DIMACS Series in Discrete Mathematics
and Theoretical Computer Science

On the Mixing Time of the Triangulation Walk and other Catalan Structures

Lisa McShine and Prasad Tetali

1. Introduction

Consider a graph on the set of triangulations of a convex polygon with n sides, wherein two triangulations are considered adjacent if (and only if) one can be obtained from the other by "flipping" an internal diagonal – note that each internal diagonal in a triangulation is in a unique quadrilateral, and by "flipping" we mean replacing one diagonal by the other diagonal of the same quadrilateral. A random walk can be easily defined on this graph with the property that, eventually, the probability of being at any given triangulation is independent of the choice of triangulation. This gives a Monte Carlo method of generating a triangulation of a convex n-gon *uniformly at random*. However, the efficiency of such a scheme depends crucially on the rate of approach to stationarity of the random walk. Although there are other ways of obtaining randomly such a triangulation in an efficient way, the analysis of this particular scheme has remained open.

The main objective in this paper is to show that $O(n^5 \log(n/\epsilon))$ steps are sufficient to get close (i.e. within ϵ in variation distance) to the stationary distribution, which happens to be uniform over the set of triangulations. Independently, Molloy et al. [**12**] have recently shown that at least $\Omega(n^{3/2})$ steps are necessary and also that $O(n^{23} \log(n/\epsilon))$ steps are sufficient. While our upper bound is much more reasonable and our proof much simpler, we believe the truth to be closer to their lower bound.

The idea in [**12**] in getting an upper bound on the mixing time was to bound the so-called *conductance* of the Markov chain, a technique first introduced by Jerrum and Sinclair. Here we use the comparison technique due to Diaconis and Saloff-Coste [**3**]. Recently this technique has also been used to analyze the mixing time of certain Markov chains on tilings in [**13**]. Before describing our idea further, we need to define some other *Catalan*

1991 *Mathematics Subject Classification.* 68R05, 68Q25, 60J10, 05A05.

structures – a term we use to denote any of the possible combinatorial structures whose counting sequence is the sequence of Catalan numbers, given by $c_n = \binom{2n}{n}/(n+1)$, for $n \geq 1$, and $c_0 = 1$. Some other examples of Catalan structures include the set of binary trees on n internal nodes and the set of *Dyck paths* (also called *mountain-valley* diagrams) of length $2n$. (See Fig. 1 for illustrations and Section 2 for precise definitions.) Each such Catalan structure offers an *interchange graph* in a natural way, with an appropriate definition of a local move or a local interchange similar to the move described above for the graph on triangulations of convex n-gon. (See Fig. 2 for illustration.)

Our strategy can now be summarized as follows. We use the fact that the interchange graph on triangulations of a convex n-gon is isomorphic to the interchange graph on binary trees with $n-2$ internal nodes; this fact was established in [**15**]. We then compare, *à la* Diaconis–Saloff-Coste, the chain on binary trees with the chain on Dyck paths. That is, we use the upper bound on the mixing time of the chain on Dyck paths as established by [**17**] (see also [**11**]) to get an upper bound on the mixing time of the chain on binary trees (equivalently, the chain on triangulations).

Dobrow and Fill ([**5**], [**6**]) analyzed a Markov chain on the state space of all binary trees on n nodes, wherein the transitions between states were defined in terms of a certain *move-to-front* rule. However, the Markov chain described in [**5**] does not yield the uniform distribution as the stationary distribution. Although in principle one can generate a binary tree uniformly at random using a more direct approach (as done in [**9**], [**1**],...), it is conceivable that there exists a faster method via the generation of an equivalent Catalan structure. Note that each Catalan structure offers its own Markov chain – the interchange graphs corresponding to different Catalan structures need not be (and typically are not) isomorphic; hence the corresponding random walks on the interchange graphs can behave in significantly different ways. Finally, uniform random generation of triangulations of *non-convex* polygons is apparently of sufficient interest to the community of computer graphics and computational geometry, because so far for these problems, there are no known rigorous ways of efficient random generation. It remains to be seen whether Monte Carlo Markov chain techniques will help in the non-convex case, and in the more general case of triangulations using n points (in general position) in the plane. A particularly appealing feature of the Markov chain approach to the problem of random generation is that it uses much less space, and typically much smaller (random/pseudo-random) numbers in the simulation.

In a different spirit, the nature of the present study is akin to the research which arose out of the analysis of Markov chains based on various card-shuffling schemes. Several researchers (most notably P. Diaconis) have succeeded in obtaining sharp estimates on the rates of mixing of chains based on such schemes. In fact, the comparison technique (see [**2**], [**3**]) was a result of such investigations. There does not seem to be analogous work on

Markov chains on Catalan structures, and we view the present contribution as the beginning of a systematic study.

The following presentation is organized as follows. In Section 2 we describe the background material on Catalan structures and Markov chains relevant to our work. In Section 3 we present the proof of the bound on the mixing time of the Markov chain(s) in question. In the final section, we conclude with some remarks on further work in progress.

2. Preliminaries

In this section we first give a brief description of some Catalan structures, and associated interchange graphs on them. We then describe relevant results from the literature on rapidly mixing Markov chains which we need in the next section to analyze the chain on binary trees on n internal nodes.

2.1. Catalan structures. The nth Catalan number, c_n, satisfies the following recurrence relation (see [**8**]).

$$c_n = c_0 c_{n-1} + c_1 c_{n-2} + \cdots + c_{n-1} c_0,$$

where $c_0 = 1$, and $c_1 = 1$. Also recall that for $n \geq 1$, we have $c_n = \binom{2n}{n}/(n+1)$. There are several interesting survey papers (see e.g. [**7**], [**8**]) on Catalan sequences which describe various combinatorial interpretations of these sequences. Perhaps the best source is a list of 60 or so interpretations in [**16**]. We shall henceforth use the term *Catalan structures* to mean any of the possible combinatorial structures whose counting sequence satisfies the recurrence relation describing the Catalan sequence. Some examples include the following (see Fig. 1 for $n = 3$).

Consider a convex polygon K, with $n+2$ vertices, labeled $1, 2, \ldots, n+2$, clockwise. A triangulation of K is a dissection of K into n triangles, using nonintersecting diagonals of K. The number of such triangulations is c_n.

For the purpose of this article, a binary tree of size n is a rooted tree with n *internal nodes* (those with two descendants) and $n+1$ *external nodes* or *leaves* (those with no descendants). It is well known that there are c_n such binary trees with n internal nodes.

A Dyck path from $(0, 0)$ to $(2n, 0)$ is a lattice path with steps $(1, 1)$ and $(1, -1)$ never falling below the x-axis.

Finally, label $2n$ equally spaced points around the circumference of a circle; join the points in pairs by n nonintersecting chords. The number of such Dyck paths and such chord diagrams is also c_n. We recommend that the reader refers to a lovely exposition of some of the Catalan structures by Martin Gardner [**7**], who also describes interesting bijections between these structures.

Each Catalan structure offers its own *interchange graph* in a natural way. The Markov chains which we will study are all random walks on these interchange graphs. The principle behind the definition of each interchange graph is the same:

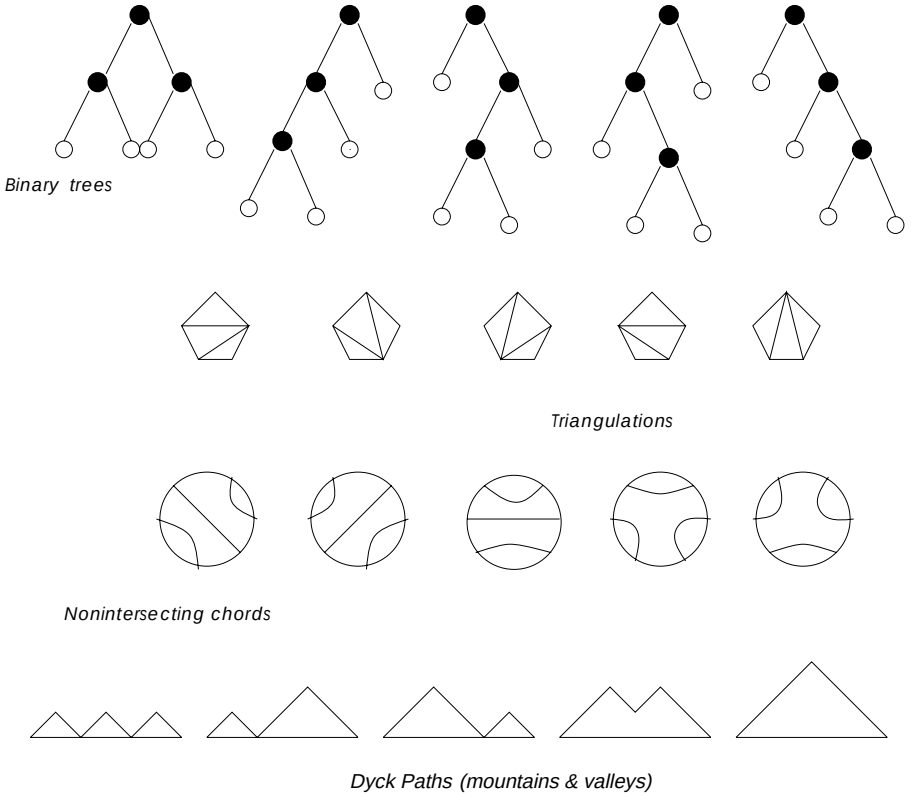

FIGURE 1. Examples of Catalan structures for $c_3 = 5$

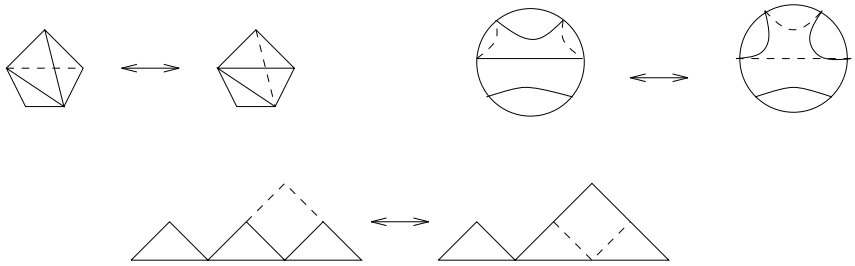

FIGURE 2. Local moves defining interchange graphs

> The vertices of the graph are labeled with the elements of a particular Catalan structure (of size n), and two vertices are adjacent in the graph if a natural (local) operation transforms the element corresponding to one of the vertices into that of the other.

Examples of such local operations/moves are illustrated in Figs. 2 and 4. Suppose the structure is the set of triangulations of a convex $(n+2)$-gon. The set of triangulations forms the vertex set of the interchange graph,

and two triangulations are adjacent if one can be obtained from the other by a diagonal flip, as described in [**15**]. Every diagonal in a triangulation of a convex polygon defines a quadrilateral. A diagonal flip replaces that diagonal with the other diagonal of the same quadrilateral. Sleator et al. [**15**] obtained, *inter alia*, tight upper and lower bounds (of $2n-6$) on the diameter of this interchange graph and other results on triangulations of the sphere (see [**10**] for a simpler proof).

Two binary trees with n internal nodes are adjacent if one can be transformed into the other by applying the rotation operation. A rotation at a node is defined as shown in Fig. 4. Sleator et al. also showed that this graph is isomorphic to the previous one on triangulations of a convex $(n+2)$-gon.

Similarly, in the collection of Dyck paths of length $2n$, two elements are adjacent if one may be changed into the other by flipping a peak into a valley (that is, changing $(1,1),(1,-1)$ to $(1,-1),(1,1)$) or a valley into a peak (that is, changing $(1,-1),(1,1)$ to $(1,1),(1,-1)$). It is easy to see that the diameter of this graph is precisely $n(n-1)/2$.

For the set of nonintersecting chords in a circle, a valid way to define interchanges is to pick a pair of chords and to replace them with two new chords obtained by matching the endpoints of the original chords, if and only if such an exchange results in a valid nonintersecting chord diagram. It is easy to show that this yields a connected undirected graph, and in fact (although less obvious) a straightforward proof by induction on n can be given to show that the diameter of such an interchange graph is $n-1$, where n is the number of chords. (Also see the remarks in Conclusions.)

2.2. Comparison of mixing times of Markov chains. Let (Ω, P, π) denote an ergodic (that is, irreducible and aperiodic) Markov chain, with finite state space Ω, transition probability matrix P, and stationary distribution π. We assume that P is reversible, that is, for all $x, y \in \Omega$,

$$\pi(x)P(x,y) = \pi(y)P(y,x)$$

We will also assume that we are considering discrete-time Markov chains. Then, for $x, y \in \Omega$, $t \in \mathbb{Z}^+$, $P^t(x,y)$ denotes the t-step probability of going from x to y. The time the Markov chain takes to be close to the stationary distribution, starting from state x, can be measured by the *variation distance*,

$$\Delta_x(t) = \frac{1}{2}\sum_{y\in\Omega} |P^t(x,y) - \pi(y)|.$$

The variation distance from the worst state is denoted by

$$\Delta(t) = \max_{x\in\Omega} \Delta_x(t).$$

For $\epsilon > 0$ (usually $0 < \epsilon < 1$), the *mixing time* from state x is defined by

$$\tau_x(\epsilon) = \min\{t : \Delta_x(t') \leq \epsilon, \forall t' \geq t\}.$$

The mixing time of the Markov chain is the mixing time from the worst state,

$$\tau(\epsilon) = \max_{x \in \Omega} \tau_x(\epsilon).$$

In the following, the mixing time of a Markov chain will always refer to the mixing time from a worst state.

Let $1 = \lambda_0 > \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{|\Omega|-1} > -1$ denote the eigenvalues of P. The following result (see [**14**]) shows the relationship between mixing time and maximum eigenvalue. Strictly speaking λ_1 should be replaced by $\lambda_{\max} = \max(\lambda_1, \lambda_{|\Omega|-1})$, but in the Markov chains we will be describing we ensure that $\lambda_1 > \lambda_{|\Omega|-1} > 0$ by adding self-loops of probability at least $1/2$.

THEOREM 1 (Sinclair). *For $\epsilon > 0$,*

(i)

$$\forall x \in \Omega, \tau_x(\epsilon) \leq \frac{1}{1-\lambda_1} \log\left(\frac{1}{\pi(x)\epsilon}\right),$$

(ii)

$$\tau(\epsilon) = \max_{x \in \Omega} \tau_x(\epsilon) \geq \frac{\lambda_1}{2(1-\lambda_1)}.$$

Let P, $\tilde{P}$ denote two reversible Markov chains on the same state space Ω, with the same stationary distribution π. Diaconis and Saloff-Coste [**3**] provide the following geometric bound between $\lambda_1(P)$, $\lambda_1(\tilde{P})$. In fact the result compares Dirichlet forms associated with P and $\tilde{P}$, so the result applies to all non-trivial eigenvalues, not just λ_1, and also to the log-Sobolev constants of P and $\tilde{P}$. Also, as stated in [**3**], the two chains need not share the same stationary distribution, it suffices if they have comparable distributions.

Let $\tilde{P}$ denote the Markov chain with known eigenvalues (or known mixing time), and let P denote the chain whose mixing time we would like to bound, by comparison with $\tilde{P}$.

Let $E(P) = \{(x, y) : P(x, y) > 0\}$ and $E(\tilde{P}) = \{(x, y) : \tilde{P}(x, y) > 0\}$ denote the sets of edges of the two chains, viewed as directed graphs.

For each (x, y) with $\tilde{P}(x, y) > 0$, define a path γ_{xy} using a fixed sequence of states, $x = x_0, x_1, \ldots, x_{k-1}, x_k = y$, with $P(x_i, x_{i+1}) > 0$. The length of such a path is denoted by $|\gamma_{xy}|$ and $|\gamma_{xy}| = k$.

Further let , $(z, w) = \{(x, y) \in E(\tilde{P}) : (z, w) \in \gamma_{xy}\}$ denote the set of paths (in P) which use the edge (z, w).

PROPOSITION 2 (Diaconis–Saloff-Coste). *With the above notation we have*

$$1 - \lambda_1(P) \geq \frac{1}{A(,)}(1 - \lambda_1(\tilde{P})),$$

where

$$A(,\ ) = \max_{(z,w)\in E(P)} \left\{ \frac{1}{\pi(z)P(z,w)} \sum_{\Gamma(z,w)} |\gamma_{xy}|\pi(x)\tilde{P}(x,y) \right\}.$$

Crucially, note that $A(,\ )$ depends on our choice of paths , $= \{\gamma_{xy}\}$, and that we need define these paths only between pairs of states which are *adjacent in the known chain.*

The strategy used in comparing mixing times based on the above proposition is described in [**13**]. First begin with a bound on the known mixing time of a chain. Part (ii) of Theorem 1 allows us to lower bound the spectral gap of such a chain. Then Proposition 2 is used to lower bound the spectral gap of an unknown chain by bounding the parameter $A(,\ )$, for a carefully chosen , . This then gives a bound on the mixing time of the unknown chain via part (i) of Theorem 1. The technique is summarized in the following proposition.

PROPOSITION 3 (Randall–Tetali). *Let $\tilde{\tau}(\epsilon)$, $\tau(\epsilon)$ denote the mixing times of $\tilde{P}$ and P respectively, and let $\pi_* = \min_{x\in\Omega} \pi(x)$. Then with $A(,\ )$ defined as in Proposition 2: for $0 < \epsilon < 1$,*

$$\tau(\epsilon) \leq \frac{4\log\left(\frac{1}{\epsilon\pi_*}\right)}{\log\left(\frac{1}{2\epsilon}\right)} A(,\ )\tilde{\tau}(\epsilon).$$

3. An upper bound on the mixing time of binary trees

Sleator et al. [**15**] gave a bijection between the set of triangulations of a convex $(n+2)$-gon and binary trees with n internal nodes so that if two triangulations differ by a diagonal flip, the corresponding binary trees differ by a rotation operation. This correspondence shows that the two interchange graphs are isomorphic. In turn it allows us to study the random walk on the interchange graph on binary trees on n internal nodes, which we will denote (following [**15**]) $RG(n)$ and from this deduce the result for triangulations.

The transition probabilities of the Markov chain on $RG(n)$ are defined as follows. For two distinct binary trees z and w,

$$\begin{aligned} P(z,w) &= 1/[2(n-1)], \quad \text{if } (z,w) \in E(RG(n)) \\ P(z,z) &= 1/2. \end{aligned}$$

The transition probabilities of the Markov chain on $DG(n)$, the interchange graph for Dyck paths of length $2n$, are defined as follows. For two distinct Dyck paths x and y,

$$\begin{aligned} \tilde{P}(x,y) &= 1/[2(2n-3)] \quad \text{if } (x,y) \in E(DG(n)) \\ \tilde{P}(x,x) &= 1 - \textstyle\sum_{y\sim x} P(x,y) \geq 1/2 \end{aligned}$$

It is easy to verify that the two Markov chains as defined above do indeed satisfy the reversibility (i.e. detailed-balance) condition and also that they

share the uniform distribution as the stationary distribution π. So for $x \in \tilde{\Omega}$ and $z \in \Omega$, we have $\tilde{\pi}(x) = \pi(z) = 1/c_n$.

Let $\tau(\epsilon)$ and $\tilde{\tau}(\epsilon)$ denote the mixing times of the Markov chains on $RG(n)$ and $DG(n)$ respectively. First we may deduce from [**17**] that the Markov chain on $DG(n)$ has mixing time

$$\tilde{\tau}(\epsilon) = O(n^3(\log n + \log 1/\epsilon)) \ , \ 0 < \epsilon < 1$$

Let our known chain, $\tilde{P}$, is the chain with the set of Dyck paths of length $2n$ as the state space, denoted $\tilde{\Omega}$; our unknown chain, P, is the chain with the set of binary trees with n internal nodes as the state space, denoted Ω. We would like to use Proposition 3 to get a bound on the mixing time $\tau(\epsilon)$. First note that Proposition 3 (and 2) require the state spaces $\tilde{\Omega}$ and Ω of $\tilde{P}$ and P, respectively, to be the same. In the present case, although they are not the same, the propositions are still applicable in view of the fact that we are able to define a bijection $f : \tilde{\Omega} \to \Omega$ between them.

• **Bijection via binary strings.** The bijection is easiest to describe through a bijection from each to another Catalan structure – the set of binary strings of length $2n$ with equal number of 1's and 0's, wherein the number of 1's in each string is always greater than or equal to the number of 0's as we count from left to right in the string. The bijection is illustrated in Fig. 3.

Given a Dyck path of length $2n$, a segment of slope $+1$ corresponds to a 1 and a segment of slope -1 corresponds to a 0. Given a binary string on n internal nodes, label the *left edges* (edges leading to left descendants) with a 1 and the *right edges* with a 0. Now the corresponding binary string is the one obtained by reading the labels as the tree is traversed, recursively, starting from the root, first visiting the left subtree and then the right subtree. Note that this provides a binary string with the above mentioned property, since every node (in the binary tree) which has a right descendant also has a left descendant. It is also easy to see that to each binary string with the above property, there is a unique Dyck path and a unique binary tree which corresponds to it.

The canonical path γ_{xy} of the comparison technique, is now a path in P, which can be described as a sequence of states, $f(x) = z_0, z_1, \ldots, z_{k-1}, z_k = f(y)$, for $(x, y) \in E(DG(n))$. The description of the paths will be simplified by using the above binary string representation of Dyck paths from now on.

Following [**15**], we will state some definitions. A *subtree* of a binary tree is either a single node or a binary tree with at least one internal node. Subtrees will be denoted by T_i, and this will stand both for the subtree and the binary string representation of T_i. The *depth* of a node in a binary tree is the length of the shortest path from the root to that node. It is convenient to view these binary trees as binary search trees with labels on the nodes, namely, with the property that the label of a node is bigger than the labels of all the nodes in its left subtree, and smaller than the labels of

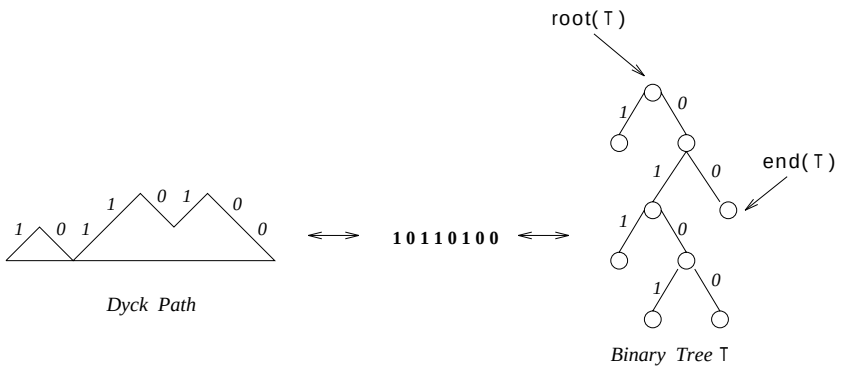

FIGURE 3. Bijection via binary strings

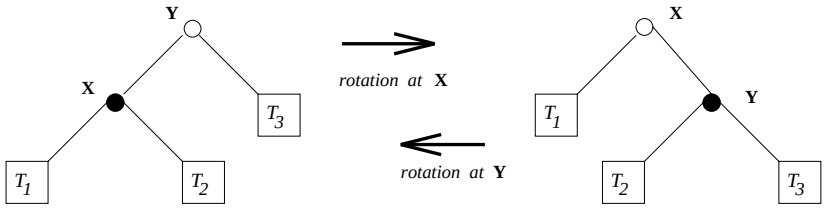

FIGURE 4. Local moves on binary trees

the nodes in its right subtree. This gives a natural ordering on the nodes of a tree. The rotation operation mentioned earlier preserves this node ordering, not surprisingly, since the rotations were invented as a way of restructuring binary search trees. Specifically, the rotation operation is defined as shown in Fig. 4.

Rotations do not change the number of internal nodes, and for a tree with n internal nodes, there are $n-1$ rotations possible at any time, one for each internal node, except for the root. The rotation at a node brings the node one step closer to the root, thus decreasing the depth of that node by exactly one. Although a rotation changes the local structure of the tree, it leaves the rest of the tree intact.

Referring to Fig. 4 above, a rotation at X decreases the depth of nodes in subtree T_1 by 1, increases the depth of nodes in subtree T_3 by 1, and leaves the depth of all other nodes the same. In the same way, a rotation at Y decreases the depth of nodes in subtree T_3 by 1, increases the depth of nodes in subtree T_1 by 1, and leaves the depth of all other nodes the same. We call each rotation either an X-rotation or a Y-rotation depending on whether it is of the form of a rotation at X in Fig. 4 above or of the form of a rotation at Y respectively.

• **Canonical paths** , . We need to define , $= \{\gamma_{xy}\}$, for $(x, y) \in E(\tilde{P})$. This can be done in a natural way, once we analyze a transition from x to

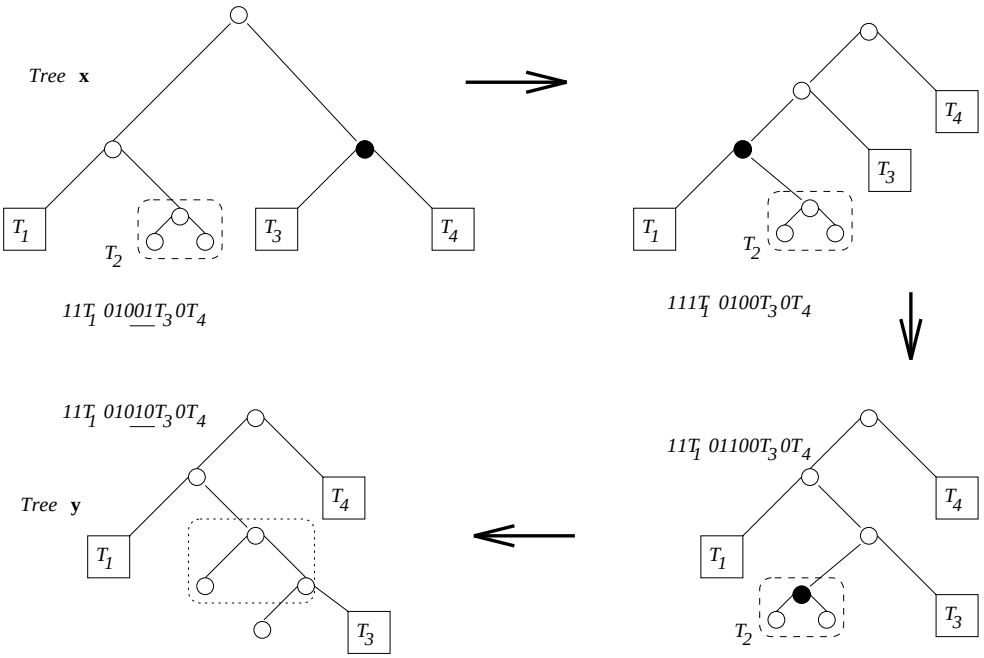

FIGURE 5. Canonical path from $\mathbf{x}$ to $\mathbf{y}$, differing in a 01–10 flip

y in $DG(n)$, according to whether it is a $01 \to 10$ flip or a $10 \to 01$ flip, and interpret the flip in terms of the corresponding binary trees, $f(x)$ and $f(y)$.

The easy case is if $f(x)$ and $f(y)$ differ by a single rotation – if $f(x) \sim f(y)$ in P, then γ_{xy} is simply the edge $(f(x), f(y)) \in E(P)$. But in general, $(x, y) \in E(\tilde{P})$ does not imply $(f(x), f(y)) \in E(P)$. See for example, Fig. 5, our final illustration which characterizes the differences in two binary trees, which have adjacent representations as Dyck paths. In such a case, we will define a unique sequence of rotations which transforms $f(x)$ into $f(y)$, and the corresponding sequence of edges in P which forms the corresponding canonical path γ_{xy}. We shall do this for the case when $x \to y$ is a $01 \to 10$ flip, and in the other case the path is just the reverse of the path in this case; we are justified in doing this since the interchange graphs can be viewed as undirected graphs.

To simplify the discussion, we will now introduce some new terms which are illustrated in Fig. 3. The *root* of a subtree T_i, $root(T_i)$, is the top vertex of that subtree. The *end* of a subtree T_i, $end(T_i)$, is the rightmost vertex of T_i, or equivalently, the node with the largest label when the tree T_i is viewed as a binary search tree. (A sure way to find the $end(T)$, starting from the $root(T)$ for any tree T, is to keep taking the right branch for as long as possible!)

The following observation is key to understanding , and to bounding $A(,)$. A $01 \to 10$ flip moves a particular left subtree (T_3 in Fig. 5) hanging from the right child of some node N to being the rightmost subtree of the

left child of N. This is easy to verify by considering the binary strings corresponding to x and y, which differ in a $01 \to 10$ flip, and then by constructing $f(x)$ and $f(y)$. We call such a subtree the *characteristic subtree* of that particular $01 \to 10$ flip. In Fig. 5, T_3 is the characteristic subtree of the flip which takes x to y. Note that such a subtree can also be a single node. (For the $10 \to 01$ flip, just the opposite happens, defining a characteristic subtree in a similar way.) Thus an edge $(x, y) \in E(\tilde{P})$ is characterized by a unique subtree of the binary trees represented by $f(x)$ and $f(y)$.

The canonical path γ_{xy} is the unique sequence of rotations which transforms the binary tree $f(x)$ into $f(y)$ – the first rotation is performed at the parent of the root of the characteristic subtree, and then every subsequent rotation is at the sibling of the root of the characteristic subtree. (In Fig. 5 the black nodes denote the nodes at which rotations are performed; also for convenience, the trees are called $\mathbf{x}$ and $\mathbf{y}$, rather than $f(x)$ and $f(y)$.)

• **Bounding** $A(,\ )$. We will show that $A(,\ )$ is at most $O(n)$, by arguing that the length of a canonical path is at most n, and that each rotation in a binary tree is used by at most one canonical path γ_{xy}, where $x \to y$ is a $01 \to 10$ flip, and by at most one canonical path corresponding to a $10 \to 01$ flip.

First notice that in a path corresponding to a $01 \to 10$ flip, the depth of the root of the characteristic subtree remains the same after the first rotation, but increases by exactly one with every subsequent rotation. At most $n - 2$ internal nodes can participate (by being siblings of the root of the characteristic subtree) in increasing the depth – the grandparent of the characteristic subtree and any nodes in the right subtree of the right child of the grandparent are the nodes where a rotation is not performed in such a canonical path. This shows that the length of a canonical path can be at most $n - 1$. (The argument for a $10 \to 01$ flip is analogous.)

Secondly, consider an arbitrary rotation $(z, w) \in E(P)$. Whether it is an X-rotation or a Y-rotation, there are always at most two choices for a subtree to play the role of a characteristic subtree. Referring to Fig. 4, if (z, w) is an X-rotation, then either T_3 is the characteristic subtree of a $01 \to 10$ flip or T_2 is the characteristic subtree of a $10 \to 01$ flip. The rotation (z, w) and the choice of either T_2 or T_3 as the characteristic subtree, uniquely identifies the pair (x, y) such that $(z, w) \in \gamma_{xy}$. Referring once again to Fig. 4, if (z, w) is an X-rotation, then T_3 will eventually end up as the right subtree of $end(T_2)$, giving us y. Knowing y and the fact that T_3 is the characteristic subtree uniquely determines x. If on the other hand, T_2 were to be the characteristic subtree, then w is in fact y, and x can be uniquely determined given that $x \to y$ is now a $10 \to 01$ flip.

Thus, for a fixed $,\ (z, w), |,\ (z, w)| \leq 2$, and when $(z, w) \in \gamma_{xy}$, $|\gamma_{xy}| \leq n$.

We have $\pi(x) = \pi(z) = (n+1)/\binom{2n}{n}$. Also, $P(z,w) = 1/(2n-2)$, for all $(z,w) \in E(P)$, and $\tilde{P}(x,y) = 1/(4n-6)$, for all $(x,y) \in E(\tilde{P})$. Thus,

$$\begin{aligned} A(,\,) &= \max_{(z,w)\in E(P)} \left\{ \frac{1}{\pi(z)P(z,w)} \sum_{\Gamma(z,w)} |\gamma_{xy}|\pi(x)\tilde{P}(x,y) \right\} \\ &\leq \frac{2n-3}{n-1} 2n = O(n). \end{aligned}$$

Applying Proposition 3, with the known bound on the mixing time of $\tilde{P}$, we can now bound the mixing time of P, the Markov chain on binary trees:

$$\begin{aligned} \tau(\epsilon) &\leq \frac{4\log\left(\frac{1}{\epsilon\pi_*}\right)}{\log\left(\frac{1}{2\epsilon}\right)} n(n^3(\log n + \log 1/\epsilon)) \\ &= \frac{4\left(\log(1/\epsilon) + \log(c_n)\right) n^4(\log n + \log 1/\epsilon)}{\log(1/2) + \log(1/\epsilon)} \\ &= O\left(\frac{n^5 \log n}{\log(1/\epsilon)} + n^4 \log n + n^5 + n^5 \log n\right) \\ &= O(n^5((\log n + \log 1/\epsilon)), \end{aligned}$$

thus establishing the following theorem.

THEOREM 4. *The mixing time of the Markov chain on triangulations of a convex $(n+2)$-gon (equivalently, on binary trees with n internal nodes) satisfies, for $0 < \epsilon < 1$,*

$$\tau(\epsilon) = O\left(n^5 \log(n/\epsilon)\right).$$

4. Conclusions

It is to be noted that, while Wilson showed his (upper) bound is tight up to a multiplicative constant, we do not believe our bound to be tight for the chain on triangulations. To get a better estimate we believe that one either needs a direct argument (say, via coupling) or a better comparison – comparison with a faster Catalan structure. Our present candidate for such a potentially faster chain is the one on the set of nonintersecting chord diagrams with $2n$ equally spaced points on a circle. It can be shown, with a proof by induction on n, that the diameter of this graph is $n-1$, whereas the diameter of $DG(n)$ is $\Theta(n^2)$. It can also be shown, with a suitable bijection, that $DG(n)$ is a proper subgraph of this graph, for $n \geq 3$. This is part of the rationale for our belief that the graph on nonintersecting chord diagrams is a better comparison candidate. Work is in progress on this problem. Of course, by comparison with $DG(n)$, we can always get *some* bound on the mixing time of this candidate problem: for example, by choosing our paths in the obvious way — it can be checked that adjacency in $DG(n)$ gives

adjacency in the chord interchange graph, making each path a path of *length one* — we get $A(,)$ is $O(1)$, and so, using Proposition 3, an upper bound for the mixing time of the Markov chain on the chord interchange graph is $O(n^5(\log n + \log 1/\epsilon))$. (Also to be noted here is that the positive transition probabilities in the chord interchange graph are lower bounded by $\Omega(1/n^2)$.) But surely, a much better bound is the truth!

We also intend to make a systematic study of several other Catalan structures, with the hope of obtaining tight bounds on the diameters, the eigenvalues, and the rates of mixing of Markov chains on the associated interchange graphs.

Acknowledgments. The second author is grateful to Bob Dobrow for bringing the triangulation walk to his (the author's) attention, and for useful discussions. The authors also thank Dana Randall and Bill Steiger for valuable comments.

References

[1] M. Atkinson and J. Sack, Generating binary trees at random, *Info. Proc. Letters*, **41** (1992), 21-23.

[2] P. Diaconis and L. Saloff-Coste, Comparison technique for random walks on finite groups, *Ann. Probability*, **21** (1993), 2131-2156.

[3] P. Diaconis and L. Saloff-Coste. Comparison theorems for reversible Markov chains. *Ann. Appl. Probability* **3**, 1993, pp. 696-730.

[4] P. Diaconis and L. Saloff-Coste, Logarithmic Sobolev inequalities for finite Markov chains, *Ann. Appl. Prob.*, **6** (1996), 695-750.

[5] R.P. Dobrow and J.A. Fill, On the Markov chain for the move-to-root rule for binary search trees, *Ann. Appl. Prob.* **5** (1995), 1-19.

[6] R.P. Dobrow and J.A. Fill, Rates of convergence for the move-to-root Markov chain for binary search trees, *Ann. Appl. Prob.* **5** (1995), 20-36.

[7] M. Gardner, Time travel and other mathematical bewilderments, Freeman, New York, NY (1987).

[8] R. Graham, D. Knuth, and O. Patashnik, *Concrete Mathematics: a foundation for computer science*, Addison-Wesley (1979), Massachusetts.

[9] B. Johnsen, Generating binary trees with uniform probability, *BIT* **31** (1991), 15-31.

[10] C. Kenyon and W. Thurston, Rotation distance using flows, preprint.

[11] M. Luby, D. Randall, and A. Sinclair, Markov chain algorithms for planar lattice structures, *Proc. of the 36th IEEE Symp. on Foundations of Computer Science*, (1995).

[12] M. Molloy, B. Reed, and W. Steiger, On the mixing rate of the triangulation walk, preprint (November, 1997).

[13] D. Randall and P. Tetali. Analyzing Glauber dynamics by comparison of Markov chains. Proceedings of *LATIN'98*; also to appear in *Lecture Notes In Computer Science.*

[14] A.J. Sinclair. *Algorithms for random generation and & counting: a Markov chain approach.* Birkhäuser, Boston, 1993, p.61.

[15] D. Sleator, R. Tarjan and W. Thurston. Rotation distance, triangulations, and hyperbolic geometry. *J.AMS.* **1**, 1988.

[16] R. Stanley. "Problems on Catalan and related numbers" in *Enumerative Combinatorics*, Volume II. Version of 12 November 1997, to be published by Cambridge University Press.

[17] D. Wilson. Mixing times of lozenge tiling and card shuffling Markov chains, draft of manuscript.

School of Mathematics, Georgia Tech, Atlanta, GA 30332-0160
E-mail address: mcshine@math.gatech.edu

School of Mathematics, Georgia Tech, Atlanta, GA 30332-0160
E-mail address: tetali@math.gatech.edu