7. Linear Extensions, Polytopes and Counting

7.1. Poset polytopes

$\boxed{\text{The order polytope}}$

(Order)-homomorphism is an order preserving map $f: P \to Q$
$\boxed{x \leq y \Rightarrow f(x) \leq f(y)}$

$P = (X, \leq)$ finite

$\Rightarrow \quad \mathcal{O}(P) = \mathrm{Hom}(P, [0,1]) \subseteq \mathbb{R}^X$

compact convex in fact a polytope

the order polytope

More down to earth description

$\mathcal{O}(P) = \{ v \in [0,1]^X : x \leq y \Rightarrow v_x \leq v_y \}$

Inequalities : $0 \leq v_x$; $v_x \leq 1$, $v_x \leq v_y$

which are essential?

- $0 \leq v_x$ for $x \in \mathrm{Min}(P)$
- $v_x \leq 1$ for $x \in \mathrm{Max}(P)$
- $v_x \leq v_y$ whenever $x \lessdot y$ cover

Remark: Let $\hat{P}$ be P with a global $\mathbb{0}$ and a global $\mathbb{1}$. Let $X = [n]$ then

$\hat{\mathcal{O}}(P) = \{ (v_0 \ldots v_{n+1}) : v_0 = 0, v_{n+1} = 1, \boxed{x \lessdot y \Rightarrow v_x \leq v_y} \} \subseteq \{0\} \times \mathbb{R}^X \times \{1\}$

unified description of the facets of $\mathcal{O}(P)$

Generic point $v = (v_1 \dots v_n)$

$\Rightarrow \exists \pi \in S_n : v_{\pi(1)} < v_{\pi(2)} < \dots < v_{\pi(n)}$

π is a linear extension of P

and $\forall \pi \in \mathcal{L}(P)$ $\quad \mathcal{O}(\pi) \subseteq \mathcal{O}(P)$

$$\mathcal{O}(P) = \bigcup_{\pi \in \mathcal{L}(P)} \mathcal{O}(\pi)$$

almost disjoint union $\mathcal{O}(\pi) \cap \mathcal{O}(\delta)$ is a subset of $\bigcup_{i \|_P j} H_{ij}$ where $H_{ij} = \{v : v_i = v_j\}$

not full dimensional in $\mathbb{R}^x$ no volume

$$vol(\mathcal{O}(P)) = \sum_{\pi \in \mathcal{L}(P)} vol(\mathcal{O}(\pi))$$

Lemma: $\pi, \delta \in S_n \Rightarrow vol(\mathcal{O}(\pi)) = vol(\mathcal{O}(\delta))$

proof: consider the permutation matrix

$A_{\delta \circ \pi^{-1}} : \mathcal{O}(\pi) \to \mathcal{O}(\delta)$ (bijective)

The map is linear and $\det(A_{\delta\pi^{-1}}) = \pm 1$

$\Rightarrow$ volume preserving $\square$

Lemma: $vol(\mathcal{O}(\pi)) = \frac{1}{n!} \quad \forall \pi \in S_n$

proof: $[0,1]^n = \bigcup_{\pi \in S_n} \mathcal{O}(\pi) = 1$

disjoint upto H_{ij} no volume

$n!$ simplices of same volume

Def: $e(P) = |\mathcal{L}(P)|$ 3

THM: $\boxed{\mathrm{Vol}(\mathcal{O}(P)) = \frac{e(P)}{n!}}$

Remark: connects combinatorial counting problem to geometric measure problem. Fertile in both directions

- $e(P)$ #P complete $\rightarrow$ hardness of volume of polytopes
- will see how geometry helps order theory

Prop: Corners of $\mathcal{O}(P)$ are the characteristic vectors of up-sets of P

proof: $\mathcal{U} \subseteq [n]$ an up-set in P

$$e_i^{\mathcal{U}} = \begin{cases} 1 & i \in \mathcal{U} \\ 0 & i \notin \mathcal{U} \end{cases} \Rightarrow e^{\mathcal{U}} \in \mathcal{O}(P)$$

$e^{\mathcal{U}}$ a corner of $[0,1]^n \Rightarrow$ corner of $\mathcal{O}(P)$

$v \in \hat{\mathcal{O}}(P)$

claim v can be written as convex combination of $\{e^{\mathcal{U}} : \mathcal{U} \text{ upset in } P\}$

$v = (v_0 v_1 \dots v_n v_{n+1})$ $v_{\pi_0} \leq v_{\pi_1} \leq \dots \leq v_{\pi_n} \leq v_{\pi_{n+1}}$

with $\pi_0 = \hat{0}$ $\pi_{n+1} = \hat{1}$

For $1 \leq j \leq n+1$

$$\mathcal{U}_j = \{\pi_j \pi_{j+1} \dots \pi_{n+1}\}$$

$$\Rightarrow \sum_1^{n+1} \underbrace{(v_{\pi(j)} - v_{\pi(j-1)})}_{\lambda_j} e^{\mathcal{U}_j} = v$$

- $\lambda_j \geq 0$
- $\sum \lambda_j = v_{\pi_{n+1}} - v_{\pi_0} = 1 - 0 = 1$

it is a convex combination □

The chain polytope

$$\mathcal{C}(P) = \{ u \in \mathbb{R}^n_+ : u^T e^C \leq 1 \text{ for every } C \text{ chain} \}$$

Prop: The corners of $\mathcal{C}(P)$ are the characteristic vectors of antichains of P

Rem: $\mathcal{O}(P)$ and $\mathcal{C}(P)$ have the same number of corners.

Proof: A antichain $\Rightarrow e^A \in \mathcal{C}(P) \subseteq [0,1]^n$

e^A a corner of $[0,1]^n \Rightarrow$ corner of $\mathcal{C}(P)$

Let $u \in \mathcal{C}(P)$ u is a weighting on $[n]$ such that $u(C) \leq 1$ $\forall$ chains

weighted canonical antichain decomp.

Formally: $w_0 = u$

$S_0 = \{ i : w_0(i) \neq 0 \}$

$A_1 = \mathrm{Min}(S_0) \quad \lambda_1 = \min\{ u_i : i \in A_1 \} \geq 0$

$w_1 = w_0 - \lambda_1 e^{A_1} \quad S_1 = \{ i : w_1(i) \neq 0 \}$

Homework 1st week weighted version of dual of Dilworth

Iterate given w_{i-1}, $S_{i-1} = \{ j : w_{i-1}(j) \neq 0 \}$

$A_i \quad \lambda_i \quad w_i = w_{i-1} - \lambda_i e^{A_i} \quad S_i$

$|S_i| < |S_{i-1}| \Rightarrow$ terminates with $S_t = \emptyset$

backwards build a chain intersecting each A_i $\Rightarrow \lambda = \sum \lambda_i \leq 1$ add $(1-\lambda) e^{\emptyset}$

$\Rightarrow$ convex combination $\square$

THM (Stanley 86) $vol(\mathcal{O}(P)) = vol(\mathcal{C}(P))$

Remark: Stanleys proof via Erhart polynomial
$E_P(k) = |kP \cap \mathbb{Z}^d|$ they are the same
Leading coefficient is the volume

Proof: A map $\phi : \mathcal{O}(P) \to \mathcal{C}(P)$

$x \in \mathcal{O}(P)$ we define $y = \phi(v)$
bottom up (along a lin ext)

$$y_j = \begin{cases} x_j & \text{if } j \in Min(P) \\ x_j - \max_{i<j} x_i & \text{otherwise} \end{cases}$$

Claim $(y_1 \dots y_n) \in \mathcal{C}(P)$ Let $i_1 < i_2 < \dots < i_k$ be a max chain

$\Rightarrow y_{i_1} = x_{i_1}$ $\quad y_{i_2} = x_{i_2} - \max_{i<i_2} x_i \leq x_{i_2} - x_{i_1}$

$y_{i_k} \leq x_{i_k} - x_{i_{k-1}}$

$\Rightarrow \sum y_{i_j} \leq x_{i_k} \leq 1$

Given $y \in \mathcal{C}(P)$ we can define x with $\phi(x) = y$: bottom up

$$x_j = \begin{cases} y_j & j \in Min \\ y_j + \max_{i<j} x_i & \end{cases}$$

$\in \mathcal{O}(P)$ ✓
inverse ✓

On $\mathcal{O}(\pi)$ the index j^* $\boxed{\text{such that } y_j = x_j - x_{j^*}}$ is constant

$\Rightarrow$ can write ϕ as a linear map
$M_\pi x = \phi(x)$

$$M_\pi = \begin{array}{c} \\ \\ \\ j \\ \\ \\ \\ \end{array} \begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & -1 & & 1 & & \\ & & & & 1 & \\ & & & & & 1 \end{pmatrix} \implies \det M_\pi = 1$$
$$\begin{array}{cccccc} \quad & j^* & \quad & j & \quad & \quad \end{array}$$

map triangulation of $\mathcal{O}(P)$ to a triang. of $\mathcal{C}(P)$ linear on the simplices of the triangulation.

The same map: $\mathcal{U}$ upset $\mathcal{U} \to \operatorname{Min} \mathcal{U}$ antichain.

corner to corner. $x \in \mathcal{O}(P)$ generic

$$[n] = \mathcal{U}_0 \supseteq \mathcal{U}_1 \supseteq \dots \supseteq \mathcal{U}_n = \emptyset$$
$$A_0 < A_1 < \dots < A_n$$

linear map between simplices. $\square$

Convex corners and antiblockers

Def: A convex corner is a convex - compact - fulldimensional - down set in $\mathbb{R}_+^X$ ← nonnegative quadrant.

Obs: $\mathcal{C}(P)$ is a convex corner

Def: The antiblocker of a convex corner K is

$$K^* = \{ x \in \mathbb{R}_+^X : x^T y \leq 1 \ \forall y \in K \}$$

Obs: K^* is a convex corner

Obs: $K \subseteq K^{**} = (K^*)^*$

Prop: $K = K^{**}$

proof: Suppose $\exists z \in K^{**}$ $z \notin K$. K compact convex

$\Rightarrow \exists$ hyperplane H separating K and z — Rödl ⇒

K a down set $H = \{ x^T w = 1 \}$ with $\boxed{w \in \mathbb{R}^X_+}$

$\Rightarrow x^T w \leq 1 \ \forall x \in K \Rightarrow w \in K^* \Rightarrow z \notin K^{**}$ ↯ □

$$\mathcal{C}^*(P) = \{ z \in \mathbb{R}^X_+ : z^T y \leq 1 \ \forall y \in \mathcal{C}(P) \}$$

$$= \{ z \in \mathbb{R}^X_+ : z^T e^A \leq 1 \ \forall A \text{ antichain} \}$$

e^A — corners of $\mathcal{C}(P)$

$$=: \mathcal{A}(P)$$

The antichain polytope. Note that $e^C \in \mathcal{A}(P) \ \forall C$ chain

Now $\mathcal{A}^*(P) = \{ x \in \mathbb{R}^X_+ : x^T y' \leq 1 \ \forall y' \text{ corner of } \mathcal{A}(P) \}$

$\Vert$ — shows that $\mathcal{A}(P)$ has no add. corners

$$\mathcal{C}(P) = \{ x \in \mathbb{R}^X_+ : x^T e^C \leq 1 \ \forall C \text{ a chain} \}$$

Cor: The corners of $\mathcal{A}(P)$ are the characteristic vectors of chains of P

A thm from convex geometry (Saint-Raymond)

K convex corner $\Rightarrow \mathrm{vol}(K)\,\mathrm{vol}(K^*) \geq \frac{1}{n!}$

Now let P be 2 dim $\Rightarrow \mathcal{A}(P) = \mathcal{C}(\bar{P})$ conjugate

$$\left.\frac{1}{n!} \leq \mathrm{vol}(\mathcal{C}_P)\,\mathrm{vol}(\mathcal{A}_P) = \frac{e(P)}{n!}\,\frac{e(\bar{P})}{n!}\right\} \Rightarrow \boxed{e(P) \cdot e(\bar{P}) \geq n!}$$

⇒

Betrachte Gerade $l : \langle 0, z \rangle$
und den Austrittspunkt p aus K

$\exists$ Tangentialebene T_p an K in p

- T_p trennt
- T_p schneidet Achse x_i in Pkt $t_i > 0$

$\Rightarrow$ Normale w ist nicht-negativ

In der Vorlesungsvorbereitung von 2010? 2016
Seiten 96 - 99 sind noch 2
weitere Teile aus Brightwell
ausgearbeitet, insbes ein rein
komb Beweis für
$e(P)\, e(\bar{P}) \geq n!$

$\boxed{\text{Upper bound for } e(P)}$

Proposition: $P = (X, <)$ a poset, $b \in \mathbb{R}_+^X$
such that $b^T e^A \le 1$ $\forall A$ antichain
(ie $b \in \mathcal{A}(P)$)

$$\Rightarrow e(P) \le \prod_{x \in X} 1/b_x$$

---

proof 1. Take the generic alg for lin ext.
When it comes to choose x_{i+1} from $\mathrm{Min}(P_i) = M_i$
Take $x \in \mathrm{Min}(P_i)$ with prob $\boxed{b_x / \sum_{y \in M_i} b_y} \ge b_x$
the denominator ≤ 1 M_i antichain!
Consider $L \in \mathcal{L}(P)$

$L = x_1 \dots x_n$

$$\mathrm{Prob}(L) = \prod_{j=1}^{n} \frac{b_{x_j}}{\sum_{y \in M_{j-1}} b_y} \ge \prod b_{x_j} = \beta \quad \text{indep of } L$$

$$\Rightarrow 1 = \sum \mathrm{Prob}(L) \ge e(P) \cdot \beta$$

$\square$

Proof 2. $b \in \mathcal{A}(P) \Rightarrow \forall a \in \mathcal{E}(P) \quad a^T b \le 1$

$$\Rightarrow \mathcal{E}(P) \subseteq \{x \in \mathbb{R}_+^X : x^T b \le 1\} = S$$

simplex with corners
$0 \quad \frac{1}{b_i} e_i \quad i \in X$

$$\mathrm{Vol}(S) = \frac{1}{n!} \cdot \prod \frac{1}{b_x}$$

$$\Rightarrow e(P) \underset{\text{Stanley's Thm}}{=} n!\, \mathrm{Vol}(\mathcal{E}(P)) \le \prod \frac{1}{b_x}$$

$\square$

Optimizing b

The max point of $A(P)$ yields best bound on $e(P)$

Def: $K \subseteq \mathbb{R}_+^X$ convex corner

$$\psi(K) = \max_{a \in K} \prod a_x$$

the maximizing point is the optimal point of K

THM: (K, K^*) an antiblocking pair with optimal points a and b resp.

$$\Leftrightarrow \quad a_x b_x = \frac{1}{n} \ \forall x \qquad \left(\psi(K) \cdot \psi(K^*) = \left(\frac{1}{n}\right)^n \right)$$

Proof: ineq of arithm and geom mean

$$\prod c_x \leq \left[\frac{1}{n} \sum c_x\right]^n$$

$$\left(\prod a_x b_x\right)^{1/n} \leq \frac{1}{n} \sum a_x b_x \leq \frac{1}{n}$$

$a \in K$, $b \in K^*$ antiblock

$$\Rightarrow \boxed{\psi(K)\,\psi(K^*) \leq \left(\frac{1}{n}\right)^n}$$

$\boxed{\text{Claim } \exists\, a \in K \text{ and } b \in K^* \text{ with } a_x b_x = \frac{1}{n} \ \forall x}$

If they exist they are optimal points.

Let a be optimal for K : $\prod a_x = \psi(K)$

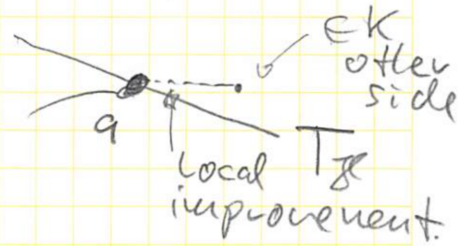
Hypersurface $\mathcal{H} = \{ c \in \mathbb{R}_+^n : \prod c_x = \psi(K) \}$
is kind of a hyperbola

partial derivat. $\dfrac{\partial(\prod x_j)}{\partial x_i} = \dfrac{\prod x_j}{x_i}$

$\Rightarrow$ gradient at the point a is a multiple of $(1/a_1, 1/a_2 \ldots 1/a_n) =: a^{-1}$

$\Rightarrow$ the tangent plane $T_{\mathcal{H}}(a) = \{ y : y^T b = 1 \}$, $b = \frac{1}{n} a^{-1}$

optimality of a and convexity of K imply that



$$K \subseteq \{y : y^T b \leq 1\}$$

$$\Rightarrow \quad b \in \mathcal{K}^* \qquad \square$$

THM For all P

can be considered a small error $n! \sim n^n/e^n$

$$n!\, \Psi(\mathcal{C}(P)) \leq e(P) \leq n^n\, \Psi(\mathcal{C}(P))$$

Proof • a optimal point of $\mathcal{C}(P)$

$\Rightarrow$ $\mathcal{C}(P)$ contains a box with volume $\Psi(\mathcal{C}(P))$

$$\Rightarrow \quad \Psi(\mathcal{C}(P)) \leq vol(\mathcal{C}(P)) = \frac{e(P)}{n!}$$

• $$e(P) \leq \min_{b \in \mathcal{A}(P)} \prod \frac{1}{b_x} = \frac{1}{\Psi(\mathcal{A}(P))} \leq n^n\, \Psi(\mathcal{C}(P))$$

prev. Thm.

Good approximations of $\Psi(\mathcal{C}(P))$ are useful for approximations of $e(P)$ (Kahn + Kim 92)

Sidorenko's inequality:

Thm: If A is an antichain in P then

$$e(P) \geq \sum_{x \in A} e(P-x)$$

with equality if $A \cap C \neq \emptyset$ $\forall C$ maximal chain

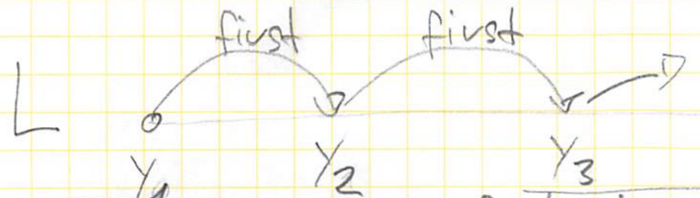
---

Rem:
- obvious cases with equality $A = \mathrm{Min}(P)$ or $A = \mathrm{Max}(P)$
- with induction this implies that $e(P)$ is a comparability invariant

---

Proof 1 [Edelm. Hibi Stanley '89] Via chain pushing

With L we associate its greedy chain $g(L)$



$L = x_1 \dots x_n$

$y_1 = x_1$ if

$y_i = x_j \Rightarrow y_{i+1} = x_k$ with $k > j$ min st $x_j < x_k$

Obs:
- This is a maximal chain.
- Every maximal chain is $g(L)$ for some L

$x \in A$ $\quad \mathcal{L}_x = \{ L \in \mathcal{L}(P) : x \in g(L) \}$

$\underset{\in A}{x \neq y}$: $\mathcal{L}_x \cap \mathcal{L}_y = \emptyset \Rightarrow |\mathcal{L}(P)| \geq \sum_{x \in A} |\mathcal{L}_x|$

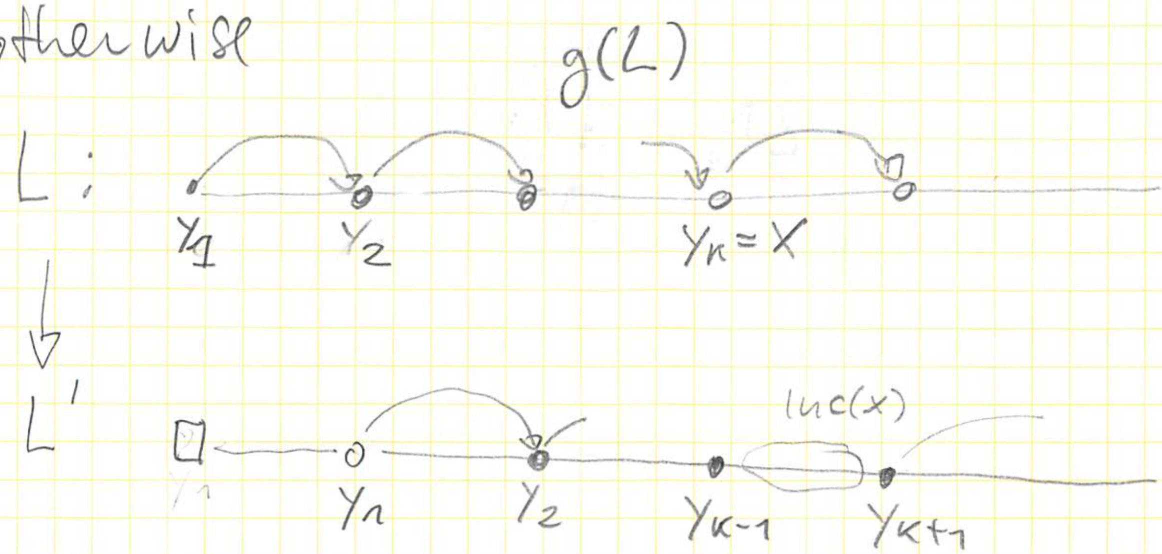
equality iff each max chain contains an element of A

Claim $|\mathcal{L}_x| = e(P-x)$

We build a bijection $\mathcal{L}_x \leftrightarrow \mathcal{L}(P-x)$

If $x \in Min(P)$ the bijection is trivial

otherwise



- L' = promotion(L) is a lin extension of $P - x$
- y_{k-1} is latest predecessor of x in L'
- $y_{k-1}\, y_{k-2} \dots y_2\, y_1$ is backwards greedy-chain of y_{k-1} $\square$

---

proof 2 [Sidorenko '91] <u>Via network flow</u>

Network N_P vertices $X \cup \{s,t\}$ $(P = (X, \le))$

Edges: directed

$x \to t$ $x \in Max(P)$

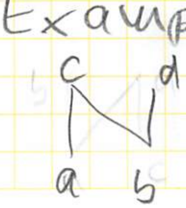
$s \to x$ $x \in Min(P)$

$x \to y$ $x \prec y$ cover

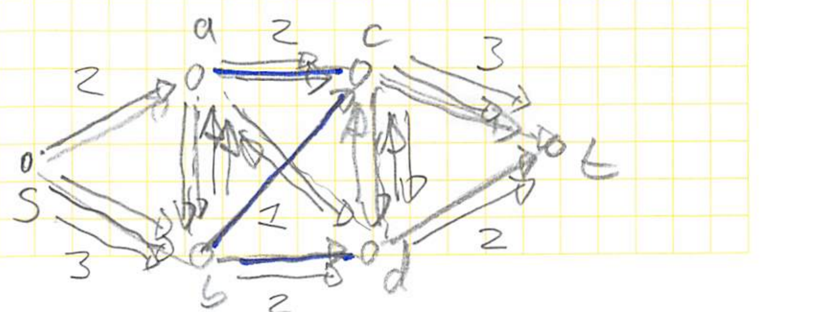
bidirected $x \leftrightarrows y$ $\forall\, x \parallel y$

In N_P we define an $s \to t$ flow λ by superimposing unit flows, one for each $L \in \mathcal{L}(P)$ $L = x_1 \dots x_n$ $s \to x_1 \to x_2 \dots \to x_n \to t$

Example:



a b c d
a b d c
b a c d
b a d c
b d a c



Observations

- $|\lambda| = e(P)$ (total $s \to t$ flow)
- if $x \| y \Rightarrow \lambda(xy) = \lambda(yx)$

$\Rightarrow$ λ is a flow on the network $\hat{N}_P$ restricted to uni-directed edges

Let $\lambda(x) = \sum\limits_{\substack{y \\ y \to x \text{ in } \hat{N}_P}} \lambda(yx)$

Lemma: $\lambda(x) = e(P-x)$

proof: $\lambda(x) = \#(L \in \mathcal{L}(P) : L = \dots yx \dots$ with $y < x)$

$\Rightarrow$ y is latest predecessor of x in L

$\Rightarrow$ $L = \dots yx \dots \leftrightarrow L' = \dots y \dots \in \mathcal{L}(P-x)$ is a bijection

---

The flow in $\hat{N}_P$ can be decomposed into $|\lambda|$ paths each carrying a unit
These paths correspond to maximal chains in P. Let Δ_x be the set of paths containing x

$\Rightarrow$ $x \| y \Rightarrow \Delta_x \cap \Delta_y = \emptyset$

A antichain $\Rightarrow$ $\underset{\substack{\| \\ e(P)}}{|\lambda|} \geq \sum\limits_{x \in A} |\Delta_x| = \sum\limits_{x \in A} \underset{\substack{\| \\ e(P-x)}}{\lambda(x)}$

if each max chain intersects A $\Rightarrow$ partition

$\Rightarrow$ $\lambda = \sum\limits_{x \in A} |\Delta_x|$ $\square$

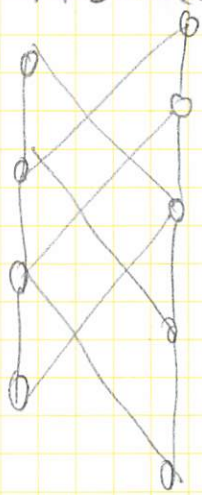
Enumeration – some easy cases

(1) Antichains $e(A_n) = n!$

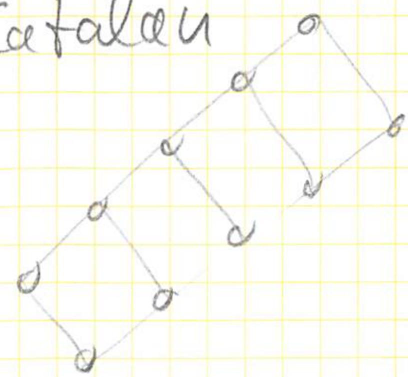
(2) series parallel $P \oplus Q$ parallel $\quad P \odot Q$ serial

$$e(P \odot Q) = e(P) \cdot e(Q)$$
$$e(P \oplus Q) = \binom{n_P + n_Q}{n_P} e(P) \cdot e(Q)$$

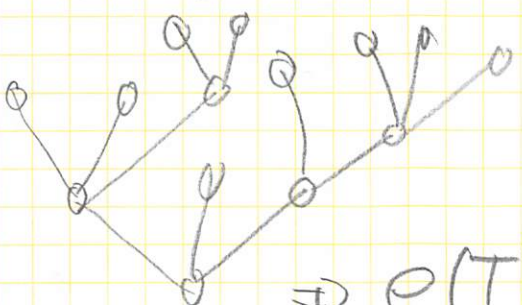
(3) Fibonacci



(4) Catalan



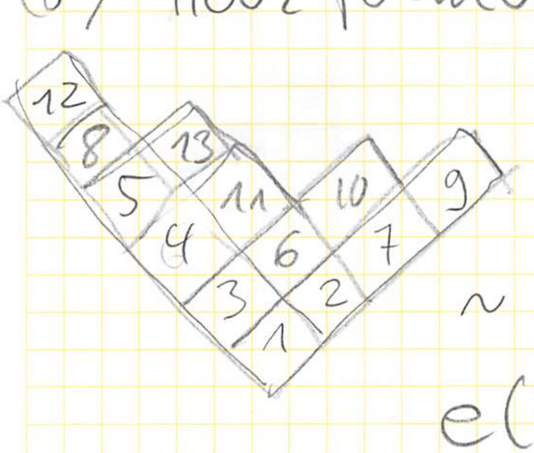
(5) Tree-hook formula



$h_x = |U[x]|$ closed upset of x

$$\Rightarrow e(T) = \frac{n!}{\prod_{x \in T} h_x}$$

(6) Hook formula for Young-Tableaux



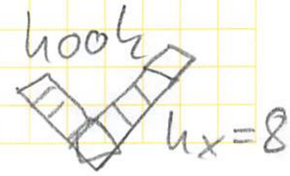
Ferrer's shape: Down set of grid

Young tableaux: Filling monotone in rows + columns

$\sim$ Linear extension of F

$$e(F) = \frac{n!}{\prod_{x \text{ cell}} h_x}$$

h_x hook



7.3 Random generation of linear extensions the Markov chain approach

Markov chains in our context:

Memoryless discrete time discrete space stochastic processes

- finite state space S
- sequence $X_0, X_1, X_2 \ldots$ random events with $X_i \in S$
- $Pr(X_t = a \mid X_{t-1} = b, X_{t-2} = b_2, \ldots X_0 = b_t)$

history

$= Pr(X_t = a \mid X_{t-1} = b) = P_{ab}$

time

$P = (P_{ab})$ transition matrix

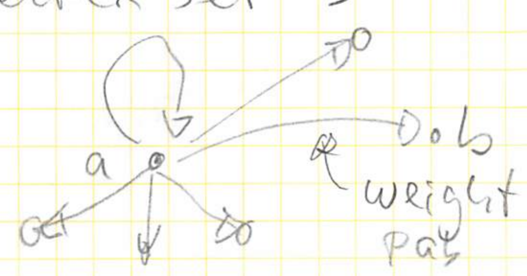
$p_0 \in \mathbb{R}_+^S$ initial distribution

$\Rightarrow P^t p_0$ distribution at time t

Markov chains and random walks

emphasis on the sequence of random events

vertex set S



From $\mathbb{1}^T = \mathbb{1}^T P$ (P is stochastic) we get $\exists\, \pi$ with $\pi = P\pi$ stationary distrib.

II

Def: M is ergodic $\iff$
transition graph is connected + aperiodic
$\iff \forall a,b \in S \ \exists T$ such that $\forall t \geq T$

$$\Pr(X_t = b \mid X_0 = a) > 0$$

Fundamental Theorem

M ergodic $\Rightarrow$

- $\exists$ unique stationary distrib π
  - if M is symmetric ($P_{ab} = P_{ba} \ \forall ab$) $\Rightarrow \pi$ is the uniform distr on S
- $\forall p_0 \quad \lim\limits_{t \to \infty} P^t p_0 = \pi$

Can lead to effective sampling from large state spaces.

→ Example hypercube Q_n: $P_{ab} = \begin{cases} 1/2 & \text{if } a = b \\ 1/2n & \text{if } a = b \oplus e_i \\ 0 & \text{otherwise} \end{cases}$
as a random walk

- how long do we have to stroll around to shake off the initialization bias? mixing time

Total variational distance

A measure for the distance of distributions

$$\|\mu - \pi\|_{TV} := \max_{A \subseteq S} |\mu(A) - \pi(A)|$$

Note that

$$\max_{A \subseteq S} |\mu(A) - \pi(A)| = \frac{1}{2} \sum_{s \in S} |\mu(s) - \pi(s)| = \frac{1}{2} \|\mu - \pi\|_1$$

We are interested in time needed for ε-approx mixing

$$\rightarrow \tau(\varepsilon) = \max_{a \in S} \min_t \left(\|P^t e_a - \pi\|_{TV} \leq \varepsilon \right)$$

Auf einer Tafel und noch ein bisschen Lesen

Couplings

A coupling of Markov chain M with transitions P_{ab} is a pair (X_t, Y_t) of copies of M running in parallel, ie

$$\Pr(X_t = a \mid (X_{t-1} Y_{t-1}) = (b, b')) = P_{ab}$$
$$\Pr(Y_t = a' \mid (X_{t-1} Y_{t-1}) = (b, b')) = P_{a'b'}$$

The coupling Lemma:

Let $(X_t Y_t)$ be a coupling. If $\forall a, b \in S$

$$\Pr(X_T \neq Y_T \mid X_0 = a, Y_0 = b) < \varepsilon$$

then $\tau(\varepsilon) \leq T$

Example I: Consider the Markov chain on Q_n

A coupling: Given (X_i, Y_i)

- choose a position $j \in [n]$ uniformly at rand
- choose a value $v \in \{0,1\}$ uniformly

$X_{i+1} = X_i$ with coordinate j replaced by v
$Y_{i+1} = Y_i$ — " —.

Obs. as soon as all coordinates have been touched $X_t = Y_t$.

The coupling time is given by the coupon collector problem $\sim n \log n$

Example II: consider two copies of with an edge joining top of Q^1 to bottom of Q^2

The prob for edge traversal $M_i \to M_{i+1}$ is $\leq 2^n/n$ $\Rightarrow$ slowly mixing.

Proof of the coupling lemma

Choose $b = Y_0$ according to the stationary distr. and $a = X_0$ arbitrary. Let $A \subseteq S$

$$\Pr(X_T \in A) \geq \Pr((X_T = Y_T) \cap (Y_T \in A))$$
$$= 1 - \Pr((X_T \neq Y_T) \cup (Y_T \notin A))$$
$$\overset{\text{union bound}}{\geq} 1 - \Pr(X_T \neq Y_T) - \Pr(Y_T \notin A)$$
$$= \pi(A) - \varepsilon$$

$$\boxed{\begin{array}{l} \Pr(Y_T \notin A) = \pi(\bar{A}) \\ \pi(A) = 1 - \pi(\bar{A}) \end{array}}$$

From

$\Pr(X_T \in \bar{A}) \geq \pi(\bar{A}) - \varepsilon$ we get

$\Pr(X_T \in A) \leq \pi(A) + \varepsilon$

$$\Rightarrow \underbrace{\max_{A \subseteq S} \left| P_a^T(A) - \pi(A) \right|}_{= \|P_a^T - \pi\|_{TV}} \leq \varepsilon$$

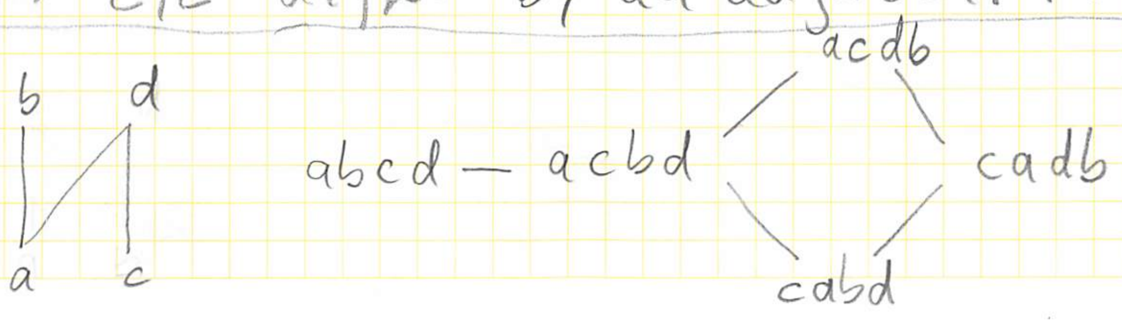
$\square$

Linear extensions

$P = (X, \leq)$ a poset on n-elements

The graph of linear extensions $\mathcal{G}(P)$ has $V = \mathcal{L}(P)$ and $(L, L') \in E$ $\Leftrightarrow$ L, L' differ by an adjacent transposition

Ex



V

$d(L)$ = degree of L in $\mathcal{G}(P)$

The Karzanov-Khachiyan chain (1991)

$X_0 X_1 \ldots$ with $X_i \in \mathcal{L}(P)$

and $$M_{LL'} = \begin{cases} 1/_{2(n-1)} & \text{if } LL' \text{ adj transp.} \\ 1 - \frac{d(L)}{2(n-1)} & \text{if } L = L' \end{cases}$$

! symmetric

KK prove mixing in $O(n^6 \log n)$ using a geometric argument to bound the conductance of $\mathcal{G}(P)$ (order polytope)

Bubley Dyer (1997) use path coupling and a modified chain

$p_1 \ldots p_{n-1}$ probabilities $\sum_1^{n-1} p_i = 1$

For $L = (x_1 x_2 \ldots x_n)$ we let

$$\tau_i(L) = (x_1 \ldots x_{i-1} x_{i+1} x_i x_{i+2} \ldots x_n)$$

Transition prob.

$$M_{LL'} = \begin{cases} p_i & \text{if } L' = \tau_i(L) \in \mathcal{L}(P) \\ 1 - \sum p_i \, \delta[\tau_i(L) \in \mathcal{L}(P)] & \text{if } L = L' \end{cases}$$

~~For the coupling we additionally need~~

~~L_0 an arbitrary reference lin ext.~~

~~we let~~ $$\cancel{\tau_i^0(L) = \begin{cases} \tau_i(L) & \text{if } \tau_i(L) \in \mathcal{L}(P) \text{ and } x_{i+1} <_{L_0} x_i \\ L & \text{otherwise} \end{cases}}$$

~~towards L_0~~

For $c \in \{0,1\}$ Let $\tau_i^c(L) = \begin{cases} \tau_i(L) & \text{if } \tau_i(L) \in \mathcal{L}(P) \text{ and } c=1 \\ L & \text{otherwise} \end{cases}$

NICHT HIER

The coupling along a path.

- A transposition path for $X, Y \in \mathcal{L}(P)$

  $X = z_0 z_1 \dots z_r = Y$ with $z_k \in \mathcal{L}(P)$

  such that $z_k = (z_1 \dots z_n)$ and $\exists i < j$ with

  $z_{k+1} = (z_1 .. z_j .. z_i .. z_n) = \tau_{ij} z_k$

- weights $\omega(z, \tau_{ij} z) = j - i$

  $\omega(X,Y) = \min\left(\sum_k \omega(z_k z_{k+1}) : \begin{matrix} z_1 z_2 \dots \text{ a transp} \\ \text{path for } X\,Y \end{matrix}\right)$

Obs: $\omega(X,Y) \leq \binom{n}{2}$ upper bound on diameter of $\mathcal{G}(P)$

- Path coupling

  Given $(X_t Y_t)$ with min weight transp path

  $X = z_0 z_1 .. z_r = Y$

  choose • $i \in [n-1]$ with prob p_i • $c_0 \in \{0,1\}$ prob $\frac{1}{2}$

  For $j = 1$ to r if $z_j = \tau_i(z_{j-1})$ in $\mathcal{G}(P)$

  then $c_j = 1 - c_{j-1}$

  else $c_j = c_{j-1}$

Hierher

Now let $X_{t+1} = \tau_i^{c_0} X_t$ $\quad Y_{t+1} = \tau_i^{c_r} Y_t$

and in general $z_k' = \tau_i^{c_k} z_k$

We will use $\boxed{\omega(X_{t+1} Y_{t+1}) = \sum_k \omega(z_k z_{k+1})}$

Proposition if $z_{k+1} = \tau_{ab} z_k$ then

$$E[\omega(z_k' \, z_{k+1}')] \leq \omega(z_k z_{k+1}) + \frac{1}{2}(p_{a-1} - p_a - p_{b-1} + p_b)$$

proof:

$$z_k = (z_1 \dots z_a \dots z_b \dots z_n)$$
$$z_{k+1} = (z_1 \dots z_b \dots z_a \dots z_n)$$

$\omega(z_k, z_{k+1}) = b - a$

- if chosen position $i \neq \{a-1, a, b-1, b\}$
  then τ_i^c acts in the same way on both
  $\Rightarrow \omega' = \omega(z_k', z_{k+1}') = b - a$
- if $i = a-1$
  τ_i^c active on both $\Rightarrow \omega' = \omega + 1$
  τ_i^c active on one, say on z_k,
  $\Rightarrow \omega(z_k', z_{k+1}') \leq \omega(z_k', z_k) + \omega(z_k, \underset{z_{k+1}'}{\underbrace{z_{k+1}}})$
  hence $\omega' \leq \omega + 1$
  τ_i^c inactive on both $\omega' = \omega$ with prob $\geq \frac{1}{2}$
  $\Rightarrow E(\omega') \leq \omega + \frac{1}{2}$
- if $i = b$ same analysis as with $i = a-1$
- if $i = a$ and $b \neq a+1$
  Note that $z_{a+1} \parallel z_a$ and $z_{a+1} \parallel b$
  also $c_k = c_{k+1}$
  $\Rightarrow$ with prob $1/2$ $c_k = 1$ and $\omega' = \omega - 1$
  with prob $1/2$ both stay $\omega' = \omega$
  $\Rightarrow E(\omega') = \omega - \frac{1}{2}$
- if $i = b-1$ same analysis as with $i = a$

Summing up: If $b \neq a+1$ then

$$E(\omega') \leq \omega + \frac{1}{2}\left(p_{a-1} - p_a - p_{b-1} + p_b\right)$$

same

- If $b = a+1$ then
  - if $i \notin \{a-1, a, a+1\}$ then $\omega' = \omega$
  - if $i \in \{a-1, a+1\}$ then $E(\omega') \leq \omega + 1/2$ as before
  - if $i = a$ then $z_a \parallel z_{a+1}$ and $c_k \neq c_{k+1}$

  $\Rightarrow$ The transposition will be applied to exactly one $z'_k = z'_{k+1}$ and $E\omega' = \omega - 1$

$$E(\omega') \leq \omega + \frac{1}{2}\left(p_{a-1} - 2p_a + p_b\right) = \omega + \frac{1}{2}\left(p_{a-1} - p_a - p_{b-1} + p_b\right)$$

$\square$

Remark • if $p_i = \frac{1}{n-1}$ $\forall i$ $\Rightarrow$ $E(\omega') = \omega$

$\Rightarrow$ $E(\omega(X_t Y_t))$ behaves like a random walk on a line $0 \;\cdots\; \binom{n}{2}$

mixing time $\sim n \binom{n}{2}^2 \in O(n^5)$ tries for a real move

- With $p_i = 1/(n-1)$ (this is the KK chain) Wilson uses a sinusoidal weight to show mixing in $O(n^3 \log n)$

We now fix $p_i = \frac{i(n-i)}{K}$ $\qquad K = \sum_{1}^{n-1} i(n-i) = \frac{(n-1)n(n+1)}{6}$

this is optimized concave

$$p_{a-1} - p_a = \frac{-n+2a-1}{K} < 0$$

$$\frac{1}{2}\left(p_{a-1} - p_a - p_{b-1} + p_b\right) = \frac{a-b}{K}$$

dazwischen $\sum \left(1 - \frac{1}{K}\right) \underbrace{\omega(z_k z_{k+1})}_{b-a}$

$$\Rightarrow E(\omega(X_{t+1} Y_{t+1})) \leq \sum E(\omega(z'_k z'_{k+1})) = \left(1 - \frac{1}{K}\right)\omega(X_t Y_t)$$

IX

$$\Pr(X_t \neq Y_t) \leq E(\omega(X_t, Y_t)) \leq \left(1 - \frac{1}{K}\right)^t \cdot \binom{n}{2}$$

$$\leq e^{-t/K} \binom{n}{2}$$

this is $< \varepsilon$ when

$t > K \ln\left(\binom{n}{2} \varepsilon^{-1}\right)$

with $K \approx \frac{1}{6} n^3$ this yields mixing in $O(n^3 \log n)$

Consequences for counting

Let $P_0 = P$ and $P_{K+1} = \text{trans}(P_K + (x_K y_K))$

finally let P_m be a chain

$$\Rightarrow e(P) = \frac{e(P_0)}{e(P_1)} \frac{e(P_1)}{e(P_2)} \cdots \frac{e(P_{m-1})}{e(P_m)} = \prod_0^{m-1} \frac{1}{\Pr\left(\begin{smallmatrix} x_K < y_K \\ \text{in } P_K \end{smallmatrix}\right)}$$

these probabilities can be estimated by looking at random linear extensions.

Yields FPRAS for #LinExt

complexity $O(n^5 \varepsilon^{-2} \operatorname{polylog}(n, \varepsilon))$