Smith Normal Form and Combinatorics

Richard P. Stanley

Smith normal form

A: $n \times n$ matrix over commutative ring ***R*** (with 1)

Suppose there exist ***P***, ***Q*** $\in \mathrm{GL}(n, R)$ such that

$$PAQ := B = \mathrm{diag}(d_1, d_1 d_2, \ldots, d_1 d_2 \cdots d_n),$$

where $d_i \in R$. We then call B a **Smith normal form (SNF)** of A.

Smith normal form

A: $n \times n$ matrix over commutative ring ***R*** (with 1)

Suppose there exist ***P***, ***Q*** $\in \mathrm{GL}(n, R)$ such that

$$PAQ := B = \mathrm{diag}(d_1, d_1 d_2, \ldots, d_1 d_2 \cdots d_n),$$

where $d_i \in R$. We then call B a **Smith normal form (SNF)** of A.

NOTE. (1) Can extend to $m \times n$.

$$(2)\ \mathrm{unit} \cdot \det(A) = \det(B) = d_1^n d_2^{n-1} \cdots d_n.$$

Thus SNF is a refinement of $\det$.

Row and column operations

Can put a matrix into SNF by the following operations.

- Add a multiple of a row to another row.
- Add a multiple of a column to another column.
- Multiply a row or column by a **unit** in R.

Row and column operations

Can put a matrix into SNF by the following operations.

- Add a multiple of a row to another row.
- Add a multiple of a column to another column.
- Multiply a row or column by a **unit** in R.

Over a field, SNF is **row reduced echelon form** (with all unit entries equal to 1).

Existence of SNF

PIR: principal ideal ring, e.g., $\mathbb{Z}$, $K[x]$, $\mathbb{Z}/m\mathbb{Z}$.

If R is a PIR then A has a unique SNF up to units.

Existence of SNF

PIR: principal ideal ring, e.g., $\mathbb{Z}$, $K[x]$, $\mathbb{Z}/m\mathbb{Z}$.

If R is a PIR then A has a unique SNF up to units.

Otherwise A "typically" does not have a SNF but may have one in special cases.

Algebraic note

Not known in general for which rings R does every matrix over R have an SNF.

Algebraic note

Not known in general for which rings R does every matrix over R have an SNF.

Necessary condition: R is a **Bézout ring**, i.e., every finitely generated ideal is principal.

Example. ring of entire functions and ring of all algebraic integers (not PIR's)

Algebraic note

Not known in general for which rings R does every matrix over R have an SNF.

Necessary condition: R is a **Bézout ring**, i.e., every finitely generated ideal is principal.

Example. ring of entire functions and ring of all algebraic integers (not PIR's)

Open: every matrix over a Bézout domain has an SNF.

Algebraic interpretation of SNF

R: a PID

A: an $n \times n$ matrix over R with rows $v_1, \dots, v_n \in R^n$

$\mathrm{diag}(e_1, e_2, \dots, e_n)$: SNF of A

Algebraic interpretation of SNF

R: a PID

A: an $n \times n$ matrix over R with rows $v_1, \dots, v_n \in R^n$

$\mathrm{diag}(e_1, e_2, \dots, e_n)$: SNF of A

Theorem.

$$R^n/(v_1, \dots, v_n) \cong (R/e_1R) \oplus \cdots \oplus (R/e_nR).$$

Algebraic interpretation of SNF

R: a PID

A: an $n \times n$ matrix over R with rows $v_1, \dots, v_n \in R^n$

$\mathrm{diag}(e_1, e_2, \dots, e_n)$: SNF of A

Theorem.

$$R^n/(v_1, \dots, v_n) \cong (R/e_1R) \oplus \cdots \oplus (R/e_nR).$$

$R^n/(v_1, \dots, v_n)$: **(Kasteleyn) cokernel** of A

An explicit formula for SNF

R: a PID

A: an $n \times n$ matrix over R with $\det(A) \neq 0$

$\mathrm{diag}(e_1, e_2, \ldots, e_n)$: SNF of A

An explicit formula for SNF

R: a PID

A: an $n \times n$ matrix over R with $\det(A) \neq 0$

$\mathrm{diag}(e_1, e_2, \ldots, e_n)$: SNF of A

Theorem. $e_1 e_2 \cdots e_i$ *is the gcd of all* $i \times i$ *minors of* A*.*

minor: determinant of a square submatrix.

Special case: e_1 is the gcd of all entries of A.

An example

Reduced Laplacian matrix of K_4:

$$A = \begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix}$$

An example

Reduced Laplacian matrix of K_4:

$$A = \begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix}$$

Matrix-tree theorem $\Longrightarrow \det(A) = 16$, the number of spanning trees of K_4.

An example

Reduced Laplacian matrix of K_4:

$$A = \begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix}$$

Matrix-tree theorem $\Longrightarrow \det(A) = 16$, the number of spanning trees of K_4.

What about SNF?

An example (continued)

$$\begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & -1 \\ -4 & 4 & -1 \\ 8 & -4 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & -1 \\ -4 & 4 & 0 \\ 8 & -4 & 0 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 0 & 0 & -1 \\ 0 & 4 & 0 \\ 4 & -4 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & -1 \\ 0 & 4 & 0 \\ 4 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

Reduced Laplacian matrix of K_n

$$\begin{aligned} \boldsymbol{L_0(K_n)} &= nI_{n-1} - J_{n-1} \\ \det L_0(K_n) &= n^{n-2} \end{aligned}$$

Reduced Laplacian matrix of K_n

$$\begin{aligned} \boldsymbol{L_0(K_n)} &= nI_{n-1} - J_{n-1} \\ \det L_0(K_n) &= n^{n-2} \end{aligned}$$

Trick: 2×2 submatrices (up to row and column permutations):

$$\begin{bmatrix} n-1 & -1 \\ -1 & n-1 \end{bmatrix}, \quad \begin{bmatrix} n-1 & -1 \\ -1 & -1 \end{bmatrix}, \quad \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix},$$

with determinants $n(n-2)$, $-n$, and 0. Hence $e_1 e_2 = n$. Since $\prod e_i = n^{n-2}$ and $e_i | e_{i+1}$, we get the SNF $\text{diag}(1, n, n, \ldots, n)$. ∎

Laplacian matrices of general graphs

SNF of the Laplacian matrix of a graph: very interesting

connections with sandpile models, chip firing, abelian avalanches, etc.

Laplacian matrices of general graphs

SNF of the Laplacian matrix of a graph: very interesting

connections with sandpile models, chip firing, abelian avalanches, etc.

no time for further details

Laplacian matrices of general graphs

SNF of the Laplacian matrix of a graph: very interesting

connections with sandpile models, chip firing, abelian avalanches, etc.

no time for further details 

SNF of random matrices

Huge literature on random matrices, mostly connected with eigenvalues.

Very little work on SNF of random matrices over a PID.

Is the question interesting?

$\mathbf{Mat}_k(n)$: all $n \times n$ $\mathbb{Z}$-matrices with entries in $[-k, k]$ (uniform distribution)

$p_k(n, d)$: probability that if $M \in \mathrm{Mat}_k(n)$ and SNF$(M) = (e_1, \ldots, e_n)$, then $e_1 = d$.

Is the question interesting?

$\mathbf{Mat}_k(n)$: all $n \times n$ $\mathbb{Z}$-matrices with entries in $[-k, k]$ (uniform distribution)

$p_k(n, d)$: probability that if $M \in \mathrm{Mat}_k(n)$ and SNF$(M) = (e_1, \dots, e_n)$, then $e_1 = d$.

Recall: $e_1 =$ gcd of 1×1 minors (entries) of M

Is the question interesting?

$\mathbf{Mat}_k(n)$: all $n \times n$ $\mathbb{Z}$-matrices with entries in $[-k, k]$ (uniform distribution)

$p_k(n, d)$: probability that if $M \in \mathrm{Mat}_k(n)$ and $\mathrm{SNF}(M) = (e_1, \ldots, e_n)$, then $e_1 = d$.

Recall: $e_1 = $ gcd of 1×1 minors (entries) of M

Theorem. $\lim_{k\to\infty} p_k(n, d) = \dfrac{1}{d^{n^2}\zeta(n^2)}$

Specifying some e_i

with **Yinghui Wang**

Specifying some e_i

with **Yinghui Wang** (王颖慧)

Specifying some e_i

with **Yinghui Wang** (王颖慧)

Two general results.

- Let $\alpha_1,\dots,\alpha_{n-1}\in\mathbb{P}$, $\alpha_i|\alpha_{i+1}$.

  $\boldsymbol{\mu_k(n)}$: probability that the SNF of a random $A\in\mathrm{Mat}_k(n)$ satisfies $e_i=\alpha_i$ for $1\leq\alpha_i\leq n-1$.

$$\boldsymbol{\mu(n)}=\lim_{k\to\infty}\mu_k(n).$$

  Then $\mu(n)$ exists, and $0<\mu(n)<1$.

Second result

- Let $\alpha_n \in \mathbb{P}$.

  $\boldsymbol{\nu_k(n)}$: probability that the SNF of a random $A \in \mathrm{Mat}_k(n)$ satisfies $e_n = \alpha_n$.

  Then
$$\lim_{k\to\infty} \nu_k(n) = 0.$$

Sample result

$\boldsymbol{\mu_k(n)}$: probability that the SNF of a random $A \in \mathrm{Mat}_k(n)$ satisfies $e_1 = 2$, $e_2 = 6$.

$$\boldsymbol{\mu(n)} = \lim_{k\to\infty} \mu_k(n).$$

Conclusion

$$
\begin{aligned}
\mu(n) \;&=\; 2^{-n^2}\left(1-\sum_{i=(n-1)^2}^{n(n-1)} 2^{-i} + \sum_{i=n(n-1)+1}^{n^2-1} 2^{-i}\right)\\
&\cdot\frac{3}{2}\cdot 3^{-(n-1)^2}(1-3^{(n-1)^2})(1-3^{-n})^2\\
&\cdot\prod_{p>3}\left(1-\sum_{i=(n-1)^2}^{n(n-1)} p^{-i} + \sum_{i=n(n-1)+1}^{n^2-1} p^{-i}\right).
\end{aligned}
$$

Cyclic cokernel

$\boldsymbol{\kappa(n)}$: probability that an $n \times n$ $\mathbb{Z}$-matrix has SNF diag$(e_1, e_2, \ldots, e_n)$ with $e_1 = e_2 = \cdots = e_{n-1} = 1$

Cyclic cokernel

$\boldsymbol{\kappa(n)}$: probability that an $n \times n$ $\mathbb{Z}$-matrix has SNF $\mathrm{diag}(e_1, e_2, \ldots, e_n)$ with $e_1 = e_2 = \cdots = e_{n-1} = 1$

Theorem. $$\kappa(n) = \frac{\prod_p \left(1 + \frac{1}{p^2} + \frac{1}{p^3} + \cdots + \frac{1}{p^n}\right)}{\zeta(2)\zeta(3)\cdots}$$

Cyclic cokernel

$\boldsymbol{\kappa(n)}$: probability that an $n \times n$ $\mathbb{Z}$-matrix has SNF $\mathrm{diag}(e_1, e_2, \ldots, e_n)$ with $e_1 = e_2 = \cdots = e_{n-1} = 1$

Theorem.
$$\kappa(n) = \frac{\prod_p \left(1 + \frac{1}{p^2} + \frac{1}{p^3} + \cdots + \frac{1}{p^n}\right)}{\zeta(2)\zeta(3)\cdots}$$

Corollary.
$$\lim_{n\to\infty} \kappa(n) = \frac{1}{\zeta(6)\prod_{j\geq 4}\zeta(j)}$$
$$\approx 0.846936\cdots .$$

Small number of generators

g: number of generators of cokernel (number of entries of SNF $\neq 1$) as $n \to \infty$

previous slide: $\mathrm{Prob}(g = 1) = 0.846936 \cdots$

Small number of generators

g: number of generators of cokernel (number of entries of SNF $\neq 1$) as $n \to \infty$

previous slide: Prob$(g = 1) = 0.846936\cdots$

$$\text{Prob}(g \leq 2) = 0.99462688\cdots$$

Small number of generators

g: number of generators of cokernel (number of entries of SNF $\neq 1$) as $n \to \infty$

previous slide: Prob$(g = 1) = 0.846936\cdots$

$$\mathrm{Prob}(g \leq 2) = 0.99462688\cdots$$

$$\mathrm{Prob}(g \leq 3) = 0.99995329\cdots$$

Small number of generators

g: number of generators of cokernel (number of entries of SNF $\neq 1$) as $n \to \infty$

previous slide: $\text{Prob}(g = 1) = 0.846936\cdots$

$$\text{Prob}(g \leq 2) = 0.99462688\cdots$$

$$\text{Prob}(g \leq 3) = 0.99995329\cdots$$

Theorem. $\text{Prob}(g \leq \ell) =$

$$1 - (3.46275\cdots)2^{-(\ell+1)^2}(1 + O(2^{-\ell}))$$

Small number of generators

g: number of generators of cokernel (number of entries of SNF $\neq 1$) as $n \to \infty$

previous slide: $\text{Prob}(g = 1) = 0.846936\cdots$

$$\text{Prob}(g \le 2) = 0.99462688\cdots$$

$$\text{Prob}(g \le 3) = 0.99995329\cdots$$

Theorem. $\text{Prob}(g \le \ell) =$

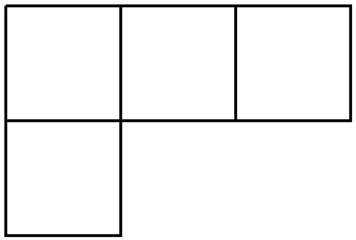
$$1 - (\mathbf{3.46275\cdots})2^{-(\ell+1)^2}(1 + O(2^{-\ell}))$$

$3.46275\cdots$

$$3.46275\cdots = \frac{1}{\prod_{j\geq 1}\left(1-\frac{1}{2^j}\right)}$$

Example of SNF computation

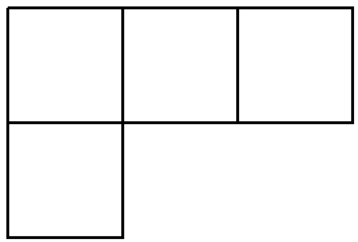
$\boldsymbol{\lambda}$: a partition $(\lambda_1, \lambda_2, \dots)$, identified with its Young diagram



(3,1)

Example of SNF computation

λ: a partition $(\lambda_1, \lambda_2, \ldots)$, identified with its Young diagram

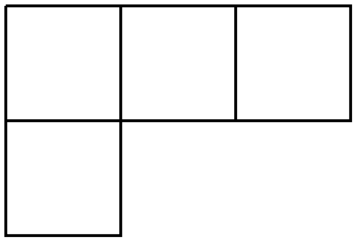


(3,1)

λ*: λ extended by a border strip along its entire boundary

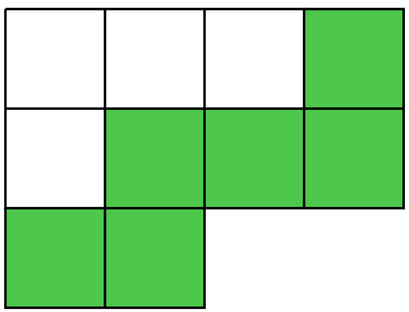
Example of SNF computation

$\boldsymbol{\lambda}$: a partition $(\lambda_1, \lambda_2, \ldots)$, identified with its Young diagram



(3,1)

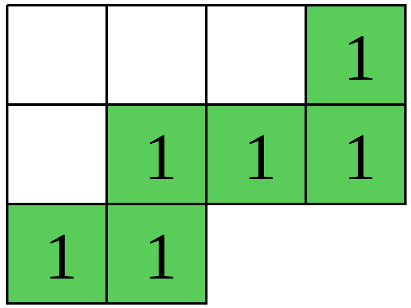
$\boldsymbol{\lambda^*}$: λ extended by a border strip along its entire boundary



(3,1)* = (4,4,2)

Initialization

Insert 1 into each square of λ^*/λ.



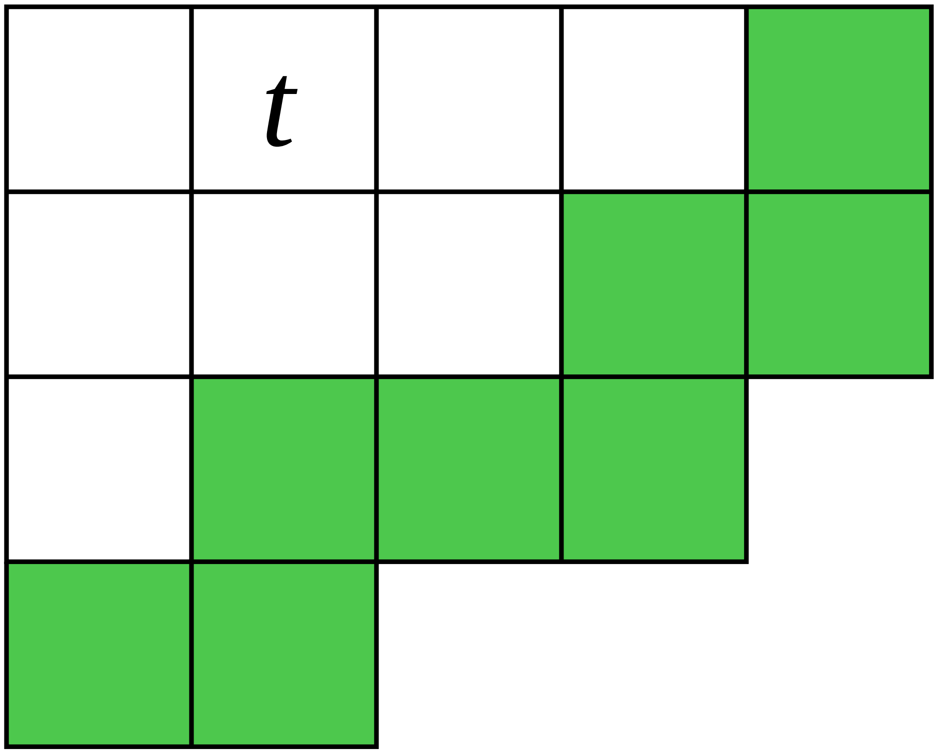
$(3,1)^* = (4,4,2)$

M_t

Let $t \in \lambda$. Let M_t be the largest square of λ^* with t as the upper left-hand corner.

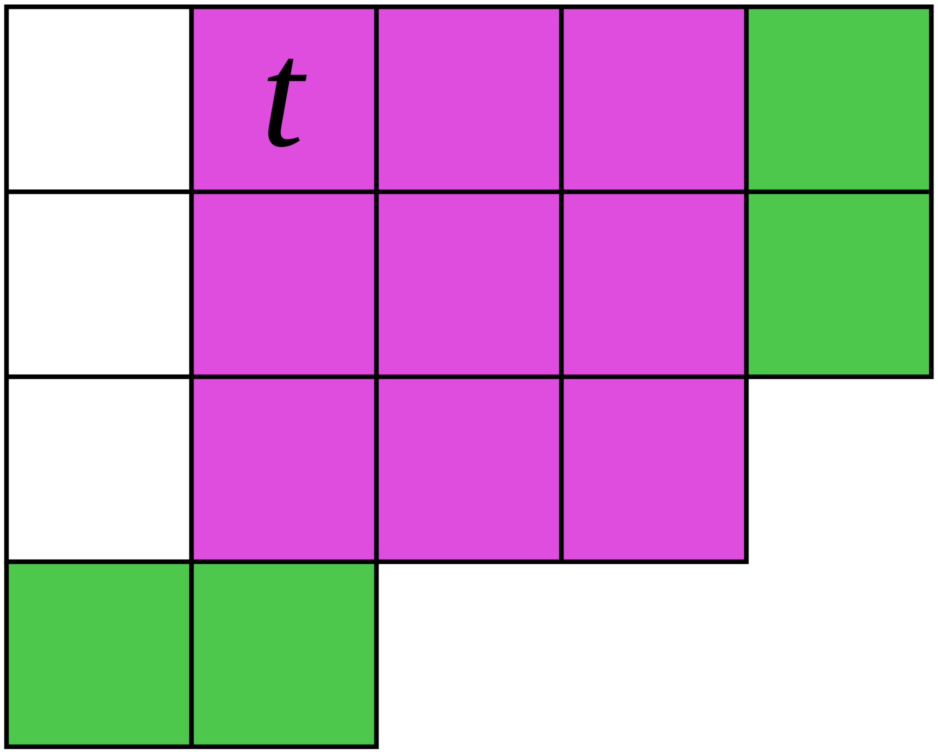
M_t

Let $t \in \lambda$. Let M_t be the largest square of λ^* with t as the upper left-hand corner.



M_t

Let $t \in \lambda$. Let M_t be the largest square of λ^* with t as the upper left-hand corner.

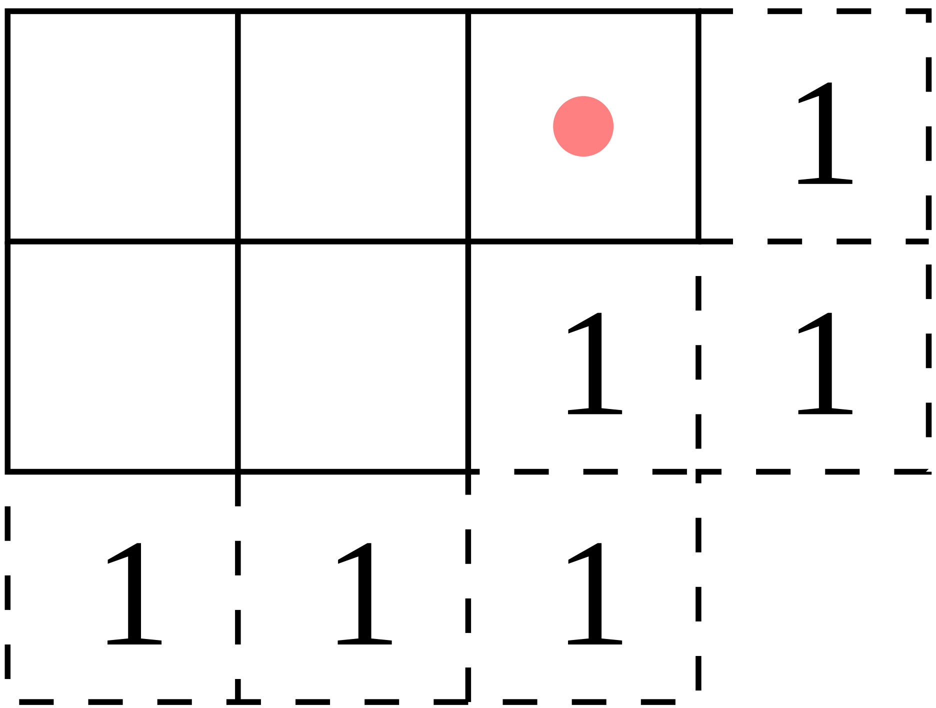


Determinantal algorithm

Suppose all squares to the southeast of t have been filled. Insert into t the number $\boldsymbol{n_t}$ so that $\det M_t = 1$.

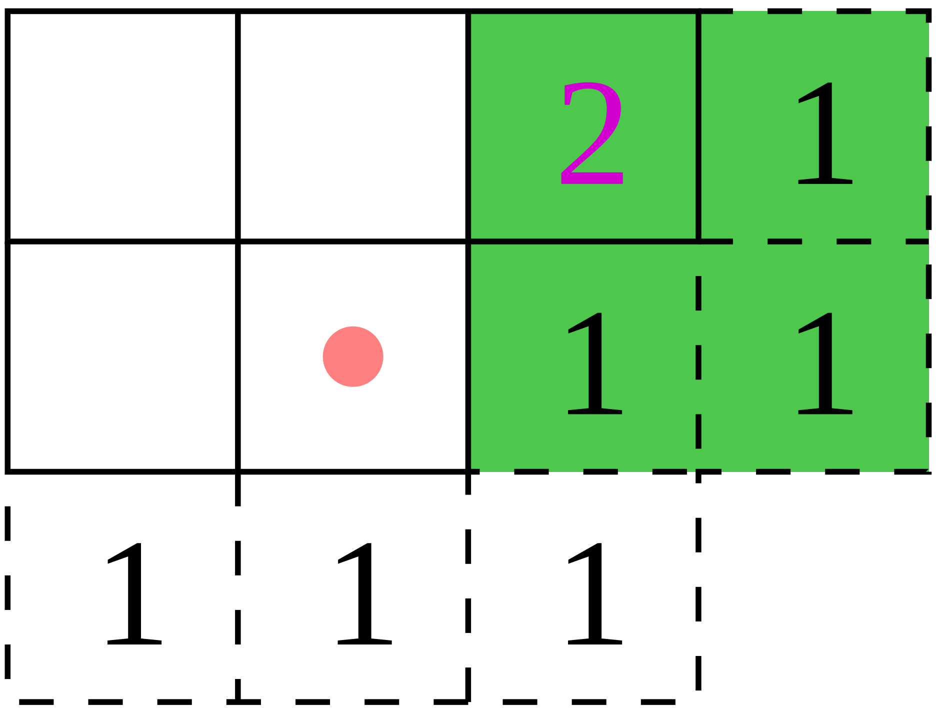
Determinantal algorithm

Suppose all squares to the southeast of t have been filled. Insert into t the number $\boldsymbol{n_t}$ so that $\det M_t = 1$.



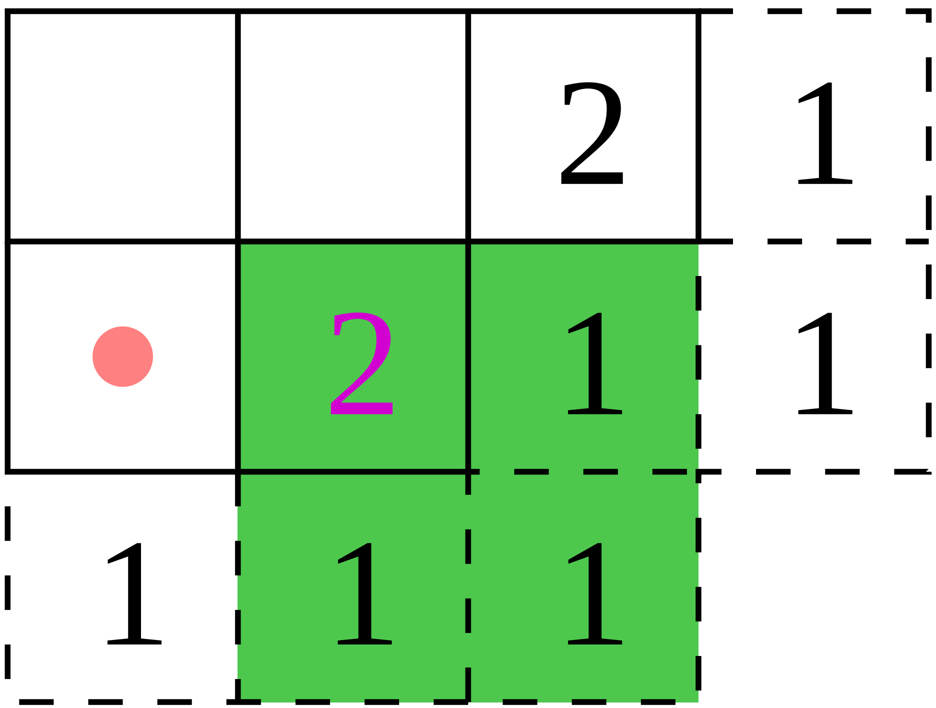
Determinantal algorithm

Suppose all squares to the southeast of t have been filled. Insert into t the number $\boldsymbol{n_t}$ so that $\det M_t = 1$.



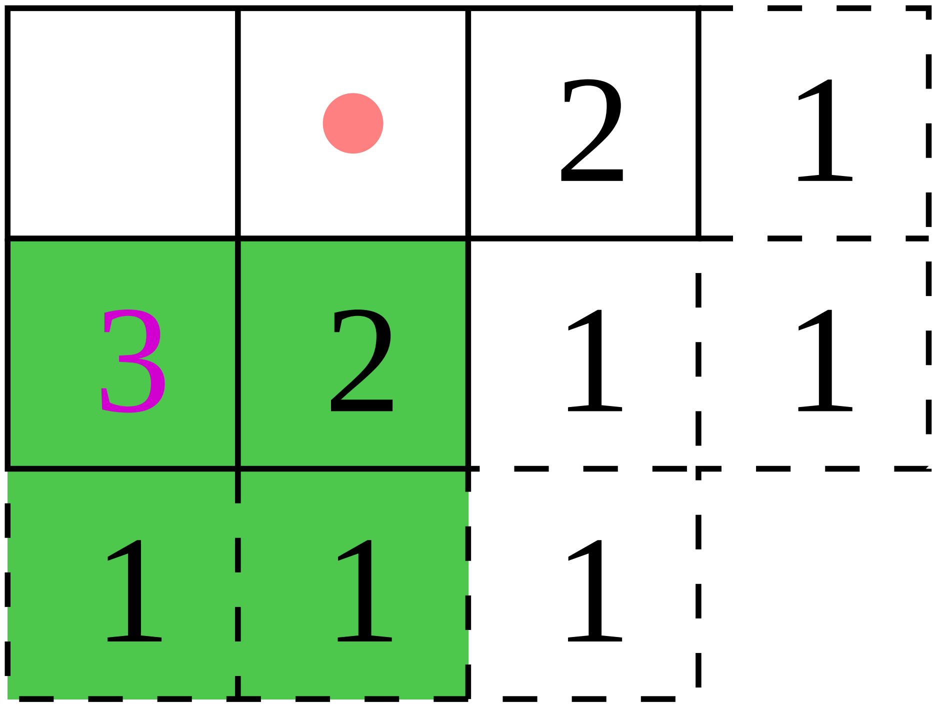
Determinantal algorithm

Suppose all squares to the southeast of t have been filled. Insert into t the number $\boldsymbol{n_t}$ so that $\det M_t = 1$.



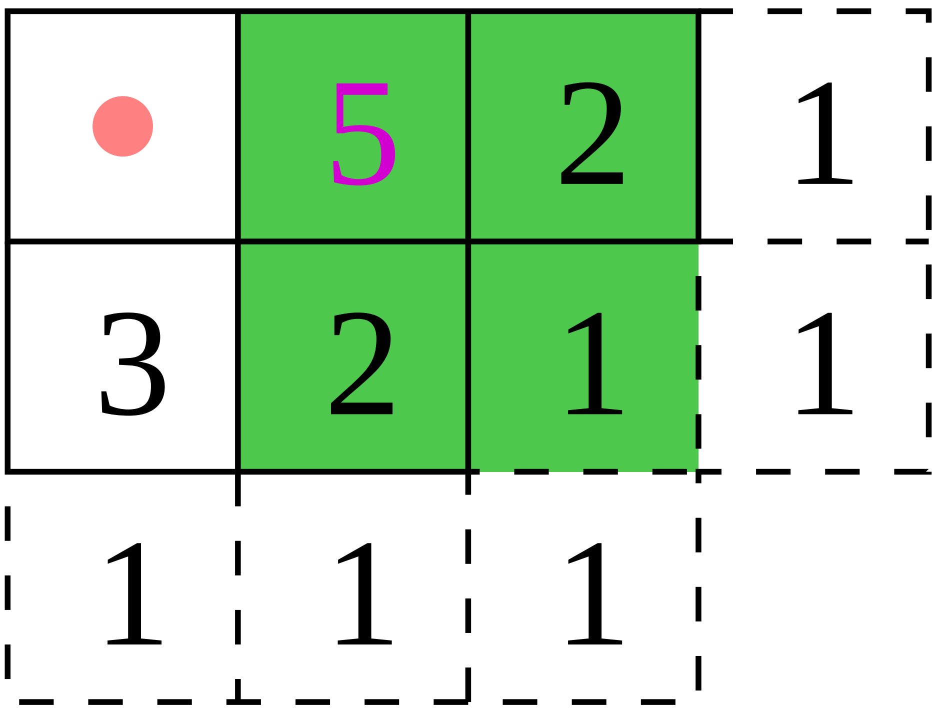
Determinantal algorithm

Suppose all squares to the southeast of t have been filled. Insert into t the number $\boldsymbol{n_t}$ so that $\det M_t = 1$.



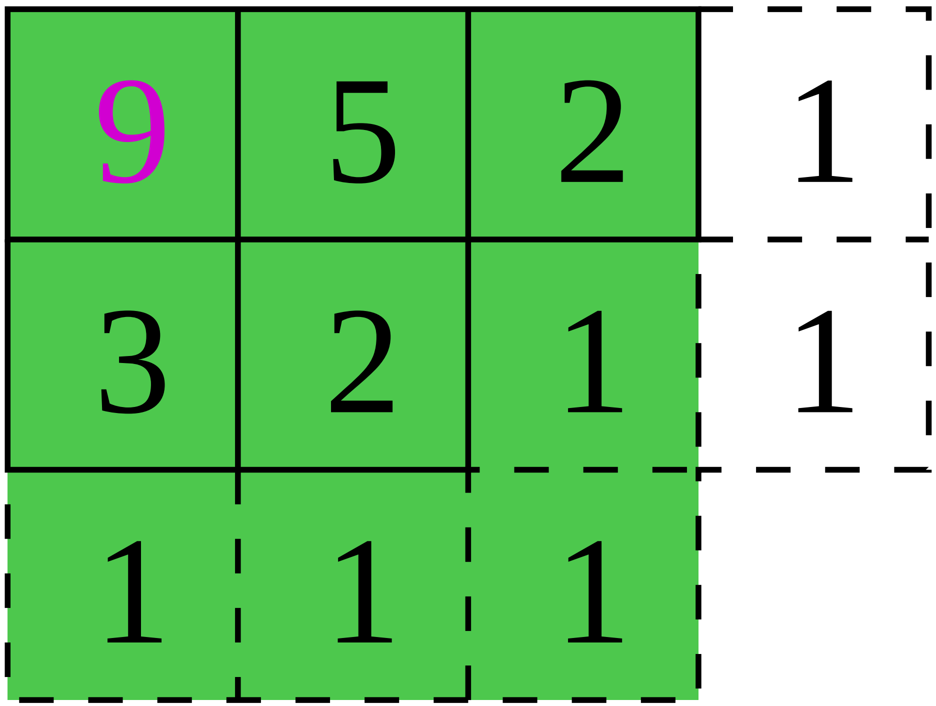
Determinantal algorithm

Suppose all squares to the southeast of t have been filled. Insert into t the number $\boldsymbol{n_t}$ so that $\det M_t = 1$.



Determinantal algorithm

Suppose all squares to the southeast of t have been filled. Insert into t the number $\boldsymbol{n_t}$ so that $\det M_t = 1$.



Uniqueness

Easy to see: the numbers n_t are well-defined and unique.

Uniqueness

Easy to see: the numbers n_t are well-defined and unique.

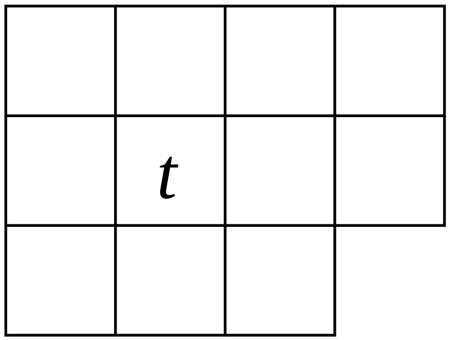
Why? Expand $\det M_t$ by the first row. The coefficient of n_t is 1 by induction.

$\boldsymbol{\lambda(t)}$

If $t \in \lambda$, let $\boldsymbol{\lambda(t)}$ consist of all squares of λ to the southeast of t.

$\boldsymbol{\lambda(t)}$

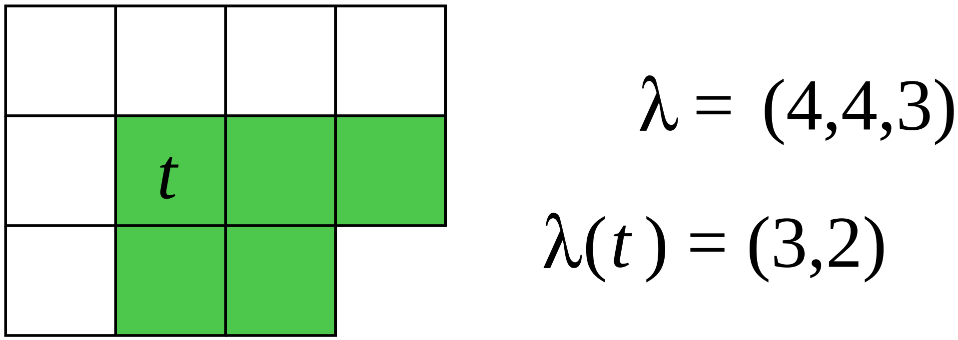
If $t \in \lambda$, let $\boldsymbol{\lambda(t)}$ consist of all squares of λ to the southeast of t.



$\lambda = (4,4,3)$

$\boldsymbol{\lambda(t)}$

If $t \in \lambda$, let $\boldsymbol{\lambda(t)}$ consist of all squares of λ to the southeast of t.



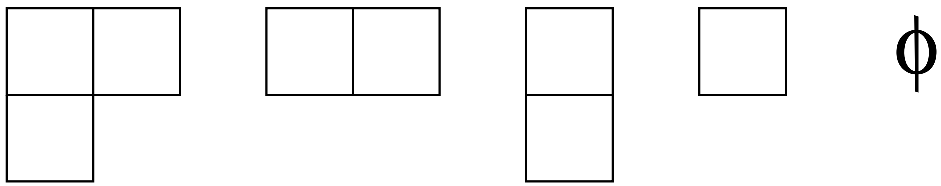
u_λ

$$u_\lambda = \#\{\mu : \mu \subseteq \lambda\}$$

u_λ

$$u_\lambda = \#\{\mu : \mu \subseteq \lambda\}$$

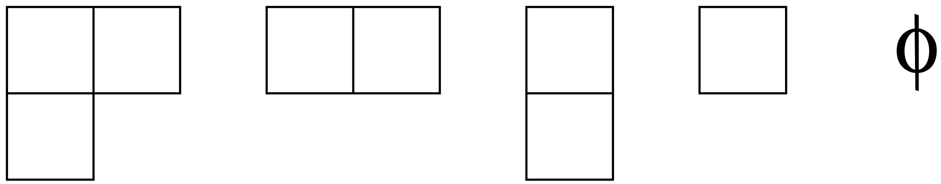
Example. $u_{(2,1)} = 5$:



u_λ

$$u_\lambda = \#\{\mu \,:\, \mu \subseteq \lambda\}$$

Example. $u_{(2,1)} = 5$:



There is a determinantal formula for u_λ, due essentially to **MacMahon** and later **Kreweras** (not needed here).

Carlitz-Scoville-Roselle theorem

- **Berlekamp** (1963) first asked for n_t (mod 2) in connection with a coding theory problem.
- **Carlitz-Roselle-Scoville** (1971): combinatorial interpretation of n_t (over $\mathbb{Z}$).

Carlitz-Scoville-Roselle theorem

- **Berlekamp** (1963) first asked for n_t (mod 2) in connection with a coding theory problem.
- **Carlitz-Roselle-Scoville** (1971): combinatorial interpretation of n_t (over $\mathbb{Z}$).

Theorem. $n_t = u_{\lambda(t)}$

Carlitz-Scoville-Roselle theorem

- **Berlekamp** (1963) first asked for n_t (mod 2) in connection with a coding theory problem.
- **Carlitz-Roselle-Scoville** (1971): combinatorial interpretation of n_t (over $\mathbb{Z}$).

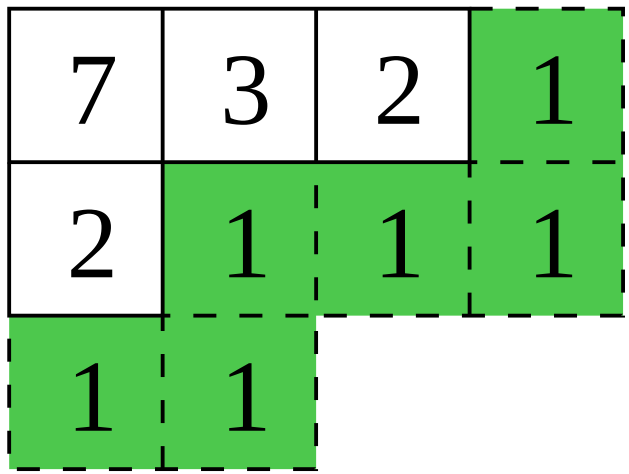
Theorem. $n_t = u_{\lambda(t)}$

Proofs. 1. Induction (row and column operations).

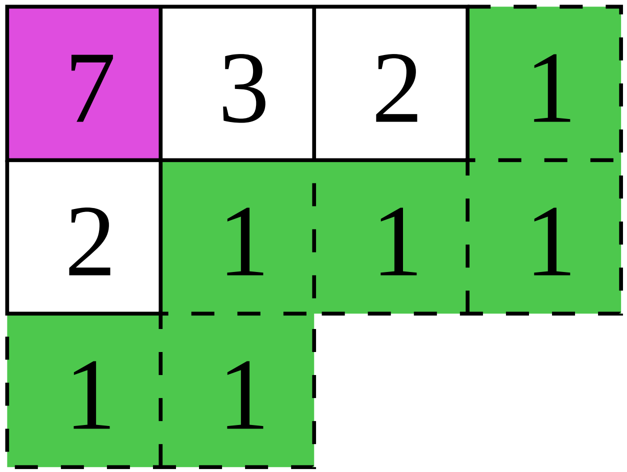
2. Nonintersecting lattice paths.

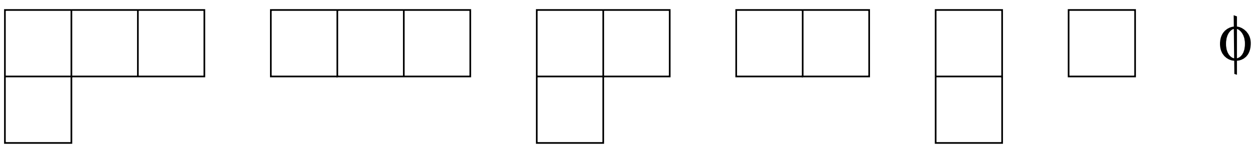
$\square$

An example



An example



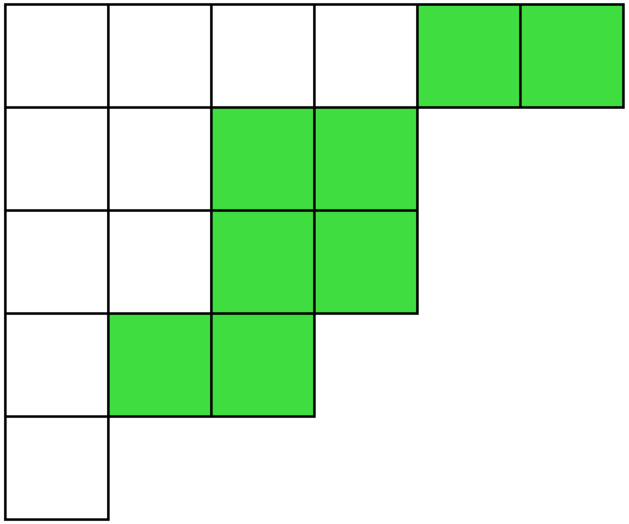


A q-analogue

Weight each $\mu \subseteq \lambda$ by $q^{|\lambda/\mu|}$.

A q-analogue

Weight each $\mu \subseteq \lambda$ by $q^{|\lambda/\mu|}$.

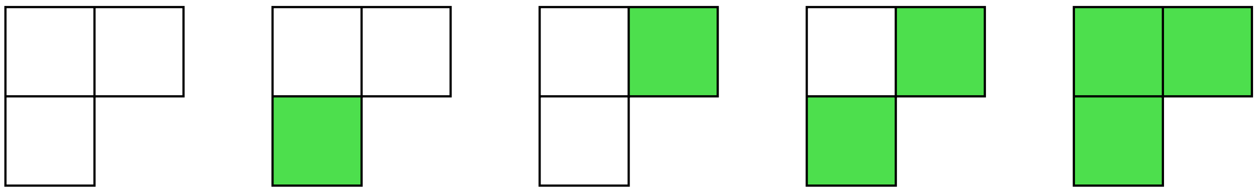


$$\lambda = 64431, \quad \mu = 42211, \quad q^{\lambda/\mu} = q^8$$

$u_\lambda(q)$

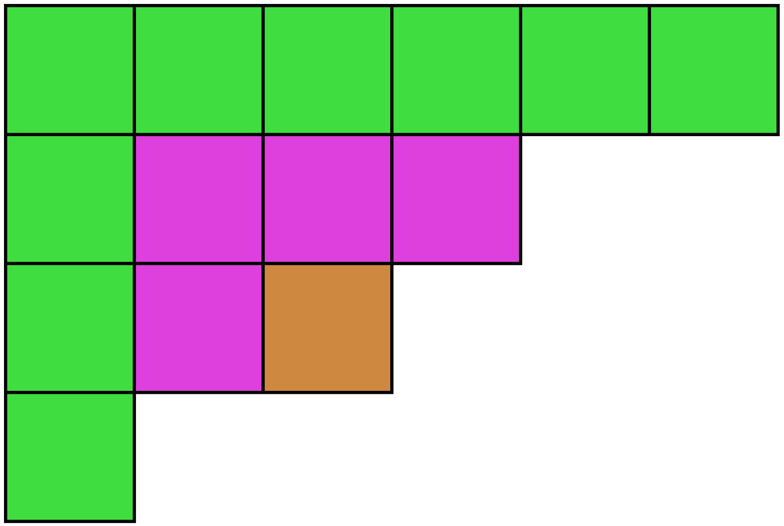
$$u_\lambda(q) = \sum_{\mu \subseteq \lambda} q^{|\lambda/\mu|}$$

$$u_{(2,1)}(q) = 1 + 2q + q^2 + q^3 :$$



Diagonal hooks

$$d_i(\lambda) = \lambda_i + \lambda'_i - 2i + 1$$



$$d_1 = 9, \quad d_2 = 4, \; d_3 = 1$$

Main result (with C. Bessenrodt)

Theorem. *M_t has an SNF over $\mathbb{Z}[q]$. Write $d_i = d_i(\lambda_t)$. If M_t is a $(k+1)\times(k+1)$ matrix then M_t has SNF*

$$\mathrm{diag}(1, q^{d_k}, q^{d_{k-1}+d_k}, \dots, q^{d_1+d_2+\cdots+d_k}).$$

Main result (with C. Bessenrodt)

Theorem. *M_t has an SNF over $\mathbb{Z}[q]$. Write $d_i = d_i(\lambda_t)$. If M_t is a $(k+1)\times(k+1)$ matrix then M_t has SNF*

$$\mathrm{diag}(1, q^{d_k}, q^{d_{k-1}+d_k}, \dots, q^{d_1+d_2+\cdots+d_k}).$$

Corollary. $\det M_t = q^{\sum i d_i}$.

Main result (with C. Bessenrodt)

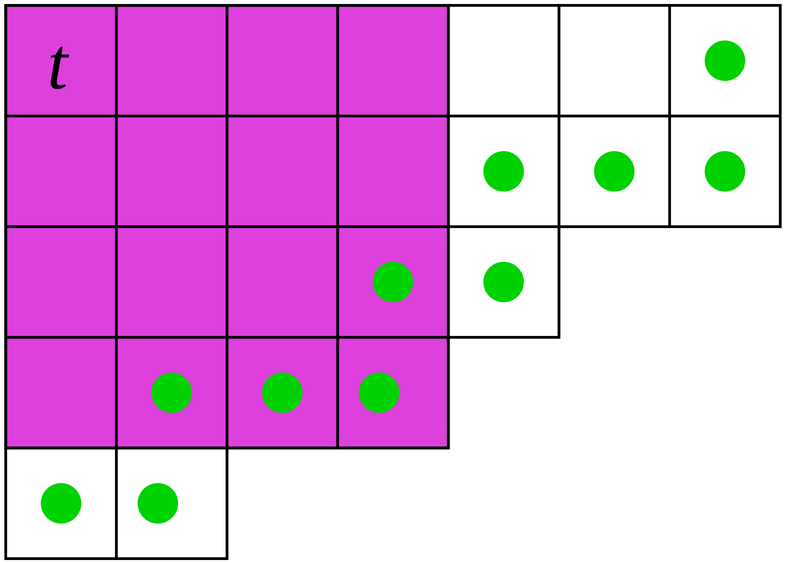
Theorem. *M_t has an SNF over $\mathbb{Z}[q]$. Write $d_i = d_i(\lambda_t)$. If M_t is a $(k+1)\times(k+1)$ matrix then M_t has SNF*

$$\mathrm{diag}(1, q^{d_k}, q^{d_{k-1}+d_k}, \ldots, q^{d_1+d_2+\cdots+d_k}).$$

Corollary. $\det M_t = q^{\sum i d_i}$.

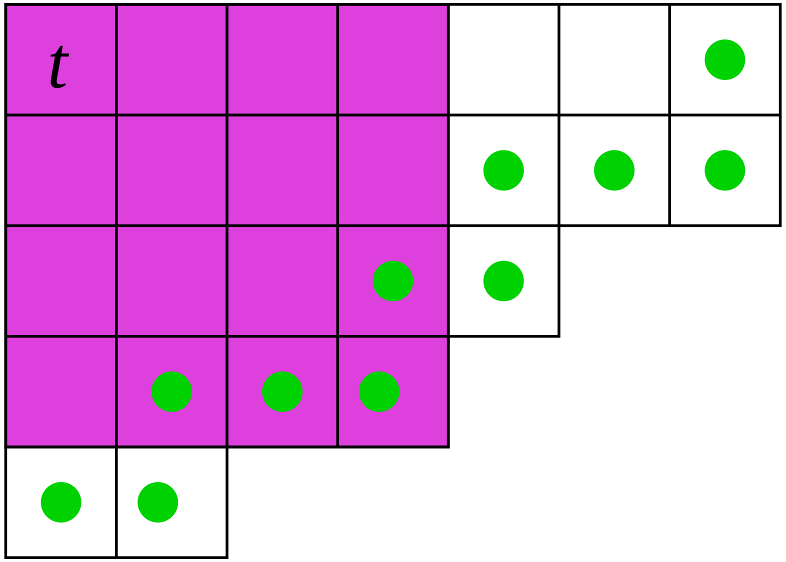
Note. There is a multivariate generalization.

An example



$$\lambda = 6431, \quad d_1 = 9, \quad d_2 = 4, \quad d_3 = 1$$

An example



$$\lambda = 6431, \quad d_1 = 9, \quad d_2 = 4, \quad d_3 = 1$$

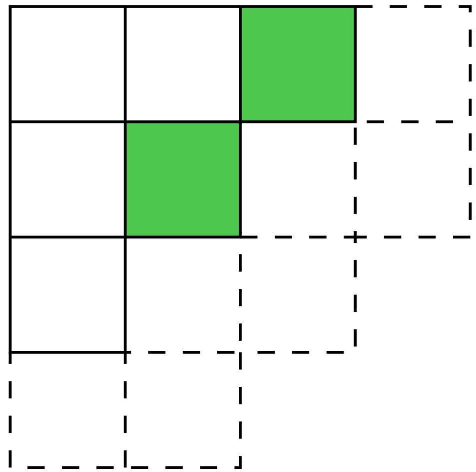
$$\text{SNF of } M_t: \ (1, q, q^5, q^{14})$$

A special case

Let λ be the **staircase** $\boldsymbol{\delta_n} = (n-1, n-2, \ldots, 1)$.

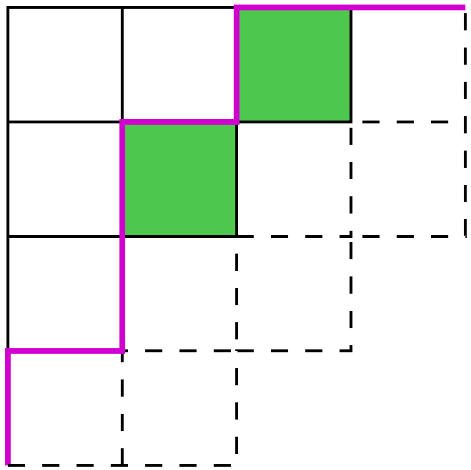
A special case

Let λ be the **staircase** $\boldsymbol{\delta_n} = (n-1, n-2, \ldots, 1)$.



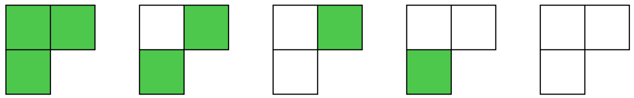
A special case

Let λ be the **staircase** $\delta_n = (n-1, n-2, \ldots, 1)$.



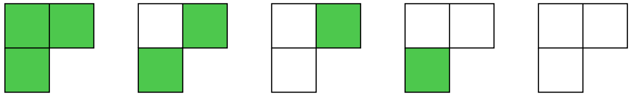
$u_{\delta_{n-1}}(q)$ counts Dyck paths of length $2n$ by (scaled) area, and is thus the well-known q-analogue $C_n(q)$ of the Catalan number C_n.

A q-Catalan example



$$C_3(q) = q^3 + q^2 + 2q + 1$$

A q-Catalan example

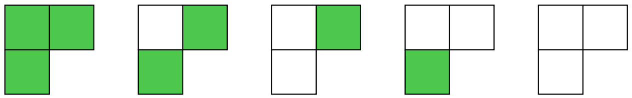


$$C_3(q) = q^3 + q^2 + 2q + 1$$

$$\begin{vmatrix} C_4(q) & C_3(q) & 1+q \\ C_3(q) & 1+q & 1 \\ 1+q & 1 & 1 \end{vmatrix} \overset{\text{SNF}}{\sim} \operatorname{diag}(1, q, q^6)$$

since $d_1(3,2,1) = 1$, $d_2(3,2,1) = 5$.

A q-Catalan example



$$C_3(q) = q^3 + q^2 + 2q + 1$$

$$\begin{vmatrix} C_4(q) & C_3(q) & 1+q \\ C_3(q) & 1+q & 1 \\ 1+q & 1 & 1 \end{vmatrix} \overset{\text{SNF}}{\sim} \operatorname{diag}(1, q, q^6)$$

since $d_1(3,2,1) = 1$, $d_2(3,2,1) = 5$.

- q-Catalan determinant previously known
- SNF is new

Ramanujan

$\sum_{n\geq 0} C_n(q)x^n =$

$$\cfrac{1}{1-\cfrac{x}{1-\cfrac{qx}{1-\cfrac{q^2x}{1-\cdots}}}}.$$

Ramanujan

$$\sum_{n\geq 0} C_n(q)x^n =$$

$$\cfrac{1}{1-\cfrac{x}{1-\cfrac{qx}{1-\cfrac{q^2x}{1-\cdots}}}}.$$

THE END