(x, y) = \sum_{i,k=0}^{\infty} a_{i,k} x^i y^k$$

converges at least for $|x|, |y| < 1/2$. Let $f_k(x) = \sum_{i=0}^{\infty} a_{i,k} x^i$ denote the generating function for the k^{th} row of a. Then from (15) it follows that $F(x, y)$ satisfies

$$\begin{aligned} (x^3 + y^3 - xy)F(x, y) &= xy(a_{0,0} + a_{1,0}x + a_{0,1}y + a_{1,1}xy) + x(x^2 - y)\left(f_0(x) + y f_1(x)\right) \\ &+ y(y^2 - x)\left(f_0(y) + x f_1(y)\right) + x^2 y^2 \left(x f_2(x) + y f_2(y)\right), \end{aligned} \tag{17}$$

while (16) implies that $f_0(x) = f_1(x) = 1/(1 - x)$. Writing $f(x)$ for $x f_2(x)$, equation (17) thus turns into

$$(x^3 + y^3 - xy)F(x, y) = \frac{x^3 + y^3 - xy + x^2y^2(xy - x - y)}{(1 - x)(1 - y)} + x^2y^2(f(x) + f(y)). \tag{18}$$

By restricting (17) to the cubic curve

$$x^3 + y^3 - xy = 0 \tag{19}$$

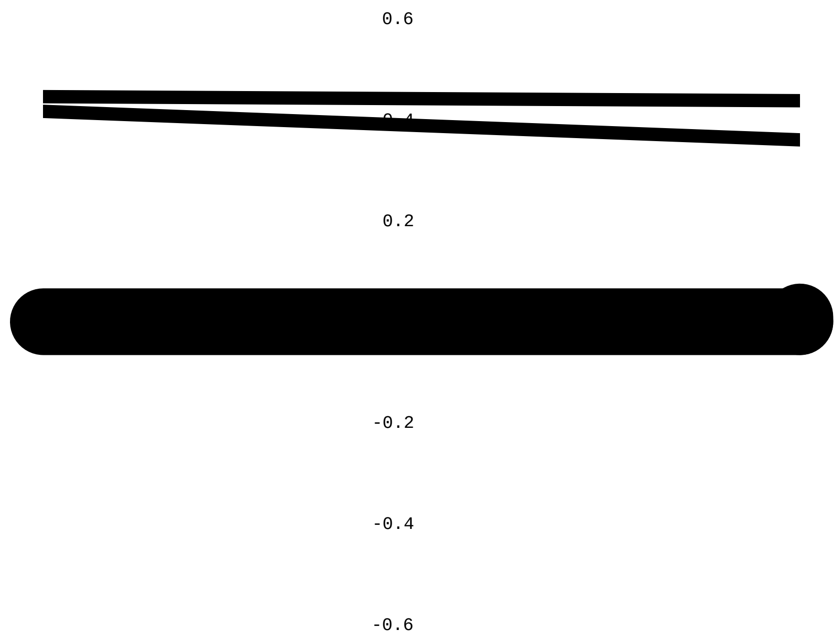

Figure 2: The leaf of Descartes ($x^3 + y^3 = xy$).

shown in Fig. 2, an additional functional equation satisfied by f is obtained:

$$f(x) + f(y) = \frac{1}{(1-x)(1-y)} - 1 \qquad (\text{when } x^3 + y^3 = xy). \tag{20}$$

This equation uniquely determines $f(x) = x \sum_{k=0}^{\infty} a_{2,k} x^k$.

Theorem 6 *The formal power series f defined by (20) is not rational.*

Proof: Assume that $f(x)$ is a rational function of x. Let $x(t) = t/(1+t^3)$, $y(t) = t^2/(1+t^3)$ be a rational parameterization of (19). Then

$$f(x(t)) + f(y(t)) = \frac{1}{(1-x(t))(1-y(t))} - 1 \tag{21}$$

as an equality of rational functions of t.

Let $t_0 = 1/\sqrt[3]{2}$, $x_0 = x(t_0) = \sqrt[3]{4}/3$, $y_0 = y(t_0) = \sqrt[3]{2}/3$. We claim that $f(x)$ is singular at $x = x_0$. Differentiating both sides of (21) w.r.t. t we find

$$f'(x) \;=\; \frac{1}{(1-x)^2(1-y)} + \frac{\dot{y}(t)}{\dot{x}(t)}\left(\frac{1}{(1-x)(1-y)^2} - f'(y)\right) \qquad (x = x(t), y = y(t)). \tag{22}$$

The derivative $f'(y)$ is regular at $y = y_0$, because the series $f(y)$ converges for $|y| < 1/2$ and $y_0 < 1/2$. By using the rough upper bound $a_{2,k} \leq 2^{k+2}$ for $k \geq n$, we can estimate

$$f'(y_0) \;=\; \sum_{k=0}^{\infty} (k+1) a_{2,k} y_0^k$$

$$\begin{aligned}
&\le \sum_{k=0}^{n-1}(k+1)a_{2,k}y_0^k \;+\; \sum_{k=n}^{\infty}(k+1)2^{k+2}y_0^k \\
&= \sum_{k=0}^{n-1}(k+1)a_{2,k}y_0^k \;+\; 8(n(1-2y_0)+1)\frac{(2y_0)^n}{(1-2y_0)^2}
\end{aligned}$$

which, using a computer algebra system and taking $n = 99$, gives $f'(y_0) < 4.726$. On the other hand, $1/((1-x_0)(1-y_0)^2) > 6.312$, proving that $1/((1-x_0)(1-y_0)^2) - f'(y_0) \neq 0$. Likewise $\dot{y}(t_0) = \sqrt[3]{4}/3 \neq 0$ and $(1-x_0)^2(1-y_0) \neq 0$, while $\dot{x}(t_0) = 0$. From (22) it follows that the rational function $f'(x(t))$ is the sum of two terms, one of which is regular and the other singular at $t = t_0$. Therefore $f'(x)$, and hence $f(x)$, is singular at $x = x_0$ as claimed.

On the other hand, from (21) it follows that $f(x(t)) = 1/((1-x(t))(1-y(t))) - f(y(t)) - 1$. All three terms on the right are regular at $t = t_0$, hence so is $f(x(t))$. Therefore $f(x)$ should be regular at $x = x_0$. This contradiction shows that $f(x)$ is not a rational power series. □

We conjecture that $f(x)$ (and therefore $F(x,y)$) is not algebraic, and, moreover, not even D-finite [9, 5].

Note that by replacing the initial conditions (16) by

$$\begin{aligned}
a_{i,0} &= 2^i, \quad a_{i,1} = 2^{i+1} \qquad (i \ge 0),\\
a_{0,k} &= 2^k, \quad a_{1,k} = 2^{k+1} \qquad (k \ge 0),
\end{aligned}$$

the solution of (15) changes to $a_{i,k} = 2^{i+k}$ with rational generating function $F(x,y) = \frac{1}{(1-2x)(1-2y)}$.

Acknowledgement. The author wishes to express his thanks to Philippe Flajolet, Bruno Salvy and Ivan Vidav for helpful discussions of the problem of the knight. He is also indebted very much to an anonymous referee who pointed out applications to lattice path problems and provided further candidates for nonalgebraic generating function.

References

[1] I. Gessel, A factorization for formal Laurent series and lattice path enumeration, *J. Comb. Th. A* **28** (1980) 321–337.

[2] D. E. Knuth, *Art of Computer Programming, Vol. 1: Fundamental Algorithms*, Addison-Wesley, Reading Mass., 1968.

[3] J. Labelle, Y.-N. Yeh, Dyck paths of knight moves, *Discr. Appl. Math.* **24** (1989) 213–221.

[4] J. Labelle, Y.-N. Yeh, Generalized Dyck paths, *Discr. Math.* **82** (1990) 1–6.

[5] L. Lipshitz, D-finite power series, *J. Algebra* **122** (1989) 353 – 373.

[6] P. Luksch, M. Petkovšek, An explicit formula for $|FM(1+1+n)|$, *Order* **6** (1990) 319 – 324.

[7] M. Petkovšek, *Finding Closed-Form Solutions of Difference Equations by Symbolic Methods*, Ph.D. Thesis, Carnegie Mellon University, Pittsburgh PA, 1991 (CMU-CS-91-103).

[8] J. Riordan, *Combinatorial Identities*, John Wiley & Sons, New York, 1968.

[9] R. P. Stanley, Differentiably finite power series, *European J. Combin.* **1** (1980) 175–188.

[10] R. Suarez, Difference equations and a principle of double induction, *Math. Magazine* **62** (1989) 334 – 339.