Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Combinatorial Theory, Series A 104 (2003) 49–62

Journal of Combinatorial Theory Series A

http://www.elsevier.com/locate/jcta

The Catalan matroid

Federico Ardila

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 17 July 2002

Abstract

We show how the set of Dyck paths of length $2n$ naturally gives rise to a matroid, which we call the "Catalan matroid" $\mathbf{C}_n$. We describe this matroid in detail; among several other results, we show that $\mathbf{C}_n$ is self-dual, it is representable over $\mathbb{Q}$ but not over finite fields $\mathbb{F}_q$ with $q \leqslant n-2$, and it has a remarkably nice Tutte polynomial. We then generalize our construction to obtain a family of matroids, which we call "shifted matroids". They are precisely the matroids whose independence complex is a shifted simplicial complex.

Keywords: Dyck path; Matroid; Tutte polynomial; Shifted complex

1. Introduction

A Dyck path of length $2n$ is a path in the plane from $(0,0)$ to $(2n,0)$, with steps $(1,1)$ and $(1,-1)$, that never passes below the x-axis. It is a classical result (see for example [14, Corollary 6.2.3.(iv)]) that the number of Dyck paths of length $2n$ is equal to the Catalan number $C_n = \frac{1}{n+1}\binom{2n}{n}$.

Each Dyck path P defines an *up-step set*, consisting of the integers i for which the ith step of P is $(1,1)$. The starting point of this paper is Theorem 2.1. It states that the collection of up-step sets of all Dyck paths of length $2n$ is the collection of bases of a matroid. Most of this paper is devoted to the study of this matroid, which we call the *Catalan matroid*, and denote $\mathbf{C}_n$.

Section 2 starts by proving Theorem 2.1. As we know, there are many equivalent ways of defining a matroid: in terms of its rank function, its independent sets, its flats, and its circuits, among others. The rest of Section 2 is devoted to describing some of these definitions for $\mathbf{C}_n$.

E-mail address: fardila@math.mit.edu.

0097-3165/$ - see front matter
doi:10.1016/S0097-3165(03)00121-3

In Section 3, we compute the Tutte polynomial of the Catalan matroid. We find that it enumerates Dyck paths according to two simple statistics. Some nice enumerative results are derived as a consequence.

In Section 4, we generalize our construction of $\mathbf{C}_n$ to a wider class of matroids, which we call *shifted matroids*. They are precisely the matroids whose independence complex is a shifted simplicial complex. We describe the homotopy type of the independence complex of Catalan and shifted matroids. We then generalize our construction in a different direction to obtain, for any finite poset P and any order ideal I, a shifted family of sets. This family is not always the set of bases of a matroid.

Finally, in Section 5 we address the question of representability of the matroids we have constructed. We show that the Catalan matroid, and more generally any shifted matroid, is representable over $\mathbb{Q}$. In the opposite direction, we show that $\mathbf{C}_n$ is *not* representable over the finite field $\mathbb{F}_q$ if $q \leqslant n-2$.

Throughout this paper, we will assume some familiarity with the basic concepts of matroid theory. For instance, Chapter 1 of [10] should be enough to understand most of the paper. We also highly recommend Section 6.2 and Exercises 6.19–6.37 of [14] for an encyclopedic treatment of Catalan numbers and related topics.

2. The matroid

Let n be a fixed positive integer. Consider all paths in the plane which start at the origin and consist of $2n$ steps, where each step is either $(1,1)$ or $(1,-1)$. We will call such steps *up-steps* and *down-steps*, respectively. From now on, the word *path* will always to refer to a path of this form.

Such paths are in bijection with subsets of $[2n]$. To each path P, we can assign the set of integers i for which the ith step of P is an up-step. We call this set the *up-step set* of P. Conversely, to each subset $A \subseteq [2n]$, we can assign the path whose ith step is an up-step if and only if i is in A.

To simplify the notation later on, we will omit the brackets when we talk about subsets of $[2n]$. We will also use subsets of $[2n]$ and paths interchangeably. For example, for $n=3$, the path 13 will be the path with up-steps at steps 1 and 3, and down-steps at steps 2, 4, 5 and 6.

A useful statistic to keep track of will be the *height of path P at x*; i.e., the height of the path after taking its first x steps. We shall denote it $\mathrm{ht}_P(x)$; it is equal to $2|P_{\leqslant x}| - x$, where $P_{\leqslant x}$ denotes the set of elements of P which are less than or equal to x. Also, let minht_P and maxht_P be the minimum and maximum heights that P achieves, respectively.

Theorem 2.1. *Let $\mathscr{B}_n$ be the collection of up-step sets of all Dyck paths of length $2n$. Then $\mathscr{B}_n$ is the collection of bases of a matroid on the set $[2n]$.*

To prove Theorem 2.1, we will use the following lemma.

Lemma 2.2. *The collection $\mathcal{B}_n$ consists of all the sets of positive integers $\{a_1 < a_2 < \cdots < a_n\}$ such that $a_i \leqslant 2i - 1$ for $1 \leqslant i \leqslant n$.*

Proof. A path $\{a_1 < \cdots < a_n\}$ is Dyck if and only if, for each i with $1 \leqslant i \leqslant n$, the ith up-step comes before the ith down-step; that is, if and only if $a_i \leqslant 2i - 1$. $\square$

Proof of Theorem 2.1. It is easy to check directly that $\mathcal{B}_n$ satisfies the basis exchange axiom; we invite the reader to carry out this proof. We now present a shorter proof, using a connection with transversal matroids suggested by several people, including the authors of [3] and one of the anonymous referees.

Recall that, given a finite set S and a collection $\mathcal{A} = (A_1, A_2, \dots, A_n)$ of subsets of S, a *transversal* of $\mathcal{A}$ is a set of n distinct elements of S which can be labeled $s_1, \dots, s_n$ so that $s_i \in A_i$ for $1 \leqslant i \leqslant n$. It is well-known that the collection of all transversals of $\mathcal{A}$ is the collection of bases of a matroid on S. Such a matroid is called a *transversal matroid*, and the collection $\mathcal{A}$ is called a *presentation* of the matroid. For more information on transversal matroids, see for example Chapter 7 of [18].

We will prove that $\mathcal{B}_n$ is precisely the collection of transversals of the collection $\mathcal{A}_n = ([1], [3], [5], \dots, [2n-1])$. It will follow that $\mathcal{B}_n$ is the collection of bases of the transversal matroid with presentation $\mathcal{A}_n$.

Lemma 2.2 makes it clear that every set in $\mathcal{B}_n$ is a transversal of $\mathcal{A}_n$. For the converse, consider a transversal A of $\mathcal{A}_n$; say $A = \{a_1, a_2, \dots, a_n\}$, where $a_i \in [2i-1]$ for $1 \leqslant i \leqslant n$. Let $b_1 < b_2 < \cdots < b_n$ be the increasing rearrangement of A. Since $a_1, a_2, \dots, a_i$ are all less than or equal to $2i - 1$, it follows that b_i, which is the ith smallest element of A, must also be less than or equal to $2i - 1$; i.e., $b_i \in [2i-1]$. It follows that $A \in \mathcal{B}_n$. $\square$

From Theorem 2.1, we have a unique matroid on the ground set $[2n]$ whose collection of bases is $\mathcal{B}_n$. We will call it the *Catalan matroid of rank n* (or simply the *Catalan matroid*), and denote it by $\mathbf{C}_n$. This paper is mostly devoted to the study of this matroid.

Proposition 2.3. *The rank function of $\mathbf{C}_n$ is given by*

$$r(A) = n + \lfloor \operatorname{minht}_A / 2 \rfloor$$

for each $A \subseteq [2n]$.

Proof. Fix a subset $A \subseteq [2n]$, and let $\operatorname{minht}_A = -y$, where y is a non-negative integer. Also, let x be the smallest integer such that $\operatorname{ht}_A(x) = \operatorname{minht}_A$.

Recall that the rank of a subset A of $[2n]$ is equal to the largest possible size of an intersection $A \cap B$, where B is a basis of $\mathbf{C}_n$.

The path A is at height $-y$ after taking $|A_{\leqslant x}|$ up-steps and $x - |A_{\leqslant x}|$ down-steps, so $|A_{\leqslant x}| = (x-y)/2$. Also, for any basis B, we have that $|B_{>x}| \leqslant n - x/2$, since $\mathrm{ht}_B(x) \geqslant 0$. Hence

$$|A \cap B| = |(A \cap B)_{\leqslant x}| + |(A \cap B)_{>x}| \leqslant |A_{\leqslant x}| + |B_{>x}| \leqslant n - y/2.$$

We conclude that $r(A) \leqslant n + \lfloor \mathrm{minht}_A/2 \rfloor$.

Now we need a basis B with $|A \cap B| = n + \lfloor \mathrm{minht}_A/2 \rfloor$. We construct it as follows. First, add to A the smallest $a = \lceil y/2 \rceil$ numbers that it is missing, to obtain the set A'. Then $\mathrm{ht}_{A'}(x) = 2a - y \geqslant 0$; in fact, it is clear that the path A' never crosses the x-axis. Let $|A| = n + h$ for some integer h; then $\mathrm{ht}_A(2n) = 2h$ and $\mathrm{ht}_{A'}(2n) = 2h + 2a$. Now remove from A' the largest $h + a$ numbers that it contains, to obtain the set B. It is again easy to see that the path B never crosses the x-axis, and ends at $(2n, 0)$. So B is Dyck, and

$$|A \cap B| = |A \cap A'| - (h + a) = |A| - (h + a) = n - a$$

as desired. $\square$

Now that we know the rank function of $\mathbf{C}_n$, we describe several important classes of subsets of the matroid in Propositions 2.4–2.8. We will only provide a proof for Proposition 2.4; the remaining proofs are similar in flavor. The interested reader may want to complete the details to get better acquainted with the matroid $\mathbf{C}_n$.

Proposition 2.4. *The flats of* $\mathbf{C}_n$ *are: the set* $[2n]$, *and the subsets* $A \subseteq [2n]$ *such that*

(i) minht_A *is odd, and*
(ii) *if* $\mathrm{ht}_A(x) = \mathrm{minht}_A$, *then* $\{x+1, \ldots, 2n\} \subseteq A$.

Proof. Let A be a flat of $\mathbf{C}_n$ other than $[2n]$, and let x be such that $\mathrm{ht}_A(x) = \mathrm{minht}_A$. If some integer y with $x + 1 \leqslant y \leqslant n$ was not in A, then we would clearly have $\mathrm{minht}_{A \cup y} = \mathrm{minht}_A$ and thus $r(A \cup y) = r(A)$, contradicting the assumption that A is a flat. Therefore, any flat must satisfy condition (ii).

Also, if we had a flat A with $\mathrm{minht}_A = -2h$ achieved at $\mathrm{ht}_A(x)$, then we would have $x \notin A$, and $\mathrm{minht}_{A \cup x} = -2h + 1$ would be achieved at $\mathrm{ht}_{A \cup x}(x - 1)$. We would then have $r(A \cup x) = r(A)$, again a contradiction. So any flat A must also satisfy condition (i).

Conversely, assume that A satisfies conditions (i) and (ii). Let $\mathrm{minht}_A = -(2k+1)$, which can only be achieved once, say at $\mathrm{ht}_A(x)$. Any y which is not in A must be less than or equal to x; and we have $\mathrm{minht}_{A \cup y} = -(2k - 1)$ if $y < x$, or $\mathrm{minht}_{A \cup y} = -2k$ if $y = x$. In either case, $r(A \cup y) = r(A) + 1$. This completes the proof. $\square$

Proposition 2.5. *The independent sets of* $\mathbf{C}_n$ *are the subsets* $A \subseteq [2n]$ *such that* $\mathrm{minht}_A = \mathrm{ht}_A(2n)$.

Proposition 2.6. *The spanning sets of* $\mathbf{C}_n$ *are the subsets* $A \subseteq [2n]$ *such that* $\mathrm{minht}_A = 0$.

Proposition 2.7. *The circuits of* $\mathbf{C}_n$ *are the subsets* $A \subseteq [2n]$ *of the form* $A = \{2k, 2k + b_1, \dots, 2k + b_{n-k}\}$, *for some positive integer* $k \leqslant n$ *and some Dyck path* $\{b_1, \dots, b_{n-k}\}$ *of length* $2(n-k)$.

Proposition 2.8. *The cocircuits of* $\mathbf{C}_n$ *are the subsets* $A \subseteq [2n]$ *such that*

(i) $\mathrm{maxht}_A = 1$, *and*
(ii) *if* $\mathrm{ht}_A(x) = 1$, *then* A *has no elements greater than* x.

We complete this section with an observation which is interesting in itself, and will also be important to us in Section 3.

Proposition 2.9. *The Catalan matroid is self-dual.*[1]

Proof. Say $B = \{b_1, \dots, b_n\}$ is a basis of $\mathbf{C}_n$, and let $[2n] - B = \{c_1, \dots, c_n\}$ be the corresponding basis of the dual matroid $\mathbf{C}_n^*$. Then $\{2n+1-c_n, \dots, 2n+1-c_1\}$ is a Dyck path; in fact, it is the path obtained by reflecting the Dyck path B across a vertical axis. So the bases of $\mathbf{C}_n^*$ are simply the up-step sets of all Dyck paths of length $2n$, under the relabeling $x \to 2n+1-x$. Thus $\mathbf{C}_n^* \cong \mathbf{C}_n$. $\square$

3. The Tutte polynomial

Given a matroid M over a ground set S, its Tutte polynomial is defined as

$$T_M(q,t) = \sum_{A \subseteq S} (q-1)^{r(S)-r(A)} (t-1)^{|A|-r(A)}.$$

For our purposes, it is more convenient to define the Tutte polynomial in terms of the internal and external activity of the bases. We recall this definition now.

We first need to fix an arbitrary linear ordering of S.

For any basis B and any element $e \notin B$, the set $B \cup e$ contains a unique circuit. If e is the smallest element of that circuit with respect to our fixed linear order, then we say that e is *externally active* with respect to B. The number of externally active elements with respect to B is called the *external activity* of B; we shall denote it by $e(B)$.

Dually, for any basis B and any element $i \in B$, the set $S - B \cup i$ contains a unique cocircuit. If i is the smallest element of that cocircuit, then we say that i is *internally active* with respect to B. The number of internally active elements with respect to B is called the *internal activity* of B; we shall denote it by $i(B)$.[2]

[1] We follow Oxley [10] in calling a matroid M *self-dual* if $M \cong M^*$. It is worth mentioning, however, that some authors reserve the term "self-dual" for matroids M such that $M = M^*$.

[2] The internally active elements with respect to a basis B of M are precisely the externally active elements with respect to the basis $S - B$ of the dual matroid M^*. That is why we say that internal activity and external activity are dual concepts.

Proposition 3.1 (Crapo [4]). *For any matroid M and any linear order of its ground set*,

$$T_M(q,t) = \sum_{B \text{ basis}} q^{i(B)} t^{e(B)}.$$

We will use Proposition 3.1 to study the Tutte polynomial of the Catalan matroid. The first thing to do is to fix a linear order of its ground set, $[2n]$. We will use the most natural choice: $1 < 2 < \cdots < 2n$. Now we compute the internal and external activity of each basis of $\mathbf{C}_n$.

Lemma 3.2. *The internal activity of a Dyck path B is equal to the number of up-steps that B takes before its first down-step.*

Proof. Let $i \in B$. The path $[2n] - B$ never goes above height 0; the path $[2n] - B \cup i$ goes up to height 2. Let j be the smallest integer such that $\mathrm{ht}_{[2n]-B\cup i}(j) = 1$. Clearly $j \geqslant i$.

Let D be the unique cocircuit of $\mathbf{C}_n$ which can be obtained by deleting some elements of $[2n] - B \cup i$. We cannot delete any element less than or equal to j, or else the resulting path will not reach height 1. We must delete any element larger than j by Proposition 2.8. So $D = ([2n] - B)_{\leqslant j}$.

Therefore, i is the smallest element of D if and only if B contains all of $1, 2, \ldots, i-1$. This completes the proof. $\square$

Lemma 3.3. *The external activity of a Dyck path B is equal to the number of positive integers x for which $\mathrm{ht}_B(x) = 0$.*

Proof. Let $e \notin B$. The path $B \cup e$ ends at height 2; let $2k-1$ be the largest integer such that $\mathrm{ht}_{B\cup e}(2k-1) = 1$. Clearly $2k - 1 < e$.

We start by showing that the unique circuit C of $\mathbf{C}_n$ contained in $B \cup e$ is $(B \cup e)_{\geqslant 2k}$.

Since $C \subseteq B \cup e$, we have that $\mathrm{ht}_C(2n) - \mathrm{ht}_C(2k-1) \leqslant \mathrm{ht}_{B\cup e}(2n) - \mathrm{ht}_{B\cup e}(2k-1) = 1$. Equality holds if and only if every up-step of $B \cup e$ after the $(2k-1)$th is also an up-step of C; i.e., when $(B\cup e)_{\geqslant 2k} = C_{\geqslant 2k}$.

But it is clear from Proposition 2.7 that $\mathrm{ht}_C(2n) - \mathrm{minht}_C = 1$, and that minht_C is only achieved at $\mathrm{ht}_C(\min C - 1)$. So the above inequality can only hold if $\min C = 2k$. Thus $C = C_{\geqslant 2k} = (B \cup e)_{\geqslant 2k}$ as desired.

Now we know that $\min C = 2k$, so e is externally active if and only if $e = 2k$. If $\mathrm{ht}_B(e) = 0$, this is clearly the case. On the other hand, if $\mathrm{ht}_B(e) \geqslant 1$, then $\mathrm{ht}_{B\cup e}(e-1) = \mathrm{ht}_B(e-1) \geqslant 2$, so this is not the case. This completes the proof. $\square$

Theorem 3.4. *For a Dyck path P, let $a(P)$ denote the number of up-steps that P takes before its first down-step, and let $b(P)$ denote the number of positive integers x for which $\mathrm{ht}_P(x) = 0$.*

Then the Tutte polynomial of the Catalan matroid $\mathbf{C}_n$ *is equal to*

$$T_{\mathbf{C}_n}(q,t) = \sum_{P \text{ Dyck}} q^{a(P)} t^{b(P)},$$

where the sum is over all Dyck paths of length $2n$.

Proof. This follows immediately from Proposition 3.1 and Lemmas 3.2 and 3.3. □

Corollary 3.5. *The polynomial*

$$\sum_{P \text{ Dyck}} q^{a(P)} t^{b(P)},$$

is symmetric in q *and* t.

Proof. It is well-known that, for any matroid M, we have $T_{M^*}(q,t) = T_M(t,q)$. The result follows from Proposition 2.9 and Theorem 3.4. □

It is a known fact that the statistics $a(P)$ and $b(P)$ are equidistributed over the set of Dyck paths of length $2n$: the number of paths with $a(P) = k$ and the number of paths with $b(P) = k$ are both equal to $\frac{k}{2n-k}\binom{2n-k}{n}$. For the first equality, see for example [16]; for the second, see [9, Eq. (7)].

Corollary 3.5 was also discovered independently by James Haglund [6]. It is not difficult to prove it directly; in fact, it will be an immediate consequence of our next theorem.

Theorem 3.6. *Let* $C(x) = \frac{1}{2}(1 - \sqrt{1-4x}) = C_0 + C_1 x + C_2 x^2 + \cdots$ *be the generating function for the Catalan numbers. Then*

$$\sum_{n \geqslant 0} T_{\mathbf{C}_n}(q,t) x^n = \frac{1 + (qt - q - t)xC(x)}{1 - qtx + (qt - q - t)xC(x)}.$$

Proof. A Dyck path P of length $2n \geq 2$ can be decomposed uniquely in the standard way: it starts with an up-step, then it follows a Dyck path P_1 of length $2r$, then it takes a down-step, and it ends with a Dyck path P_2 of length $2s$, for some non-negative integers r, s with $r + s = n - 1$. More precisely, and necessarily more confusingly,

$$P = \{1, 1 + p_1, 1 + p_2, \ldots, 1 + p_r, 2r + 2 + q_1, 2r + 2 + q_2, \ldots, 2r + 2 + q_s\}$$

for some Dyck paths $\{p_1, \ldots, p_r\}$ and $\{q_1, \ldots, q_s\}$ with $r + s = n - 1$.

It is clear that in this decomposition we have $a(P) = a(P_1) + 1$ and $b(P) = b(P_2) + 1$. Therefore,

$$\begin{aligned} T_{\mathbf{C}_n}(q,t) &= \sum_{r+s=n-1} \sum_{P_1 \in \mathscr{B}_r} \sum_{P_2 \in \mathscr{B}_s} q^{a(P_1)+1} \, t^{b(P_2)+1} \\ &= qt \sum_{r+s=n-1} T_{\mathbf{C}_r}(q,1) T_{\mathbf{C}_s}(1,t) \end{aligned}$$

for $n \geqslant 1$; so if we write $\mathbf{T}(q,t,x) = \sum_{n \geqslant 0} T_{\mathbf{C}_n}(q,t)x^n$, we have

$$\mathbf{T}(q,t,x) = 1 + qtx\mathbf{T}(q,1,x)\mathbf{T}(1,t,x). \tag{1}$$

Now observe that $\mathbf{T}(1,1,x) = C(x)$. Setting $q = 1$ in (1) gives a formula for $\mathbf{T}(1,t,x)$, and setting $t = 1$ gives a formula for $\mathbf{T}(q,1,x)$. Substituting these two formulas back into (1), we get the desired result. $\square$

4. Shifted matroids

We now generalize our construction of $\mathbf{C}_n$ to a larger family of matroids, which we call *shifted matroids*. There is one shifted matroid for each non-empty set $S = \{s_1 < \cdots < s_n\}$ of positive integers, which we shall denote $\mathbf{SM}(s_1, \ldots, s_n)$.

Theorem 4.1. *Let $S = \{s_1 < \cdots < s_n\}$ be a set of positive integers, and let $\mathscr{B}_S$ be the collection of sets of positive integers $\{a_1 < \cdots < a_n\}$ such that $a_1 \leqslant s_1, \ldots, a_n \leqslant s_n$. Then $\mathscr{B}_S$ is the collection of bases of a matroid $\mathbf{SM}(s_1, \ldots, s_n)$ on the set $[s_n]$.*

Proof. We can repeat the argument of Theorem 2.1 to conclude that $\mathbf{SM}(s_1, \ldots, s_n)$ *is the transversal matroid with presentation $\mathscr{A}_S = ([s_1], [s_2], \ldots, [s_n])$.* $\square$

The Catalan matroid is a member of the shifted matroid family. From Lemma 2.2 we know that the Catalan matroid $\mathbf{C}_n$ is exactly the shifted matroid $\mathbf{SM}(1, 3, 5, \ldots, 2n-1)$, with an additional loop $2n$.

Shifted matroids have been discovered several times in the past. Welsh [17] used them to prove a lower bound for the number of non-isomorphic matroids on $[n]$. Oxley et al. [11] gave different characterizations of them. Bonin and de Mier [2], Bonin et al. [3] are currently studying a wider class of matroids which provides an excellent level of generality for matroid-theoretic structural considerations. It is worth pointing out that the *generalized Catalan matroids* of [3] are precisely our shifted matroids.

Now we present a new characterization of shifted matroids. Recall that an *abstract simplicial complex Δ* on $[n]$ is a family of subsets of $[n]$ (called *faces*) such that if $G \in \Delta$ and $F \subseteq G$, then $F \in \Delta$. A simplicial complex Δ is *shifted* if, for any face $F \in \Delta$ and any pair of elements $i < j$ such that $i \notin F$ and $j \in F$, the subset $F - j \cup i$ is also a face of Δ.

The family of independent sets of a matroid M is always a simplicial complex, called the *independence complex* or *matroid complex* of M. For shifted matroids, we have the following simple observation.

Proposition 4.2. *The independence complex of a shifted matroid* $\mathbf{SM}(s_1, \ldots, s_n)$ *is a shifted complex.*

Proof. If $F \subseteq [s_n]$ is independent, it is contained in some basis B. Now assume that we have two elements $i < j$ such that $i \notin F$ and $j \in F$, and let $G = F - j \cup i$. If the basis B

contains i, then it contains G. Otherwise, $B - j \cup i$ is also a basis: for any $1 \leqslant k \leqslant n$, its kth smallest is less than or equal to the kth smallest element of B, which is less than or equal to s_k. This basis contains G. In both cases, we conclude that G is independent. □

Klivans [8] characterizes *shifted matroid complexes*: shifted complexes which are independence complexes of matroids. Her results and ours were discovered almost simultaneously. When we sat down to discuss them, we realized that her shifted matroid complexes were precisely the independence complexes of the matroids $\mathbf{SM}(s_1, \dots, s_n)$. This is why these matroids were baptized "shifted matroids".

Proposition 4.3 (Klivans, [8]). *If the independence complex of a loop-less matroid M is a shifted complex, then $M \cong \mathbf{SM}(s_1, \dots, s_n)$ for some positive integers $s_1 < \cdots < s_n$.*

Corollary 4.4. *The independence complex of a matroid M is a shifted complex if and only if M is a shifted matroid.*

One of the main reasons to study shifted complexes is the simplicity of their topology. For any simplicial complex Γ, one can construct a shifted simplicial complex $\Delta(\Gamma)$. This shifted complex preserves many combinatorial and topological properties of Γ, but is much easier to study [7]. In particular, any shifted complex is homotopically equivalent to a wedge of spheres. Therefore, its homology groups have no torsion and its cohomology ring is trivial. We now use these facts to discuss the topology of Catalan matroid complexes and, more generally, shifted matroid complexes.

The Catalan matroid complex is contractible, since every basis contains 1. So is the complex of any shifted matroid $\mathbf{SM}(1, s_2, \dots, s_n)$. Instead, consider the *reduced Catalan matroid complex*: the independence complex of the Catalan matroid with the coloop 1 deleted.

Proposition 4.5. *The reduced Catalan matroid complex is homotopically equivalent to a wedge of C_{n-1} $(n-2)$-dimensional spheres.*

Let $1 < s_1 < \cdots < s_n$ be integers. The independence complex of $\mathbf{SM}(s_1, \dots, s_n)$ is homotopically equivalent to a wedge of $(n-1)$-dimensional spheres. The number of spheres is equal to the number of bases of $\mathbf{SM}(s_1 - 1, \dots, s_n - 1)$.

Proof. We use Lemma 3.1 of [7]. Any shifted complex is homotopically equivalent to a wedge of spheres, possibly of different dimensions. In our case, since matroid complexes are pure, the spheres must have the same dimension.

For the reduced Catalan matroid complex, the spheres are $(n-2)$-dimensional. The number of them is equal to the number of maximal faces of the complex which do not contain 2; that is, the number of Dyck paths of length $2n$ whose second step is a down-step. There are C_{n-1} such paths.

For the independence complex of $\mathbf{SM}(s_1, \dots, s_n)$, the spheres are $(n-1)$-dimensional. The number of them is equal to the number of maximal faces of the complex which do not contain 1. Subtracting 1 from the labels of these faces puts them in bijective correspondence with the bases of $\mathbf{SM}(s_1-1,\dots,s_n-1)$. $\square$

Theorem 4.1 and Propositions 4.2 and 4.3 have a nice application to Young tableaux. Recall that a *partition* $\lambda=(\lambda_1,\dots,\lambda_k)$ of n is a weakly decreasing sequence of positive integers which add up to n. We associate to it a *Young diagram*: a left-justified array of unit squares, which has λ_i squares on the ith row from top to bottom.[3] A *standard Young tableaux* is a placement of the integers $1,\dots,n$ in the squares of the Young diagram, in such a way that the numbers are increasing from left to right and from top to bottom.

These definitions will be sufficient for our purposes. For a much deeper treatment of the theory of Young tableaux, we refer the reader to [5].

Corollary 4.6. *Let λ be a partition. Define the* first row set *of a standard Young tableau T of shape λ to be the set of entries which appear in the first row of T. Then the collection of first row sets of all standard Young tableaux of shape λ is the collection of bases of a shifted matroid.*

Proof. Let $\lambda'=(\lambda_1{}',\dots,\lambda_n{}')$ be the conjugate partition of λ, so $\lambda_i{}'$ is the number of squares on the ith column of the Young diagram of λ. Let $s_i=1+\lambda_1{}'+\cdots+\lambda_{i-1}{}'$ for $1\leqslant i\leqslant n$.

Let $\{b_1<\cdots<b_n\}$ be the first row set of a standard Young tableau T of shape λ. The first entry on the ith column of T is b_i; it is smaller than every entry to its southeast. There are only $\lambda_1{}'+\cdots+\lambda_{i-1}{}'$ cells which are not to its southeast, so $b_i\leqslant s_i$.

Conversely, if $B=\{b_1<\cdots<b_n\}$ is such that $b_i\leqslant s_i$ for $1\leqslant i\leqslant n$, then we can construct a standard Young tableau with first row set B. To do it, we first put the elements of B in order on the first row of λ. Then we put the remaining numbers from 1 to $|\lambda|$ on the remaining cells going in order down the columns, starting with the leftmost column. The inequalities $b_i\leqslant s_i$ guarantee that this process does indeed give a Young tableau T.

It follows that the collection in question is simply the collection of bases of the matroid $\mathbf{SM}(s_1,\dots,s_n)$. $\square$

We might try to generalize Corollary 4.6, replacing the first row of λ by any partition $\mu\subseteq\lambda$. Define the *μ-set* of a standard Young tableau T of shape λ to be the set of entries which appear in the sub-shape μ in T.

It is not too difficult to see that we do not always get the collection of bases of a matroid with this construction. However, we can still say something interesting.

Proposition 4.7. *Let $\mu\subseteq\lambda$ be partitions. Then the collection $\mathscr{B}_{\lambda\mu}$ of μ-sets of all standard Young tableau of shape λ is a shifted family.*

[3] This is the English way of drawing Young diagrams; francophones draw them with λ_i squares on the ith row from bottom to top.

In fact, as suggested by Stanley [15], this result still holds in the following more general setting. Let P be a partially ordered set, or *poset*, of n elements. Recall that a subset I of P is an *order ideal* of P if, for any pair of elements $x, y \in P$ with $x <_P y$ and $y \in I$, we also have $x \in I$. Also recall that a *linear extension* of P is a bijection $f : P \to [n]$ such that $i <_P j$ implies that $f(i) < f(j)$. For more information on posets, we refer the reader to [13, Chapter 3].

Define the *I-set* of a linear extension f of P to be the set $\{f(i) : i \in I\}$.

Proposition 4.8. *Let P be a finite poset, and let I be an order ideal of P. Then the collection $\mathscr{B}_{P,I}$ of I-sets of all linear extensions of P is a shifted family.*

Proof of Proposition 4.8. We need to check that if we have a set $B \in \mathscr{B}_{P,I}$ and a pair of numbers $a < b$ such that $a \notin B$ and $b \in B$, then $B - b \cup a \in \mathscr{B}_{P,I}$. It is enough to show this for $a = b - 1$; the general case will then follow by induction on $b - a$.

So let f be a linear extension of P with I-set B, and let $b \in B$ be such that $b - 1 \notin B$. Let $b = f(i)$ and $b - 1 = f(p)$ where $i \in I$ and $p \in P - I$. Let $g : P \to [n]$ be defined by switching the values of f at i and p; i.e.,

$$g(x) = \begin{cases} f(x) & \text{if } x \notin \{i, p\}, \\ b - 1 & \text{if } x = i, \\ b & \text{if } x = p. \end{cases} \tag{2}$$

We claim that g is also a linear extension for P. An important observation is that i and p are incomparable in P. If we had $i < p$, then we would have $b = f(i) < f(p) = b - 1$. If we had $i > p$, then $i \in I$ would imply $p \in I$.

Since f is a linear extension, we know that $f(i) = b$ satisfies several inequalities: it must be greater than all the values that f takes on $P_{<i}$, and less than all the values that f takes on $P_{>i}$. But b is never compared to $b - 1$ here, since p and i are incomparable. Therefore, $b - 1$ also satisfies all those inequalities that b needs to satisfy.

Similarly, $b - 1$ is greater than all the values that f takes on $P_{<p}$ and less than all the values that f takes on $P_{>p}$. The number b also satisfies these inequalities.

So we can switch the values of $f(i)$ and $f(p)$, and the resulting function g defined by (2) will also be a linear extension of P. Also, the I-set of g is $B - b \cup (b - 1)$. This concludes the proof. □

Proof of Proposition 4.7. The cells of λ can be given a partial order P_λ: cell i is less than cell j in P_λ if and only if cell i is northwest of cell j in λ. The cells of μ define an order ideal I_μ of P_λ, and $\mathscr{B}_{\lambda\mu} = \mathscr{B}_{P_\lambda, I_\mu}$. Now use Proposition 4.8. □

In view of Corollary 4.6 and Proposition 4.8, a natural problem, suggested by Richard Stanley, is the following.

Problem 4.9. *Characterize the pairs (P, I) of a finite poset P and an order ideal I for which $B_{P,I}$ is the set of bases of a matroid.*

5. Representability

Piff and Welsh [12] showed that every transversal matroid is representable over fields of every characteristic. We use ideas essentially equivalent to theirs to give vector representations for the Catalan and shifted matroids, which are transversal.

To write down an explicit representation, we will use the notion of a generic collection of numbers. There are many different approaches that we could follow; one that is sufficient for our purposes is the following. Given a collection of real numbers $x_1, \dots, x_N$, let $x_S = \prod_{i \in S} x_i$ for each subset $S \subseteq [N]$ with $|S| = n$. Form all the $2^{\binom{N}{n}}$ possible sums of some of the x_S's. If these sums are all distinct, we will say that the initial collection of numbers is *n-generic*. Most N-tuples of real numbers are n-generic. A specific example is a set of algebraically independent real numbers. Another example is any sequence of positive integers which increases quickly enough; for example one such that, for each $1 \leqslant i \leqslant N$, we have $x_i > \sum x_S$ summing over all n-subsets S of $[i-1]$.

Theorem 5.1. *Let $v_1, \dots, v_{2n}$ be the columns of a matrix*

$$A = \begin{pmatrix} a_{11} & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 & 0 & \cdots & 0 & 0 \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & a_{n4} & a_{n5} & a_{n6} & \cdots & a_{n,2n-1} & 0 \end{pmatrix}$$

where the a_{ij}'s with $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant 2i-1$ are n-generic integers. Then the vector matroid of $\{v_1, \dots, v_{2n}\}$ is isomorphic to the Catalan matroid $\mathbf{C}_n$.

Theorem 5.2. *Let $s_1 < \cdots < s_n$ be arbitrary positive integers. Let $v_1, \dots, v_{s_n}$ be the columns of a matrix $A = (a_{ij})_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant s_n}$, where the a_{ij}'s with $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant s_i$ are n-generic, and the remaining a_{ij}'s are equal to 0. Then the vector matroid of $\{v_1, \dots, v_{s_n}\}$ is isomorphic to the shifted matroid $\mathbf{SM}(s_1, \dots, s_n)$.*

The above two theorems are not difficult to prove directly. Columns $c_1, \dots, c_n$ of the matrix are a basis for $\mathbb{R}^n$ if and only if the determinant of the corresponding $n \times n$ matrix M_c is non-zero. This determinant is a sum of $n!$ terms with plus and minus signs; by the n-genericity condition, if any of these terms is non-zero then $\det M_c \neq 0$. Since M_c has its non-zero entries on the bottom left, $\det M_c \neq 0$ if and only if all the diagonal entries of M_c are all non-zero; this is equivalent to $\{c_1, \dots, c_n\}$ being a basis of the shifted matroid. For more details, we refer the reader to [12].

Theorem 5.1 shows that the Catalan matroid is representable over $\mathbb{Q}$, or even over a sufficiently large finite field. In the other direction, we now show a negative result about representing $\mathbf{C}_n$ over finite fields.

Proposition 5.3. *The Catalan matroid* $\mathbf{C}_n$ *is* not *representable over the finite field* $\mathbb{F}_q$ *if* $q \leqslant n-2$.

Proof. It is known ([10, Proposition 6.5.2]) and easy to show that the uniform matroid $U_{2,k}$ is $\mathbb{F}_q$-representable if and only if $q \geqslant k-1$. A matroid containing it as a minor is not representable over $\mathbb{F}_q$ for $q \leqslant k-2$. This suggests that we should find the largest k for which $U_{2,k}$ is a minor of $\mathbf{C}_n$.

We can use the Scum theorem (Higgs, [10, Proposition 3.3.7]), which essentially says that, if a matroid has a certain minor, then it must have that same minor hanging from the top of its lattice of flats. Our question is then equivalent to finding the largest k for which there exists a rank-$(n-2)$ flat which is contained in k rank-$(n-1)$ flats.

Lemma 5.4. *Let A be a rank-$(n-2)$ flat, and let x be the smallest integer such that* $\mathrm{ht}_A(x) = -1$. *Then there are exactly* $\frac{x+3}{2}$ *rank-$(n-1)$ flats containing A.*

Proof of Lemma 5.4. We know from Propositions 2.3 and 2.4 that $\mathrm{minht}_A = -3$ and that, once the path A reaches height -3, say at $\mathrm{ht}_A(y)$, it only takes up-steps. We want to add elements to A to obtain a path which reaches a minimum height -1, and only takes up-steps after that.

Say that we add one element a to A. This new up-step at a comes before the yth, so $\mathrm{ht}_{A\cup a}(y) = -1$. If we do not want to add any more elements to A, we have to make sure that $A \cup a$ only reaches height -1 at y. For this to be true, we need the new up-step a to occur on or before the xth step. In A, there are $\frac{x+1}{2}$ down-steps up to the xth to choose from. Each one of these gives a rank-$(n-1)$ flat containing A.

On the other hand, if we are to add more elements to A to obtain a rank-$(r-1)$ flat B, they will all be less than y so we will have $\mathrm{ht}_B(y) > 0$. The minimum height in B must then be achieved at some z for which $\mathrm{ht}_A(z) = -1$. In fact, for this z to be unique, it must be the leftmost one; i.e., it must be x. So the only possibility is that $B = A_{\leqslant x} \cup \{x+1, \ldots, 2n\}$, which is indeed a rank-$(n-1)$ flat. This concludes the proof of Lemma 5.4. □

Having shown Lemma 5.4, the rest is easy. The rank-$(n-2)$ flat which is contained in the largest number of rank-$(n-1)$ flats, is the latest one to arrive to height -1. This flat is clearly $\{1, 2, \ldots, n-3, n-2, 2n\}$, which arrives to height -1 after $2n-3$ steps. It is contained in exactly n rank-$(n-1)$ flats.

Therefore $\mathbf{C}_n$ contains $U_{2,n}$ as a minor, and thus it is not representable over a field $\mathbb{F}_q$ with $q \leqslant n-2$. □

Acknowledgments

The present work is Chapter 5 of the author's Ph.D. thesis [1]. I am very grateful to my advisor, Richard Stanley, and to the Department of Mathematics at MIT for their support. I also thank Nantel Bergeron, Joseph Bonin, James Haglund, Carly Klivans, and the referees for helpful discussions and correspondence.

References

[1] F. Ardila, Enumerative and algebraic aspects of matroids and hyperplane arrangements, Ph.D. Dissertation, Massachusetts Institute of Technology, 2003.
[2] J.E. Bonin, A. de Mier, Lattice path matroids: structural aspects, in preparation.
[3] J.E. Bonin, A. de Mier, M. Noy, Lattice path matroids: enumerative aspects and Tutte polynomials, arXiv:math.CO/0211188 v1, 12 November 2002.
[4] H.H. Crapo, The Tutte polynomial, Aequationes Math. 3 (1969) 211–229.
[5] W. Fulton, Young Tableaux with Applications to Representation Theory and Geometry, Cambridge University Press, New York, 1997.
[6] J. Haglund, personal communication, 2002.
[7] G. Kalai, Algebraic shifting, in: T. Hibi (Ed.), Computational Commutative Algebra and Combinatorics, Mathematical Society of Japan, Tokyo, 2002.
[8] C. Klivans, Shifted matroid complexes, Ph.D. Dissertation, Massachusetts Institute of Technology, in preparation.
[9] G. Kreweras, Sur les éventails de segments, Cahiers BURO 15 (1970) 3–41.
[10] J.G. Oxley, Matroid Theory, Oxford University Press, New York, 1992.
[11] J.G. Oxley, K. Prendergast, D. Row, Matroids whose ground sets are domains of functions, J. Austral. Math. Soc. Ser. A 32 (1982) 380–387.
[12] M.J. Piff, D.J.A. Welsh, On the vector representation of matroids, J. London Math. Soc. 2 (1970) 284–288.
[13] R.P. Stanley, Enumerative Combinatorics, Vol. 1, Wadsworth and Brooks – Cole, Belmont, CA, 1986; reprinted by Cambridge University Press, Cambridge, 1997.
[14] R.P. Stanley, Enumerative Combinatorics, Vol. 2, Cambridge University Press, Cambridge, 1999.
[15] R.P. Stanley, personal communication, 2002.
[16] J. Vallé, Une bijection explicative de plusieurs propriétés ramarquables des ponts, European J. Combin. 18 (1997) 117–124.
[17] D.J.A. Welsh, A bound for the number of matroids, J. Combin. Theory 6 (1969) 313–316.
[18] D.J.A. Welsh, Matroid Theory, Academic Press, New York, 1976.