Order 8: 7–15, 1991.

On the Conductance of Order Markov Chains

ALEXANDER KARZANOV
Institute for Systems Studies of the U.S.S.R. Academy of Sciences, U.S.S.R.

and

LEONID KHACHIYAN*
Department of Computer Science, Rutgers University, New Brunswick, NJ 08903, U.S.A.

Communicated by I. Rival

(Received: 30 October 1990; accepted: 18 January 1991)

Abstract. Let Q be a convex solid in $\mathbb{R}^n$, partitioned into two volumes u and v by an area s. We show that $s > \min(u, v)/\text{diam } Q$, and use this inequality to obtain the lower bound $n^{-5/2}$ on the conductance of order Markov chains, which describe nearly uniform generators of linear extensions for posets of size n. We also discuss an application of the above results to the problem of sorting of posets.

AMS subject classification (1991). 06A06.

Key words. Conductance, isoperimetric inequality, linear extension, poset, uniform generation, sorting.

1. Isoperimetric Inequality on Partitions of Convex Bodies and a Lower Bound on the Conductance of Order Markov Chains

Let $A = \{1, \ldots, n; \prec\}$ be a poset with n elements, and let $E = E(A)$ be the set of linear extensions of A, i.e. the set of all total linear orders

$$e = \{e(1) \prec e(2) \prec \cdots \prec e(k) \prec e(k+1) \prec \cdots e(n)\} \tag{1.1}$$

compatible with the partial order $\prec$ on A. Thus, each linear extension $e \in E$ can be viewed as a permutation (1.1) of elements $1, \ldots, n$. Two linear extensions $e, g \in E$ are said to be *neighbors in E*, if g can be obtained by a single transposition of two consecutive elements in e, i.e. if

$$g = \{e(1) \prec e(2) \prec \cdots \prec e(k+1) \prec e(k) \prec \cdots \prec e(n)\}$$

for some $k \in [1, n-1]$. In particular, the number $n(e)$ of neighbors of e in E is at most $n - 1$.

Given a poset A, consider the *order Markov chain M of A* with states $e \in E$, and with transition probabilities

$$p(e, g) = \begin{cases} 1/(2n-2) & \text{if } e \text{ and } g \text{ are neighbors in } E \\ 1 - n(e)/(2n-2) & \text{if } e = g \\ 0 & \text{otherwise.} \end{cases} \tag{1.2}$$

* On leave from Computing Center of the USSR Academy of Sciences.

Geometrically, we consider the canonical triangulation

$$Q = \bigcup Q(e), \quad e \in E \tag{1.3}$$

of the order polytope of A

$$Q = \{x \in \mathbb{R}^n \mid 0 \leqslant x_i \leqslant 1, i = 1, \ldots, n; x_i \leqslant x_j \text{ if } i \prec j \text{ in } A\} \tag{1.4}$$

into the simplices

$$Q(e) = \{x \in \mathbb{R}^n \mid 0 \leqslant x_{e(1)} \leqslant x_{e(2)} \leqslant \cdots \leqslant x_{e(n)} \leqslant 1\}, \tag{1.5}$$

corresponding to linear extensions e of A, see, e.g., [11]. The order Markov chain (1.2) describes a random walk through the simplices in the triangulation (1.3), or equivalently through the set E of linear extensions of A. This random walk starts at an arbitrary simplex $Q(e_0)$ in the triangulation. At the t-th step, $t = 0, 1, \ldots,$ we choose with probability $1/(2n-2)$ one of $n-1$ facets

$$F_k(e) = Q(e) \cap \{x \in \mathbb{R}^n \mid x_{e(k)} = x_{e(k+1)}\}, \quad k \in [1, n-1] \tag{1.6}$$

of the current simplex $Q(e) = Q(e_t)$. If the adjacent simplex $Q(g)$, sharing the chosen facet $F_k(e)$ with $Q(e)$, belongs to the triangulation, we move to this simplex $Q(e_{t+1}) = Q(g)$, otherwise the random walk stays at the present simplex $Q(e_{t+1}) = Q(e)$.

The above construction is similar to the one considered in [4], and it is easy to see that (1.2) is an ergodic time-reversible Markov chain with uniform stationary distribution. In other words, for an arbitrary poset A and an arbitrary initial probability distribution $\pi(0, e)$, $e \in E$ on the set of linear extensions of A, the distribution after t steps of the random walk

$$\pi(t, e) = \sum_{g \in E} \pi(t-1, g) p(g, e), \quad t = 1, 2, \ldots,$$

converges to the uniform distribution on E:

$$\lim_{t \to \infty} \pi(t, e) = 1/|E|, \quad \forall e \in E.$$

Thus, for an arbitrary A and sufficiently large $T = T_A$ the following algorithm, *RandWalk*, gives a nearly uniform generator of linear extensions of the poset.

```
Input A, T;
{topological sorting} find a linear extension e = {e(1) ≺ ··· ≺ e(n)} of A;
{random walk} for t = 1, ..., T do:
  begin
    choose at random an integer k ∈ [1, ..., 2n − 2];
    if k ≤ n − 1 and not e(k) ≺ e(k + 1) in A, then
      swap e(k) and e(k + 1) in e
  end;
Output e
```

The complexity of *RandWalk* is $O(n^2 + T)$ operations plus T times the complexity of (pseudo) random uniform generation of $k \in [1, 2n-2]$.

The rate of convergence of *RandWalk* can be estimated using the following inequality of Sinclair and Jerrum [9, 10]: for any initial distribution,

$$|\pi(t, e) - 1/|E|| \leqslant (1-\alpha^2)^t, \quad \forall e \in E. \tag{1.7}$$

Here α is the *conductance* of the Markov chain (1.2), defined as

$$\alpha = \frac{1}{2n-2} \min\{C(X)/|X| \mid X \subseteq E,\ 1 \leqslant |X| \leqslant |E|/2\}, \tag{1.8}$$

where

$$C(X) \textit{ is the capacity of the cut } (X, E - X), \tag{1.9}$$

i.e. the number of pairs $e \in X$, $g \in E - X$ such that e and g are neighbors in E.

THEOREM 1. *Let M be the order Markov chain* (1.2) *of a poset with n elements. Then*

$$\alpha = \alpha(M) > 2^{-3/2} n^{-5/2}. \tag{1.10}$$

Remark 1. If A contains a chain of $n-1$ elements and a singleton, then $\alpha = ((2n-2)\lfloor n/2 \rfloor)^{-1} \approx n^{-2}$. We believe that the lower bound (1.10) can be improved by a factor of $n^{0.5}$.

Theorem 1 is a simple corollary of a more general isoperimetric inequality independently obtained in [7] and [8].

THEOREM 2. *Let Q be a convex n-dimensional solid. Suppose that Q is partitioned into two subsets U and V by an $n-1$-dimensional surface $S = \partial U - \partial Q = \partial V - \partial Q$ of area $s = vol_{n-1} S$. Then*

$$s > \min(u, v)/\text{diam}\, Q, \tag{1.11}$$

where $u = vol_n U$ and $v = vol_n V$.

Proof. For any $\varepsilon > 0$ there exists a convex solid $Q_\varepsilon \subset Q$ with a smooth boundary, such that $vol_n(Q - Q_\varepsilon) < \varepsilon$. If (1.11) holds for Q_ε and $\varepsilon \to 0$, we get (1.11) for Q. Therefore we may assume without loss of generality that the boundary of the original body Q is smooth.

Let us fix two positive numbers u and v, whose sum is the volume of Q, and consider Plateau's problem:

$$\begin{aligned} &\textit{minimize } vol_{n-1} S, \\ &\textit{where } S \textit{ is an } (n-1)\textit{-dimensional surface,} \\ &\textit{partitioning } Q \textit{ into volumes } u \textit{ and } v. \end{aligned} \tag{1.12}$$

The proof utilizes some known properties of extremal surfaces ((1.13), (1.14) and (1.16)), briefly reviewed below.

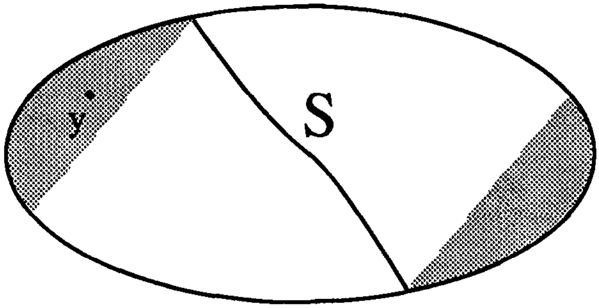

Fig. 1.

Suppose at first that there exists a smooth surface S, extremal to the problem (1.12). Then

S is transversal to the boundary of Q. (1.13)

The convexity of Q and (1.13) imply that

for any point $y \in int\, Q - S$ there exists a regular point $x \in S \cap int\, Q$, closest to y on the surface: $\|y - x\| = dist(y, S)$. (1.14)

Remark 2. If the transversality condition (1.13) is violated, then for some points $y \in int\, Q - S$ all the closest points on the surface may lie on ∂Q (see Figure 1).

In general, the extremal partitioning surface S possibly may have a closed singular set. However, it is known [1] that any point on S, closest to a point not on S, is regular (analytic). Therefore (1.14) holds for the general case.

Let $\delta(x)$ be a smooth vector field, vanishing outside of a small regular disc on $S \cap int\, Q$, and let S' be the surface, obtained by means of displacement of each point x in the disc by the vector $\lambda\delta(x)$, where λ is a scalar. Then for small λ the variations of the volumes u and v are given by the surface integrals

$$\delta u = -\delta v = \lambda \int_S \langle \delta(x), n(x) \rangle \, ds_x + O(\lambda^2),$$

where $n(x)$ is the outside normal vector to U at x, and $\langle\ ,\ \rangle$ stands for the inner product. Next, the variation of the area s of the boundary is given by (see, e.g., [3])

$$\delta s = \lambda \int_S \langle \delta(x), n(x) \rangle h(x) \, ds_x + O(\lambda^2),$$

where

$$h(x) = \sum_{i=1}^{n-1} \frac{1}{R_i(x)}, \tag{1.15}$$

and $R_i(x)$, $i = 1, \ldots, n-1$ are the radii of curvature of the U-side of S at x (we put

$1/\infty = 0$). Since S is extremal to the problem (1.12), we get the following well-known condition (see, e.g., [3]):

$$\text{at any regular point } x \in S \cap \mathit{int}\, Q \text{ of the extremal surface, the mean curvature (1.15) of the } U\text{-side of } S \text{ is constant: } h(x) = H. \tag{1.16}$$

Of course, (1.16) also holds for the V-side of S with the mean curvature $-H$.

Let $y \in \mathbb{R}^n - S$, and let $x \in S \cap \mathit{int}\, Q$ be a regular point closest to y on S. Clearly, this implies that

$$\begin{aligned}&y - x \text{ is perpendicular to } S \text{ at } x, \text{ and}\\ &\rho = \|y - x\| \leqslant \rho(x) = \min\{R_i(x) \mid i = 1, \dots, n-1 : R_i(x) > 0,\ R_i(x) \neq \infty\},\\ &\text{where } R_i(x) \text{ are the radii of curvature of the } y\text{-side of } S \text{ at } x.\end{aligned} \tag{1.17}$$

[If all $n-1$ radii of curvature are either negative or equal to ∞, we put $\rho(x) = +\infty$.]

If a point $y \in \mathbb{R}^n - S$ and a regular point $x \in S \cap \mathit{int}\, Q$ satisfy (1.17), we say that y is visible from the corresponding side of S at distance ρ. As we know from (1.14), all the interior points of U (of V) are visible from the U-side (from the V-side) of S at distance $\rho \leqslant \operatorname{diam} Q$. Note that some of the visible points may lie outside of Q.

Let ds_x be an elementary area of the extremal surface at a regular point $x \in S \cap \mathit{int}\, Q$, and let $ds_x[\rho]$ be the corresponding elementary area, visible from the U-side of ds_x at distance ρ. By (1.17)

$$ds_x[\rho] = \begin{cases} ds_x \prod_{i=1}^{n-1} (1 - \rho/R_i(x)) & \text{if } \rho \leqslant \rho(x)\\ 0 & \text{if } \rho > \rho(x),\end{cases}$$

where $R_i(x)$ are the radii of curvature of the U-side of S at x. Since $1 - r \leqslant \exp(-r)$, we get for all $\rho \geqslant 0$

$$ds_x[\rho] \leqslant ds_x \exp\left(-\rho \sum_{i=1}^{n-1} \frac{1}{R_i(x)}\right) = ds_x \exp(-\rho H), \tag{1.18}$$

(see (1.16)). Recalling that all the interior points of U are visible from the U-side of S at distance $\rho \leqslant \operatorname{diam} Q$, we obtain from (1.18) the following upper bound:

$$u = \mathit{vol}_n U < \int_S \left[\int_0^{\operatorname{diam} Q} \exp(-\rho H)\, d\rho\right] ds_x = s \int_0^{\operatorname{diam} Q} \exp(-\rho H)\, d\rho.$$

If $H \geqslant 0$, w have $u < s \operatorname{diam} Q$ and (1.11) follows. If $H < 0$, we have $v < s \operatorname{diam} Q$ and again obtain (1.11). □

Remark 3. Let Q be a cylinder of height h, and let S be the middle hyperplane parallel to the base of Q. Then $s/\min\{u, v\} = 2/h \to 2/\operatorname{diam} Q$ for $h \to \infty$.

Now we can prove Theorem 1.

Proof. Let A be an arbitrary poset with $|E| > 1$ linear extensions, and let X be a subset of E of cardinality $1 \leqslant |X| \leqslant |E|/2$. Consider the corresponding partition

$$U = \bigcup_{e \in X} Q(e), \quad V = \bigcup_{e \in E - X} Q(e)$$

of the order polytope Q of A, see (1.3)–(1.5). Clearly,

$$u = vol_n U = \frac{|X|}{n!} \leqslant v = vol_n V = \frac{|E| - |X|}{n!}.$$

Furthermore, since

$$vol_{n-1} F_k(e) = 2^{1/2}/(n-1)!$$

for each facet (1.6), we have

$$s = vol_{n-1} S = C(X) 2^{1/2}/(n-1)!,$$

where S is the common boundary of U and V, and $C(X)$ is the capacity of the cut (see (1.9)). Hence

$$\frac{1}{2n-2} \frac{C(X)}{|X|} = \frac{1}{(2n-2)2^{1/2}n} \frac{s}{\min\{u, v\}} > 2^{-3/2} n^{-2} \frac{s}{\min\{u, v\}}$$
$$by\ (1.11) > 2^{-3/2} n^{-2}/\text{diam}\, Q = 2^{-3/2} n^{-5/2},$$

and we obtain (1.10), from the definition (1.8) of α. □

By (1.7) and (1.10) the number T of steps of the algorithm *RandWalk*, sufficient to generate linear extensions of a given poset A uniformly to a given relative accuracy $\nu \in (0, 1)$

$$|\pi(T, e) - 1/|E|| \leqslant \nu/|E|, \quad \forall e \in E \tag{1.19}$$

can be bounded as

$$T = O\left(n^5 \log \frac{|E|}{\nu}\right).$$

Thus, for reasonable ν, say $\nu = 0.01$, one has

$$T = O(n^5 \log|E|).$$

We expect that these bounds can be improved by *at least* a factor of n.

2. An Application to the Problem of Sorting of Posets

Suppose that we wish to sort a given poset A with $|E(A)| > 1$ linear extensions by querying an oracle. Namely suppose that we can choose any pair $i, j \in \{1, \ldots, n\}$ of incomparable elements in A and ask the oracle to compare them. Having gotten the answer, say i precedes j, we add the relation $i \prec j$ and all its transitive consequences to A and obtain a new partial order $A^1 = A\&[i \prec j]$ on the same set $\{1, \dots, n\}$ of elements. If $|E(A^1)| > 1$, we call the oracle again, ask it to compare a new pair of element, obtain a new partial order A^2 and so on.

Thus, in a finite number q of queries we sort the original poset A, i.e. obtain a total linear order $A^q = e \in E(A)$ on $\{1, \dots, n\}$. Clearly, one has the following well-known information theory worst-case lower bound

$$q \geqslant \log_2 |E(A)| \tag{2.1}$$

on the number of queries. In 1976 Fredman [5] made the following conjecture:

For any poset A with $|E(A)| > 1$ linear extensions there exists a pair of elements $i,j \in \{1, \dots, n\}$ such that

$$\max\left\{\frac{|E(A\&[i \prec j])|}{|E(A)|}, \frac{|E(A\&[j \prec i])|}{|E(A)|}\right\} \leqslant \beta \tag{2.2}$$

with $\beta = 2/3$.
The inequality (2.2) says that in any poset A there exists a β-balanced comparison i, j which decreases the number of linear extensions by at least β.

At the present the original conjecture (2.2) with $\beta = 2/3$ remains open. However, it is known [6] that (2.2) holds with $\beta = 8/11$. The latter result implies that using 8/11-balanced test comparison one can sort an arbitrary poset A in at most

$$q \leqslant 2.2 \log_2 |E(A)| \tag{2.3}$$

queries.

It is also known [2] that computing the "balancing constants"

$$\beta_{ij} = |E(A\&[i \prec j])|/|E(A)| = probability\{i \prec j \text{ in } E(A)\}$$

is $\#P$-hard.

However, one can compute sufficiently good approximations to the balancing constants, say

$$|\hat{\beta}_{i,j} - \beta_{ij}| \leqslant 0.1 \text{ with probability } 0.99$$

in time $O(T)$, where T is the complexity of nearly uniform generation of linear extensions of A. Therefore, *a well-balanced comparison in a given poset can also be found with the same probability in time $O(T)$.*

Since, T necessary to obtain (1.19) with a fixed v, can grow at least as n^3 (see example in Remark 1), in our computational experiments (codes were written by E. Zhirova, Moscow Institute of Physics and Technology, 1989) we used another approach, based on the following fact [6]:

Let

$$\tau_i = \frac{1}{|E(A)|} \sum_{e \in E} e^{-1}(i) \tag{2.4}$$

be the average rank of $i \in \{1, \ldots, n\}$ *over the set of linear extensions of* A, *then an arbitrary pair* i, j *of elements such that*

$$|r_i - r_j| < 1 \tag{2.5}$$

is an 8/11-*balanced comparison in* A.

Remark 4. Since $|E(A)| > 1$, it is easy to see that a pair (2.5) always exists and that (2.4)–(2.5) imply the incomparability of i and j in A. In fact, it is easy to show that if instead of (2.5) a pair i, j of incomparable elements satisfies a weaker inequality, say $|r_i - r_j| \leqslant 2$, then the comparison of i and j is still β-balanced with some absolute constant $\beta < 1$.

We generated a single random trajectory of RandWalk[1] in $E(A)$ for averaging the ranks $\hat{r}_i$ of elements at the moments $t = n, 2n, \ldots, \lfloor T/n \rfloor$ for $T = 10n^2$. Due to the choice of the moments of sampling on the trajectory, these computations require only $O(T) = O(n^2)$ operations. Next we determined a possible well-balanced comparison using the strategy of minimizing $|\hat{r}_i - \hat{r}_j|$ over $i, j \in \{1, \ldots, n\}$. Though this approach does not guarantee a reliable determination of well-balanced comparisons, the computational results on the *total* number q of queries were encouraging at least for the moderate values of n. For instance, if A consists of n singletons, one gets the best known bounds on q up to $n = 16$ and the results, which differ by at most 5% from the lower bound (2.1) up to $n = 50$. This is also true in the experiments with merging two chains of lengths $n_1 + n_2 \leqslant 50$, and in some other experiments. Thus, at least for posets of moderate size, even very short random trajectories result in the total number of queries, which is much less than the theoretical upper bound (2.3) for long trajectories, and close to the information theory lower bound (2.1).

Acknowledgments

The authors would like to thank Professor F. Almgren for some explanations regarding the regularity condition used in Theorem 2.

Note

[1] Without idle loops corresponding to the values of k in $[n, 2n - 2]$.

References

1. F. Almgren (1976) Existence and regularity almost everywhere of solutions to elliptic variational problems with constraints, *Mem. Amer Math. Soc. No.* 165.
2. G. Brightwell and P. Winkler (1990) Counting linear extensions is #P-complete. Preprint, Bellcore, July.
3. Yu. D. Burago and V. A. Zalgaller (1988) *Geometric Inequalities*, Berlin; New York, Springer-Verlag.

4. M. Dyer, A. Frieze, and R. Kannan (1988) A random polynomial time algorithm for estimating volumes of convex bodies. Preprint, Department of Computer Science, Carnegie-Mellon University, September.
5. M. Fredman (1976) How good is the information theory bound in sorting, *Theoretical Computer Science* **1**, 355–361.
6. J. Kahn and M. Saks (1984) Every poset has a good comparison, *Proc. 16-th Symposium on Theory of Computing*, pp. 299–301.
7. A. Karzanov and L. Khachiyan (1990) On the conductance of order Markov chains. Technical Report DCS TR 268, Department of Computer Science, Rutgers University, June.
8. L. Lovasz and M. Simonovits (1990) The mixing rate of Markov chains, an isoperimetric inequality, and computing the volume. Preprint, Mathematical Institute of the Hungarian Academy of Sciences, May.
9. M. Mihail (1989) Conductance and convergence of Markov chain – A combinatorial treatment of expanders, *Proc. 30-th Symp. on Foundations of Computer Science*, pp. 526–531.
10. A. J. Sinclair and M. R. Jerrum (1989) Approximate counting generation and rapidly mixing Markov chains, *Information and Computation* **82**, 93–133.
11. R. P. Stanley (1986) Two order polytopes, *Discrete and Computational Geometry* **1**, 9–23.