Two-step walks on the square lattice: Full and half-plane

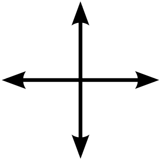
Nicholas Beaton
LIPN - Université Paris 13

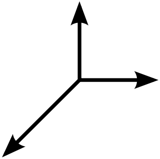
ALEA in Europe – CIRM, Luminy
25 October 2013

Introduction: Counting lattice paths

There are broadly two types of restrictions we can place on walks on an infinite lattice:

1. the types of steps allowed
   eg. on $\mathbb{Z}^2$, steps in $\mathcal{S} = \{(0,1),(1,0),(0,-1),(-1,0)\}$ = {N,E,S,W} or
   $\mathcal{S} = \{(0,1),(1,0),(-1,-1)\}$ = {N,E,SW}





2. where the walks can go
   eg. full plane, half-plane, quarter-plane, wedge of arbitrary angle α

We can then ask questions like

- how many walks of length *n*?
- asymptotics?
- generating functions?
  - functional equations?
  - explicit solutions?
  - rational/algebraic/D-finite?
- what does the "average walk" look like? where is the endpoint?
- random sampling?

Introduction: Counting lattice paths

- Banderier and Flajolet 2002: Generating functions of directed walks (all steps in positive **x** direction) in half-plane are algebraic
- Via bijections, this follows for gfs of all walks in half-plane.

Believed that gfs of walks in a quarter-plane would be D-finite. Until...

- Bousquet-Mélou and Petkovšek 2003: gf of walks taking steps in $\{(2,-1),(-1,2)\}$ and staying in first quadrant is non-D-finite
- Mishna and Rechnitzer 2007: gfs of walks taking steps in $\{(-1,1),(1,1),(1,-1)\}$ or $\{(-1,1),(0,1),(1,-1)\}$ and staying in first quadrant are non-D-finite
- Bousquet-Mélou and Mishna 2010: for steps in $\{-1,0,1\}^2 \backslash \{(0,0)\}$, there are 79 non-isomorphic cases. 23 are D-finite, conjectured that the other 56 are non-D-finite.
- Bostan and Kauers 2009: series analysis agreeing with Bousquet-Mélou and Mishna
- Raschel 2012: integral representations for the gfs of all 79 models.
- Kurkova and Raschel 2012: trivariate gfs for 51 of the 56 are non-D-finite
- Melczer and Mishna 2013: remaining 3 are non-D-finite

Two-step rules

Definition

A two-step rule $\mathcal{R}$ is a mapping

$$\mathcal{R} : \{\text{north, east, south, west}\}^2 \mapsto \{0, 1\},$$

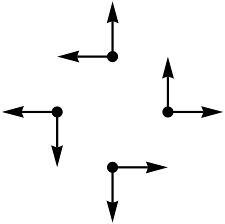
where $\mathcal{R}(i, j) = 1$ if step j can follow step i, and $\mathcal{R}(i, j) = 0$ if not.

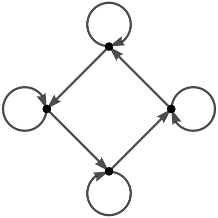
Let $\mathcal{T}$ be the set of all two-step rules.

$\mathcal{R}$ can be represented with a transfer matrix

$$\mathbf{T} \equiv \mathbf{T}(\mathcal{R}) = [\mathcal{R}(i, j)]_{\text{n,e,s,w}} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

or diagramatically

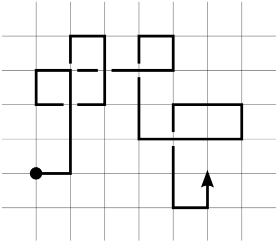


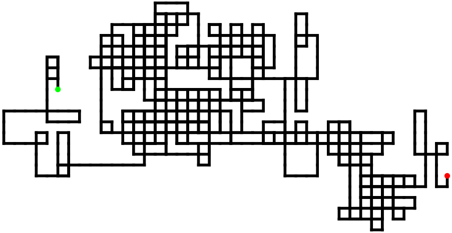


Definition

A walk $w = (v_1, v_2, \ldots, v_m) \in \{\text{n,e,s,w}\}^*$ on the edges of the square lattice obeys $\mathcal{R}$ if $\mathcal{R}(v_i, v_{i+1}) = 1$ for $i = 1 \ldots m-1$.

Let $\mathcal{W} \equiv \mathcal{W}(\mathcal{R})$ be the set of walks obeying $\mathcal{R}$, and let $\mathcal{W}_m \subset \mathcal{W}$ be the walks of length m.





Define $a_m \equiv a_m(\mathcal{R})$ to be the number of walks (up to translation) of length m obeying $\mathcal{R}$.

- What is a_m? Asymptotics?
- Generating function?
- What does the average walk look like? eg. location of endpoint?
- What if we move to half-plane? Quarter-plane?

Inspiration

Guttmann, Prellberg and Owczarek 1993: considered self-avoiding walks obeying two-step rules.

11 non-isomorphic cases, with 4 universality classes based on length-scale exponents.

Classification & isomorphisms

$|\mathcal{T}| = 2^{16} = 65536$, but many rules are trivial.

Definition

A two-step rule $\mathcal{R}$ is connected, if for $i, j \in \{\text{n,e,s,w}\}$, there is a walk of length ≥ 1 obeying $\mathcal{R}$ which starts with i and ends with j.

Let $\mathcal{C}$ be the set of connected rules.

Equivalently, for each pair i, j there exists a $k \geq 1$ such that $(\mathbf{T}^k)_{ij} \geq 1$; or, the digraph must be strongly connected.

Proposition

Of the 65536 two-step rules, 25696 are connected.

Other funny things can happen, eg. for

$$\mathbf{T}(\mathcal{R}) = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix},$$

$a_m \sim C_m \cdot \mu^m$ as $m \to \infty$, where $\mu \approx 1.55377$ and

$$C_m \approx \begin{cases} 2.42491 & m \text{ odd} \\ 2.41421 & m \text{ even.} \end{cases}$$

ie. the rule is periodic with period 2.

Definition

A two-step rule is aperiodic if there exists a $k \geq 1$ such that for all pairs i, j, $(\mathbf{T}^k)_{ij} \geq 1$.

Let $\mathcal{A}$ be the set of aperiodic rules.

Equivalently, $\mathbf{T}$ is primitive.

Proposition

Of the 25696 connected two-step rules, 25575 are aperiodic.

Lots of redundancy here, at least when it comes to enumeration.

Definition

Two rules $\mathcal{R}_1$ and $\mathcal{R}_2$ with walk sets $\mathcal{W}_1$ and $\mathcal{W}_2$ are isomorphic if there exists a permutation π of $\{$n,e,s,w$\}$ with $\pi(\mathcal{W}_1) = \mathcal{W}_2$.

In the full plane, this is the same as saying that we can permute the rows and columns of $\mathbf{T}$ to get $\mathbf{T}'$; or that the digraphs are isomorphic. Not the case in the half- or quarter-plane!

Via Burnside's lemma:

Proposition

In the full plane, there are 3044 non-isomorphic two-step rules. Of these, 1168 are connected and 1159 are aperiodic.

Enumeration in the full plane

Let n_m be the number of walks of length m ending with a north step, and likewise e_m, s_m, w_m. Take $\mathbf{c}_m = (n_m, e_m, s_m, w_m)$. Then $c_1 = (1,1,1,1)$ (ie. walks can start in any direction), and

$$\mathbf{c}_m = \mathbf{c}_{m-1} \cdot \mathbf{T} \qquad \text{for } m \geq 2.$$

By induction,

$$\mathbf{c}_m = c_1 \cdot \mathbf{T}^{m-1} \qquad \text{for } m \geq 1.$$

Taking the Jordan normal form of $\mathbf{T}$,

$$\mathbf{c}_m = c_1 \cdot \mathbf{S}\mathbf{J}^{m-1}\mathbf{S}^{-1},$$

where $\mathbf{J}$ is the Jordan normal form of $\mathbf{T}$ and $\mathbf{S}$ is the matrix of generalised eigenvectors.

Theorem (Perron-Frobenius)

The eigenvalue of a primitive matrix $\mathbf{M}$ *with greatest absolute value is real, positive, simple and unique.*

If μ is the dominant eigenvalue of $\mathbf{T}$, then μ^{m-1} will come to dominate $\mathbf{J}^{m-1}$. WLOG say $\mu = \mathbf{J}_{11}$, with corresponding eigenvector $\mathbf{v}_1$.

Proposition

As $m \to \infty$,

$$\mathbf{c}_m \sim ||\mathbf{v}_1|| \times \mu^{m-1} \times (\mathbf{S}^{-1})_{1\bullet},$$

and hence

$$a_m \sim \frac{1}{\mu} \times ||\mathbf{v}_1|| \times ||(\mathbf{S}^{-1})_{1\bullet}|| \times \mu^m.$$

Including weights

For $x, y \in (0, \infty)$, define

$$\hat{\mathbf{T}} \equiv \hat{\mathbf{T}}(x,y) = \mathbf{T} \cdot \begin{pmatrix} y & 0 & 0 & 0 \\ 0 & x & 0 & 0 \\ 0 & 0 & 1/y & 0 \\ 0 & 0 & 0 & 1/x \end{pmatrix}$$

and $\theta_m(a,b) = \#$ walks of length m obeying a given rule, which start at $(0,0)$ and end with a θ step at (a,b). Then

$$\hat{\theta}_m(x,y) = \sum_{a,b} \theta_m(a,b) x^a y^b$$

and

$$\hat{\mathbf{c}}_m(x,y) = (\hat{n}_m(x,y), \hat{e}_m(x,y), \hat{s}_m(x,y), \hat{w}_m(x,y))$$

Let $\hat{a}_m(x,y)$ be the sum of the $\hat{\theta}_m(x,y)$. Then like before,

$$\hat{\mathbf{c}}_m(x,y) = \hat{\mathbf{c}}_1(x,y) \cdot \hat{\mathbf{T}}^{m-1} \qquad \text{with } \hat{\mathbf{c}}_1(x,y) = (y, x, 1/y, 1/x).$$

$\hat{\mathbf{T}}$ has a simple dominant eigenvalue, $\mu(x,y)$.

Proposition

As $m \to \infty$,

$$\hat{\mathbf{c}}_m \sim ||\hat{\mathbf{v}}_1|| \times \mu(x,y)^{m-1} \times (\hat{\mathbf{S}}^{-1})_{1\bullet},$$

and hence

$$a_m \sim \frac{1}{\mu(x,y)} \times ||\hat{\mathbf{v}}_1|| \times ||(\hat{\mathbf{S}}^{-1})_{1\bullet}|| \times \mu(x,y)^m$$

Generating functions

Define the generating functions

$$F_\theta(x,y) \equiv F_\theta(t;x,y) = \sum_m \hat{\theta}_m(x,y)t^m = \sum_{m,a,b} \theta_m(a,b)t^m x^a y^b$$

The recursion becomes a set of equations in the F_θ. eg. for spiral walks,

$$\begin{aligned}
F_n(x,y) &= ty + tyF_n(x,y) + tyF_e(x,y)\\
F_e(x,y) &= tx + txF_e(x,y) + txF_s(x,y)\\
F_s(x,y) &= \frac{t}{y} + \frac{t}{y}F_s(x,y) + \frac{t}{y}F_w(x,y)\\
F_w(x,y) &= \frac{t}{x} + \frac{t}{x}F_n(x,y) + \frac{t}{x}F_w(x,y).
\end{aligned}$$

Equivalently,

$$(\mathbf{I} - t\hat{\mathbf{T}}^\top)\cdot \begin{pmatrix} F_n(x,y)\\ F_e(x,y)\\ F_s(x,y)\\ F_w(x,y)\end{pmatrix} = \begin{pmatrix} ty\\ tx\\ t/y\\ t/x\end{pmatrix}$$

So

$$\begin{pmatrix} F_n(x,y) \\ F_e(x,y) \\ F_s(x,y) \\ F_w(x,y) \end{pmatrix} = (\mathbf{I} - t\hat{\mathbf{T}}^\top)^{-1} \cdot \begin{pmatrix} ty \\ tx \\ t/y \\ t/x \end{pmatrix}$$

The dominant singularity of $\det(\mathbf{I} - t\hat{\mathbf{T}}^\top)$ is $1/\mu(x,y) \Rightarrow$ coefficient of t^m grows like $\mu(x,y)^m$.

What are the numerator and denominator of F_θ? Want a more combinatorial construction.

Write

$$F_\theta(t;x,y) = A_\theta(t;x,y) + B_\theta(t;x,y)F_\theta(t;x,y)$$

where A_θ counts walks with no θ steps except for the last step, and B_θ counts the subset which can follow a θ step. So

$$F_\theta(x,y) = \frac{A_\theta(x,y)}{1 - B_\theta(x,y)}$$

A_θ and B_θ have simple solutions involving $\hat{\mathbf{T}}$.

Proposition

The dominant singularity of $F_\theta(x,y)$ is a simple pole at $t = \rho(x,y) = 1/\mu(x,y)$. This is the smallest point at which $B_\theta(t;x,y) = 1$.

So

$$\hat{\theta}_m(x,y) \sim \frac{A_\theta(\rho(x,y);x,y)}{\rho(x,y)B_\theta^{(1,0,0)}(\rho(x,y);x,y)}\mu(x,y)^m$$

Location of the endpoint and drift

If γ has length m and ends at $(\mathbf{x}_m, \mathbf{y}_m) = (a, b)$, take the Boltzmann distribution on walks ending with step θ:

$$\mathbb{P}_m(\gamma) = \frac{x^a y^b}{\hat{\theta}_m(x, y)}.$$

Then we are interested in $\langle \mathbf{x}_m \rangle$ and $\langle \mathbf{y}_m \rangle$, given by

$$\langle \mathbf{x}_m \rangle = \frac{x[t^m]\frac{\partial}{\partial x} F_\theta(t; x, y)}{[t^m] F_\theta(t; x, y)} \quad \text{and} \quad \langle \mathbf{y}_m \rangle = \frac{y[t^m]\frac{\partial}{\partial y} F_\theta(t; x, y)}{[t^m] F_\theta(t; x, y)}$$

Text

Proposition

Define $P(x, y) = 1/\rho(x, y)$. Then as $m \to \infty$,

$$\langle \mathbf{x}_m \rangle = x\delta_{\mathbf{x}} m + const. + O(1/m) \quad \textit{and} \quad \langle \mathbf{y}_m \rangle = y\delta_{\mathbf{y}} m + const. + O(1/m)$$

where

$$\delta_{\mathbf{x}} = \frac{\partial}{\partial x} \log P(x, y) \quad \textit{and} \quad \delta_{\mathbf{y}} = \frac{\partial}{\partial y} \log P(x, y)$$

and the lower-order terms depend on θ.

Upper half-plane: more classifications and isomorphisms

Definition

A (connected, aperiodic) rule is rational if every walk $(\mathrm{s}, \ldots, \mathrm{s})$ must contain a north step; it is antirational if every walk $(\mathrm{n}, \ldots, \mathrm{n})$ must contain a south step. Then a rule is

- N-rational if it is rational but not antirational
- S-rational if it is antirational but not rational
- D-rational if it is both rational and antirational
- irrational if it is neither rational nor antirational

In the upper half-plane, fewer symmetries to exploit: can only swap east and west steps.

Proposition

In the upper half-plane there are

- *1525 non-isomorphic N-rational rules*
- *1525 non-isomorphic S-rational rules*
- *157 non-isomorphic D-rational rules*
- *9722 non-isomorphic irrational rules*

N-rational rules are constrained to $\mathbf{y} \geq -1$; S-rational rules are constrained to $\mathbf{y} \leq 1$; and D-rational rules are constrained to $-1 \leq \mathbf{y} \leq 1$.

Form of solution and asymptotics is different for each.

Generating functions

Upper half-plane partition functions $\hat{\theta}^{+}_{m}(x,y)$ for each step direction, with generating functions $H_\theta(x,y) \equiv H_\theta(t;x,y)$. They satisfy the system

$$(\mathbf{I} - t\,\hat{\mathbf{T}}^{\top}) \cdot \begin{pmatrix} H_n(x,y) \\ H_e(x,y) \\ H_s(x,y) \\ H_w(x,y) \end{pmatrix} = \begin{pmatrix} ty \\ tx \\ 0 \\ t/x \end{pmatrix} - t\,\mathbf{I}_s \cdot \hat{\mathbf{T}} \cdot \begin{pmatrix} H_n(x,0) \\ H_e(x,0) \\ H_s(x,0) \\ H_w(x,0) \end{pmatrix}$$

where

$$\mathbf{I}_s = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Again, we seek a more combinatorially 'clear' form:

$$H_\theta(x,y) = C_\theta(x,y) + B_\theta(x,y) H_\theta(x,y) - D_\theta(x,y) H^*(x)$$

where

- $C_\theta(x,y) \equiv C_\theta(t;x,y)$ is the gf of walks counted by A_θ (ie. in the full plane) which do not start with a south step
- $D_\theta(x,y) \equiv D_\theta(t;x,y)$ is the gf of walks counted by A_θ which do start with a south step
- $H^*(x) \equiv H^*(t;x)$ is the gf of all walks (in the upper half-plane) which (a) end on the surface, and (b) end with a step that can precede a south step.

C_θ and D_θ have simple solutions like A_θ and B_θ.

So

$$H_\theta(x,y) = \frac{C_\theta(x,y) - D_\theta(x,y)H^*(x)}{1 - B_\theta(x,y)}.$$

Need solution to $H^*(x)$!

Rearrange to kernel form:

$$(1 - B_\theta(x,y))H_\theta(x,y) = C_\theta(x,y) - D_\theta(x,y)H^*(x)$$

For D-rational rules, $B_\theta(x,y)$ has no dependence on y at all. Kernel method makes no sense! For N-rational rules, it is linear in y, but the root of $B_\theta(x,y) = 1$ is not a power series in t.

However, for rational rules (ie. N-rational and D-rational)

$$H^*(x) = \mathbf{T}_{es} H_e(x,0) + \mathbf{T}_{ws} H_w(x,0)$$

So take $y = 0$ in the appropriate equations and solve simultaneously.

Proposition

For N-rational and D-rational rules,

$$H^*(x) = \lim_{y \to 0} \frac{H^*_{num}(x,y)}{H^*_{den}(x,y)},$$

where

$$\begin{aligned} H^*_{num}(x,y) &= \mathbf{T}_{es} C_e(x,y)(1 - B_w(x,y)) + \mathbf{T}_{ws} C_w(x,y)(1 - B_e(x,y)) \\ H^*_{den}(x,y) &= (1 - B_e(x,y))(1 + \mathbf{T}_{ws} D_w(x,y)) + (1 - B_w(x,y))(1 + \mathbf{T}_{es} D_e(x,y)) \\ &\qquad - (1 - B_e(x,y)B_w(x,y)). \end{aligned}$$

For S-rational rules, the kernel is again linear in y but this time the solution is a power series in t:

$$B_\theta(t;x,\upsilon(t;x)) = 1.$$

$\upsilon(t;x)$ is the inverse of $\rho(x,y)$, ie. $\rho(x,\upsilon(t;x)) = t$.

Proposition

For an S-rational rule, let $\upsilon(t;x)$ be the unique function satisfying $B_\theta(t;x,\upsilon(t;x)) = 1$ for any of the θ. Then

$$H^*(x) = \frac{C_\theta(x,\upsilon(t;x))}{D_\theta(x,\upsilon(t;x))}.$$

Finally for irrational rules, the kernel has two roots $\upsilon^-(t;x)$ and $\upsilon^+(t;x)$. The smaller is a power series in t.

Proposition

For an irrational rule, let $\upsilon^-(t;x)$ be the smaller of the two functions satisfying $B_\theta(t;x,\upsilon^\pm(t;x)) = 1$ for any of the θ. Then

$$H^*(x) = \frac{C_\theta(x,\upsilon^-(t;x))}{D_\theta(x,\upsilon^-(t;x))}.$$

Asymptotics

For N-rational and D-rational rules, same dominant singularity a pole at $t = \rho(x, y)$:

$$\hat{\theta}_m(x, y) \sim const.\mu(x, y)^m$$

For S-rational rules, dominant singularity is now a pole of $\upsilon(t; x) \Rightarrow$ growth rate strictly smaller than full plane:

$$\hat{\theta}_m(x, y) \sim const.\tau(x, y)^m$$

with $\tau(x, y) < \mu(x, y)$.

For irrational rules, depends on sign of the vertical drift $\delta_{\mathbf{y}}$:

- if $\delta_{\mathbf{y}} > 0$, still a simple pole at $t = \rho(x, y)$:

$$\hat{\theta}_m(x, y) \sim const.\mu(x, y)^m$$

- if $\delta_{\mathbf{y}} = 0$, now a $1/\surd$ singularity at $t = \rho(x, y)$:

$$\hat{\theta}_m(x, y) \sim const.m^{-1/2}\mu(x, y)^m$$

- if $\delta_{\mathbf{y}} < 0$, now a $\surd$ singularity in $\upsilon^-(t; x) \Rightarrow$ strictly smaller growth rate:

$$\hat{\theta}_m(x, y) \sim const.m^{-3/2}\tau(x, y)^m$$

  with $\tau(x, y) < \mu(x, y)$.

Location of the endpoint

If $\delta_{\mathbf{y}} > 0$, same picture as before.

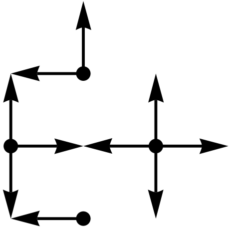
For D-rational and S-rational rules, $\langle \mathbf{y}_m \rangle \in [0, 1]$, while $\langle \mathbf{x}_m \rangle$ is $O(m)$ or $O(1)$ depending on the rule.

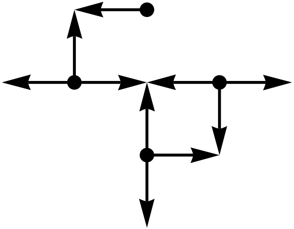
For irrational rules with $\delta_{\mathbf{y}} = 0$, now $\langle \mathbf{y}_m \rangle = O(m^{1/2})$, while $\langle \mathbf{x}_m \rangle$ can be $O(m)$, $O(m^{1/2})$ or $O(1)$, depending on the rule.

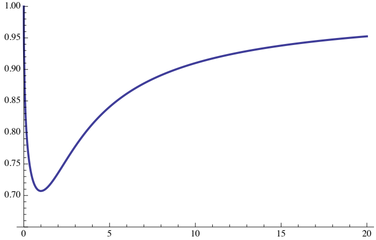
For irrational rules with $\delta_{\mathbf{y}} < 0$, now $\langle \mathbf{y}_m \rangle = O(1)$ while $\langle \mathbf{x}_m \rangle$ can be $O(m)$ or $O(1)$.

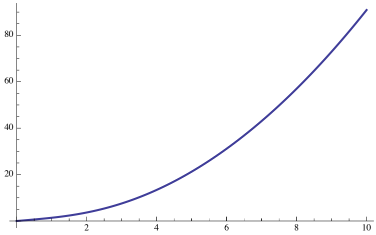
Phase boundary

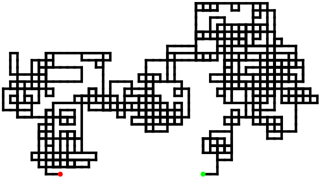
For irrational rules, can vary x and y to change the sign of $\delta_{\mathbf{y}}$. Leads to a "phase boundary" in the x-y plane, along the points (x, y) where $\delta_{\mathbf{x}} = 0$.



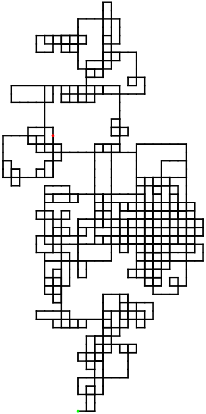




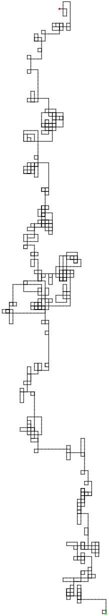




(a) $(x, y) = (1, 0.2)$



(b) $(x, y) = (1, 1)$



(c) $(x, y) = (1, 1.2)$

Quarter-plane

Reflective symmetry in $\mathbf{y} = \mathbf{x}$.

Proposition

In the first quadrant $\mathbf{x}, \mathbf{y} \geq 0$, there are 32896 non-isomorphic two-step rules, of which 12916 are connected. Of those, 12849 are aperiodic. Of those, 7520 are irrational in both axes.

Significantly more complicated. Generating functions satisfy the system

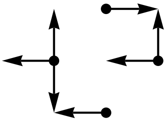
$$(\mathbf{I} - t\,\hat{\mathbf{T}}^{\top}) \cdot \begin{pmatrix} Q_n(x,y) \\ Q_e(x,y) \\ Q_s(x,y) \\ Q_w(x,y) \end{pmatrix} = \begin{pmatrix} ty \\ tx \\ 0 \\ 0 \end{pmatrix} - t\,\mathbf{I}_s \cdot \hat{\mathbf{T}} \cdot \begin{pmatrix} Q_n(x,0) \\ Q_e(x,0) \\ Q_s(x,0) \\ Q_w(x,0) \end{pmatrix} - t\,\mathbf{I}_w \cdot \hat{\mathbf{T}} \cdot \begin{pmatrix} Q_n(0,y) \\ Q_e(0,y) \\ Q_s(0,y) \\ Q_w(0,y) \end{pmatrix}$$

which becomes

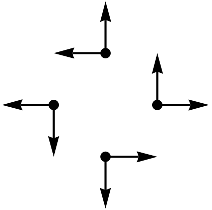
$$Q_\theta(x,y) = G_\theta(x,y) + B_\theta(x,y) Q_\theta(x,y) - D_\theta(x,y) Q^{\downarrow}(x) - J_\theta(x,y) Q^{\leftarrow}(y)$$

For some rules (eg. rational in both axes), can easily solve these equations. In general, seems to be harder.

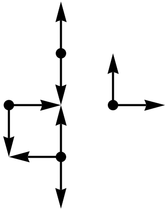
For irrational rules, different things can happen: (based on series analysis with Manuel Kauers' `Guess` package)



generating functions $Q_\theta(1,1)$ are algebraic



generating functions $Q_\theta(1,1)$ are D-finite but not algebraic



generating functions $Q_\theta(1,1)$ are non-D-finite