The Canada Day Theorem

Daniel Gomez*
Department of Mathematics and Statistics
University of Saskatchewan
106 Wiggins Road
Saskatoon SK S7N 5E6, Canada
dag857@mail.usask.ca

Hans Lundmark†
Department of Mathematics
Linköping University
SE-581 83 Linköping, Sweden
hans.lundmark@liu.se

Jacek Szmigielski‡
Department of Mathematics and Statistics
University of Saskatchewan
106 Wiggins Road
Saskatoon SK S7N 5E6, Canada
szmigiel@math.usask.ca

Submitted: Jun 19, 2012; Accepted: Jan 7, 2013; Published: Jan 29, 2013
Mathematics Subject Classifications: 15A15

Abstract

The Canada Day Theorem is an identity involving sums of $k \times k$ minors of an arbitrary $n \times n$ symmetric matrix. It was discovered as a by-product of the work on so-called peakon solutions of an integrable nonlinear partial differential equation proposed by V. Novikov. Here we present another proof of this theorem, which explains the underlying mechanism in terms of the orbits of a certain abelian group action on the set of all k-edge matchings of the complete bipartite graph $K_{n,n}$.

1 Introduction

The "Canada Day Theorem" refers to the following curious combinatorial fact:

Theorem 1.1. *Let T be the $n \times n$ matrix with entries*

$$T_{ij} = 1 + \operatorname{sgn}(i-j) = \begin{cases} 0, & i < j, \\ 1, & i = j, \\ 2, & i > j, \end{cases} \tag{1.1}$$

*Supported by grant NSERC (Natural Sciences and Engineering Research Council of Canada) USRA 414987.
†Supported by grant VR 2010-5822 from the Swedish Research Council (Vetenskapsrådet).
‡Supported by grant NSERC 163953.

and let X be an arbitrary symmetric $n \times n$ matrix. Then, for $1 \leqslant k \leqslant n$, the sum of the $k \times k$ principal minors of TX equals the sum of all $k \times k$ minors of X (principal and non-principal).

This theorem arose as an unexpected byproduct in our previous paper [8] where we studied so-called "peakon" (peaked soliton) solutions to a completely integrable nonlinear partial differential equation discovered by V. Novikov [12]; a brief account of this can be found in the appendix. The name of the theorem refers to its "birthday" (July 1, Canada's national holiday, in 2008).

A $k \times k$ minor in an $n \times n$ matrix $X = (x_{ij})$ is determined by a choice of k-element index sets I and J (row and column indices, respectively) from the set $[n] = \{1, 2, \ldots, n\}$. We will use the notation $\binom{[n]}{k}$ for the set of all k-element subsets of $[n]$, and write minors as

$$|X_{IJ}| = |X_{i_1 \ldots i_k, j_1 \ldots j_k}| = \det \begin{pmatrix} x_{i_1 j_1} & \cdots & x_{i_1 j_k} \\ \vdots & & \vdots \\ x_{i_k j_1} & \cdots & x_{i_k j_k} \end{pmatrix}, \tag{1.2}$$

for $I, J \in \binom{[n]}{k}$, where $i_1 < i_2 < \cdots < i_k$ and $j_1 < j_2 < \cdots < j_k$ are the elements of I and J listed in increasing order. A *principal minor* is a minor with $I = J$.

The sum of the $k \times k$ principal minors of a matrix A is of course very familiar as the coefficient of $(-1)^k z^{n-k}$ in the characteristic polynomial $\det(zI - A)$. The sum of *all* $k \times k$ minors is much less frequently encountered; one rare example in the literature is [13], but the results there do not seem directly related to Theorem 1.1, since they deal with arbitrary matrices where symmetry plays no role.

Example 1.2. In the case $n = 3$ we have

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & 2 & 1 \end{pmatrix}, \qquad X = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix},$$

and hence

$$TX = \begin{pmatrix} a & b & c \\ 2a+b & 2b+d & 2c+e \\ 2a+2b+c & 2b+2d+e & 2c+2e+f \end{pmatrix}.$$

The sum of the principal 1×1 minors of TX is just the trace, $(a + 2b + 2c) + (d + 2e) + f$, which indeed equals the sum of all entries of X (i.e., the sum of all 1×1 minors). The sum of the principal 2×2 minors of TX is

$$\begin{vmatrix} a & b \\ 2a+b & 2b+d \end{vmatrix} + \begin{vmatrix} a & c \\ 2a+2b+c & 2c+2e+f \end{vmatrix} + \begin{vmatrix} 2b+d & 2c+e \\ 2b+2d+e & 2c+2e+f \end{vmatrix},$$

which, as can be easily verified, equals the sum of all 2×2 minors of X,

$$\begin{vmatrix} a & b \\ b & d \end{vmatrix} + \begin{vmatrix} a & c \\ c & f \end{vmatrix} + \begin{vmatrix} d & e \\ e & f \end{vmatrix} + \begin{vmatrix} a & c \\ b & e \end{vmatrix} + \begin{vmatrix} b & d \\ c & e \end{vmatrix} + \begin{vmatrix} b & e \\ c & f \end{vmatrix} + \begin{vmatrix} a & b \\ c & e \end{vmatrix} + \begin{vmatrix} b & c \\ d & e \end{vmatrix} + \begin{vmatrix} b & c \\ e & f \end{vmatrix}.$$

And, since $\det T = 1$, the single 3×3 minor of TX (which is of course principal) equals the single 3×3 minor of X: $\det(TX) = \det X$.

Example 1.3. For $n > 3$, the cases $k = 1$ and $k = n$ of the theorem are still easy to verify: for $k = 1$ we have

$$\operatorname{tr}(TX) = \tfrac{1}{2}\operatorname{tr}(TX) + \tfrac{1}{2}\operatorname{tr}((TX)^t) = \operatorname{tr}\Big(\tfrac{1}{2}(T + T^t)X\Big) = \sum_{i,j} x_{ij}$$

since $\frac{1}{2}(T + T^t)$ is the matrix with 1 in every position, and for $k = n$ we have $\det(TX) = \det X$ as before. However, the intermediate cases $2 \leqslant k \leqslant n - 1$ are more involved. (The reader might want to check the case $n = 4$, $k = 2$ to become convinced of this.)

Of particular importance for the Canada Day Theorem are minors whose row and column indices are *interlacing*, meaning that

$$i_1 \leqslant j_1 \leqslant i_2 \leqslant j_2 \leqslant \cdots \leqslant i_k \leqslant j_k. \tag{1.3}$$

We abbreviate this condition as $I \leqslant J$. Given sets $I, J \in \binom{[n]}{k}$, let $I' = I \setminus (I \cap J)$ and $J' = J \setminus (I \cap J)$, and denote the cardinality of I' and J' by

$$p = p(I, J) = \operatorname{card}(I') = \operatorname{card}(J') = k - \operatorname{card}(I \cap J). \tag{1.4}$$

Note that I and J are interlacing if and only if I' and J' are *strictly interlacing* (abbreviated $I' < J'$),

$$i'_1 < j'_1 < i'_2 < j'_2 < \cdots < i'_p < j'_p. \tag{1.5}$$

With this notation in place, we can state the more precise version of the Canada Day Theorem that we are actually going to prove; it says that the two sums are both equal to a third sum involving only minors with interlacing index sets:

Theorem 1.4. *Let T and X be as in Theorem 1.1, and let*

$$S = \sum_{\substack{I,J \in \binom{[n]}{k} \\ I \leqslant J}} 2^{p(I,J)} \left| X_{IJ} \right|. \tag{1.6}$$

Then the following holds:

(a) The sum of the principal $k \times k$ minors of TX equals S.

(b) The sum of all $k \times k$ minors of X equals S.

The following is a short outline of the paper. In Section 2 we use the Cauchy–Binet formula and the Lindström–Gessel–Viennot path-counting lemma to prove the easier part of Theorem 1.4, namely part (a), which is true regardless of whether X is symmetric or not. Part (b) is proved in Section 3, and here it is crucial that X is a symmetric matrix. We introduce a group of "flips" and study an action of it on the set of all k-edge

matchings of the complete bipartite graph $K_{n,n}$. The main technical point is dealt with in Lemma 3.8, and this leads to the characterization of the orbit structure of the group action presented in Lemma 3.9, Corollary 3.10 and Corollary 3.11. From these results it then follows that when expanding each $k \times k$ minor according to the definition of the determinant and adding everything up, terms corresponding to orbits of a certain type ("non-interlacing") cancel out, while the other orbits ("interlacing") give contributions adding up to the sum S.

It is fair to say that the proof given in the present paper is not entirely different from the one in [8]. The central concept of (open) clusters introduced in Section 3 has its counterpart, namely linked pairs, in the original proof. However, the organization of the proof, in particular the identification of the group action on the set of all k-edge matchings as the main underpinning for the Canada Day Theorem, is novel.

2 The sum of the principal $k \times k$ minors of TX

In this section we prove part (a) of Theorem 1.4. Recall the Cauchy–Binet formula for the minors of a matrix product [6, Ch. I, § 2]:

$$|(TX)_{AB}| = \sum_{I \in \binom{[1,n]}{k}} |T_{AI}|\,|X_{IB}|\,, \qquad \text{for } A, B \in \tbinom{[1,n]}{k}. \tag{2.1}$$

Applying this with $A = B = J$ we can rewrite the sum of the principal $k \times k$ minors of TX as

$$\sum_{J \in \binom{[n]}{k}} |(TX)_{JJ}| = \sum_{I,J \in \binom{[n]}{k}} |T_{JI}|\,|X_{IJ}|\,. \tag{2.2}$$

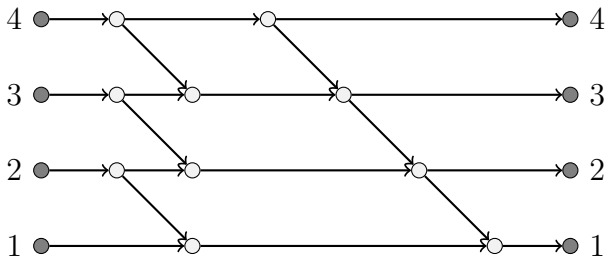
Next, we need to compute the minors $|T_{JI}|$. Here the notion of interlacing index sets (as defined in the Introduction) enters.

Lemma 2.1. *Let T be defined as in Theorem 1.1. Then, for $I, J \in \binom{[1,n]}{k}$,*

$$|T_{JI}| = \begin{cases} 2^{p(I,J)}, & \text{if } I \leqslant J, \\ 0, & \text{otherwise}, \end{cases} \tag{2.3}$$

where $p(I,J) = k - \operatorname{card}(I \cap J)$.

Proof. The matrix T is the path matrix of a planar directed graph of the form illustrated below in the case $n = 4$ (i.e., the entry T_{ab} counts the number of paths – zero, one, or two – from source vertex number a on the left to sink vertex number b on the right):



By the Lindström–Gessel–Viennot lemma [10, 11, 7, 1], the minor $|T_{JI}|$ equals the number of vertex-disjoint path families from sources indexed by J to sinks indexed by I.

Consider a vertex-disjoint path family from the source set $J = \{1 \leqslant j_1 < j_2 < \cdots < j_k \leqslant n\}$ to the sink set $I = \{1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n\}$. The planarity of of the graph allows only paths $j_m \to i_m$. Clearly $i_m \leqslant j_m$, since no path can go upwards, and moreover $j_m \leqslant i_{m+1}$, since otherwise both the path $j_m \to i_m$ and the path $j_{m+1} \to i_{m+1}$ would have to pass through the vertex immediately to the left of sink vertex i_{m+1}. Hence, if the interlacing condition $I \leqslant J$ is not satisfied, then there are no vertex-disjoint path families from J to I, and therefore $|T_{JI}| = 0$ in this case.

Suppose now that $I \leqslant J$. For each $m = 1, \ldots, k$ we draw a rectangular window over the graph, with its upper left corner at source vertex j_m and its lower right corner at sink vertex i_m. Imagine trying to construct a vertex-disjoint path family from J to I. There will be only one possible path $j_m \to i_m$ if the mth window has its height $j_m - i_m$ equal to zero, or if it shares its top edge with another window; otherwise there will be two possible paths $j_m \to i_m$. Indeed, when the height is zero there is only one way to start and finish on the same level. When the top edge is shared, there is again only one path, since the path corresponding to the window directly above must use the vertex immediately to the left of sink vertex i_{m+1} (which is on the same level as source vertex j_m). Therefore, if we remove all windows which are of height zero or share their top edge with another window, then the number of possible vertex-disjoint path systems will be given by 2 to the power of the number of remaining windows, in other words $2^{\mathrm{card}(I')}$ where $I' = I \setminus (I \cap J)$. $\square$

Now, inserting the value of $|T_{JI}|$ from this lemma into equation (2.2), we obtain

$$\sum_{J \in \binom{[n]}{k}} |(TX)_{JJ}| = \sum_{\substack{I,J \in \binom{[n]}{k} \\ I \leqslant J}} 2^{p(I,J)} \, |X_{IJ}| = S, \tag{2.4}$$

which finishes the proof of part (a) of Theorem 1.4.

Remark 2.2. There are several ways to compute the minors $|T_{JI}|$; see [8] for another argument using induction on n. Note that the proof above is implicitly taking advantage of the factorization

$$T = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 1 & \cdots & 1 & 1 & 1 \end{pmatrix}. \tag{2.5}$$

3 The sum of all $k \times k$ minors of X

In this section we prove part (b) of Theorem 1.4.

3.1 Minors and matchings

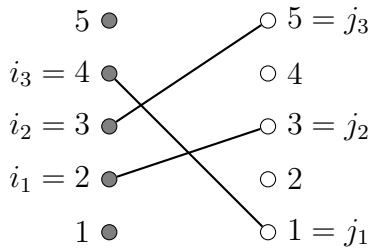
By the definition of the determinant, we have

$$|X_{IJ}| = \det \begin{pmatrix} x_{i_1 j_1} & \dots & x_{i_1 j_k} \\ \vdots & & \vdots \\ x_{i_k j_1} & \dots & x_{i_k j_k} \end{pmatrix} = \sum_{\pi \in S_k} \operatorname{sgn}(\pi) x_{i_1 j_{\pi(1)}} \cdots x_{i_k j_{\pi(k)}}, \tag{3.1}$$

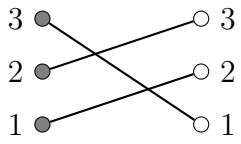
where S_k is the symmetric group of permutations on k elements (i.e., bijections $\pi\colon [k] \to [k]$). A less index-heavy notation is obtained by using that (for given I and J) each permutation $\pi\colon [k] \to [k]$ corresponds to a unique bijection $\tau\colon I \to J$ via $\tau(i_k) = j_{\pi(k)}$. Setting $\operatorname{sgn}(\tau) = \operatorname{sgn}(\pi)$ and $X_{(\tau)} = \prod_{i \in I} x_{i\tau(i)}$, we simply get

$$|X_{IJ}| = \sum_{\tau\colon I \to J} \operatorname{sgn}(\tau) X_{(\tau)} \tag{3.2}$$

where the sum runs over all bijections from I to J. We visualize such a bijection τ as a bipartite graph; for example, with $n = 5$, the bijection $\tau\colon \{2,3,4\} \to \{1,3,5\}$ given by $\tau(2) = 3$, $\tau(3) = 5$, $\tau(4) = 1$ is drawn as



and it corresponds (for the given sets I and J) to the permutation $\pi \in S_3$ represented by the graph

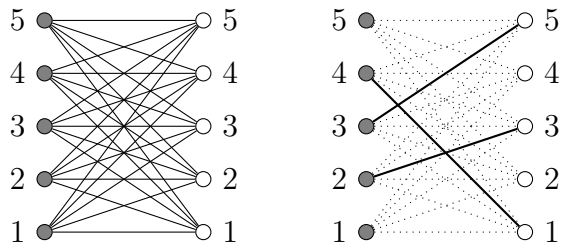


The sign of τ (and π) is $+1$ or -1 depending on whether the crossing number of the graph is even or odd. In this example there are two crossings, so $\operatorname{sgn}(\tau) X_{(\tau)} = +x_{23}x_{35}x_{41}$ (and this is one of the six terms in the 3×3 minor $|X_{234,135}|$).

Note that when composing two bijections $\tau_1\colon I \to L$ and $\tau_2\colon L \to J$, the signs obey the same rule as for the corresponding permutations:

$$\operatorname{sgn}(\tau_2 \circ \tau_1) = \operatorname{sgn}(\tau_2) \operatorname{sgn}(\tau_1).$$

Choosing k-element sets I and J together with a bijection $\tau\colon I \to J$ is equivalent to choosing a k-edge matching of the complete bipartite graph $K_{n,n}$:



(Recall that a *matching* of a graph $X = (V, E)$ with vertex set V and edge set E is a subset $F \subseteq E$ such that no two edges in F share a common vertex. One may equivalently think of the matching as being the subgraph (V, F). This habit of not distinguishing a graph from its edge set, when the underlying vertex set is understood, will be used quite frequently.)

Fix n and k, and let $\mathcal{M} = \mathcal{M}_{n,k}$ denote the set of k-edge matchings of $K_{n,n}$. We will use the same symbol τ both for such a matching and for the corresponding bijection, and slightly abuse the language by speaking about matchings $\tau\colon I \to J$. If $\tau(i) = j$ holds for the bijection τ, then we say that the matching τ contains an edge $i \to j$. (The graph $K_{n,n}$ is undirected, but the arrow notation is convenient for distinguishing the left nodes labelled $1, \ldots, n$ from the right nodes also labelled $1, \ldots, n$.)

When summing all $k \times k$ minors of an $n \times n$ matrix X and expanding each minor, the whole sum turns into an alternating sum over all k-edge matchings of $K_{n,n}$:

$$\sum_{I,J \in \binom{[n]}{k}} |X_{IJ}| = \sum_{I,J \in \binom{[n]}{k}} \sum_{\tau\colon I \to J} \operatorname{sgn}(\tau) X_{(\tau)} = \sum_{\tau \in \mathcal{M}} \operatorname{sgn}(\tau) X_{(\tau)}. \tag{3.3}$$

In order to prove that this equals

$$\sum_{I \leqslant J} 2^{p(I,J)} \, |X_{IJ}| = \sum_{I \leqslant J} \sum_{\tau\colon I \to J} 2^{p(I,J)} \operatorname{sgn}(\tau) X_{(\tau)}$$

when X is symmetric, as claimed in Theorem 1.4, we will introduce a certain abelian group $\mathcal{G}$ which acts on $\mathcal{M}$ in a way which preserves the weights $X_{(\tau)}$ but may change the signs $\operatorname{sgn}(\tau)$. Then we compute the sum $\sum_{\tau \in \mathcal{M}}$ by adding the terms for each group orbit separately and then summing over all orbits. As we will see, each orbit containing a matching $\tau\colon I \to J$ with interlacing index sets $I \leqslant J$ contributes $2^{p(I,J)}$ terms which all have the same sign, while the other orbits contain equally many positive and negative terms and therefore cancel out.

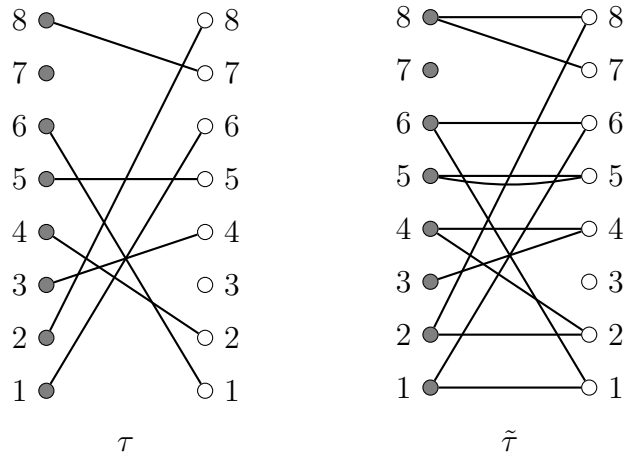
3.2 Clusters of a matching

Fix a matching $\tau\colon I \to J$ (viewed as a bipartite graph). Temporarily add auxiliary horizontal edges $r \to r$ for all $r \in I \cap J$. Split the resulting (multi)graph $\tilde{\tau}$ into connected components, and then remove the auxiliary edges. The remnants of the components of $\tilde{\tau}$ form a partition of the edges of τ into what we will call *clusters*.

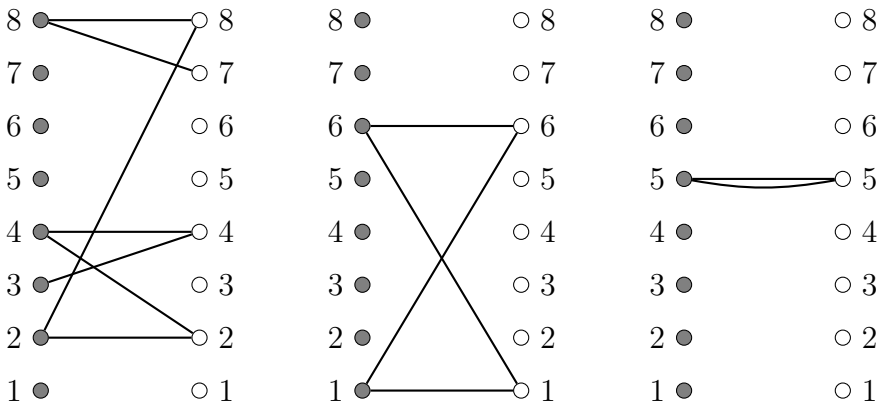
Since τ is a matching, no vertex in $\tilde{\tau}$ can have degree greater than two, and therefore the components of $\tilde{\tau}$ are paths. A cluster is either called *closed* or *open*, depending on

whether the corresponding component of $\tilde{\tau}$ is a closed or an open path. The endpoints of an open path will also be said to be the endpoints of that open cluster.

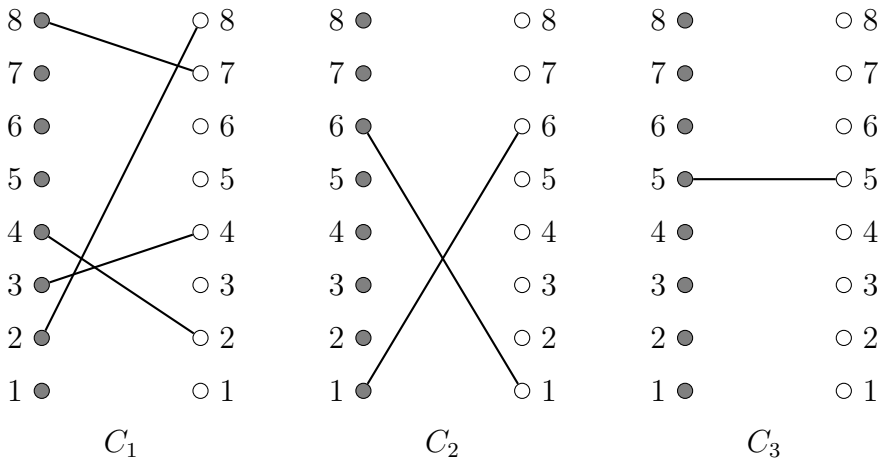
Example 3.1. Let $n = 8$ and $k = 7$. Here is a matching $\tau\colon I \to J$, where $I = \{1, 2, 3, 4, 5, 6, 8\}$ and $J = \{1, 2, 4, 6, 7, 8\}$, together with its companion $\tilde{\tau}$ obtained by adding auxiliary horizontal edges $r \to r$ for $r = I \cap J = \{1, 2, 4, 5, 6, 8\}$:



There are three connected components in $\tilde{\tau}$ (one open path with endpoints $3 \in I$ and $7 \in J$, and two closed paths):



Removing the auxiliary edges, we obtain the three clusters of the matching τ (one open and two closed):



Definition 3.2 (Endpoint separation)**.** Let $\tau\colon I \to J$ be a matching and C one of its open clusters. Consider the list consisting of all numbers in $I \cup J$, sorted in ascending order. (If $i_m = j_{m'}$ for some m and m', then include that number twice in the list.) Let the *endpoint separation* $\mathrm{sep}(C)$ of the cluster C be the number of entries in this sorted list lying strictly between the numbers labelling the two endpoints of C. For closed clusters C, we set $\mathrm{sep}(C) = 0$.

Remark 3.3. Observe that the endpoint separation is a property of both the matching and the cluster, and not just of the cluster itself.

Example 3.4. The open cluster C_1 in Example 3.1, with endpoints labelled 3 and 7, has endpoint separation $\mathrm{sep}(C_1) = 6$, since there are six numbers (4, 4, 5, 5, 6, 6) strictly between the 3 and 7 in the list $(1, 1, 2, 2, 3, 4, 4, 5, 5, 6, 6, 7, 8, 8)$. For the closed clusters, we have $\mathrm{sep}(C_2) = \mathrm{sep}(C_3) = 0$ by definition.

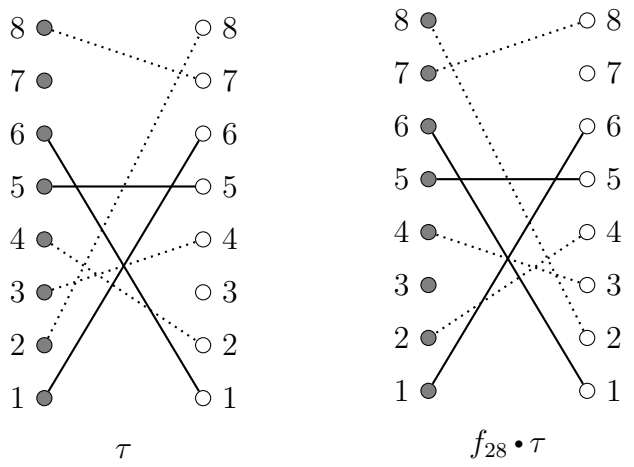
3.3 The group of flips

Given n, let $\mathcal{G} = \mathcal{G}_n$ be the abelian group obtained by taking the direct sum of $\binom{n}{2}$ copies of the two-element group $\mathbf{Z}/(2)$ (one copy for each pair (i, j) with $1 \leqslant i < j \leqslant n$). We define an action of $\mathcal{G}$ on the set $\mathcal{M}$ of k-edge matchings of $K_{n,n}$ as follows:

- Let $f_{ij} = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathcal{G}$, where the 1 is in the (i, j)th copy of $\mathbf{Z}/(2)$. The elements $\{f_{ij}\}_{i<j}$ generate $\mathcal{G}$, so it is enough to define how they act.
- If there is an open cluster C in the matching $\tau \in \mathcal{M}$ containing one of the edges $i \to j$ and $j \to i$ (note that an open cluster cannot contain both), then let $f_{ij} \bullet \tau$ be the matching obtained by *flipping* the whole cluster C, i.e., replacing each edge $a \to b$ in C by $b \to a$ (and leaving all other edges in τ as they are).
- Otherwise, let f_{ij} do nothing: $f_{ij} \bullet \tau = \tau$.

It is straightforward to verify that this really defines a group action as claimed, since flips commute and flipping the same cluster twice is the identity transformation. Note that it is the flipping of a whole cluster (rather than just an individual edge) that ensures that $f_{ij} \bullet \tau$ is still a matching.

Example 3.5. When f_{28} acts on the matching τ of Example 3.1, the open cluster C_1 (here drawn dotted) is flipped, since it is open (with endpoints 3 and 7) and contains the edge $2 \to 8$:



Lemma 3.6. *If the matrix X is symmetric, then $X_{(g \bullet \tau)} = X_{(\tau)}$ for all flips $g \in \mathcal{G}$ and all matchings $\tau \in \mathcal{M}$.*

Proof. This is immediate, since replacing the edge $a \to b$ by the edge $b \to a$ in the matching $\tau\colon I \to J$ corresponds to replacing the matrix entry X_{ab} by X_{ba} in the product $X_{(\tau)} = \prod_{i \in I} X_{i\tau(i)}$. □

Example 3.7. The flip in Example 3.5 corresponds to changing

$$X_{(\tau)} = X_{16}X_{28}X_{34}X_{42}X_{55}X_{61}X_{87}$$

into

$$X_{(f_{28} \bullet \tau)} = X_{16}X_{82}X_{43}X_{24}X_{55}X_{61}X_{78}.$$

A less trivial aspect of the group action is the relation between $\operatorname{sgn}(g \bullet \tau)$ and $\operatorname{sgn}(\tau)$, and this is where the endpoint separation and the interlacing condition $I \leqslant J$ will be of importance.

Lemma 3.8. *If the cluster C is flipped when f_{ij} acts on τ, then*

$$\operatorname{sgn}(f_{ij} \bullet \tau) = (-1)^{\operatorname{sep}(C)} \operatorname{sgn}(\tau). \tag{3.4}$$

Proof. As remarked in Section 3.1, the composition of matchings $\tau_1 \colon I \to L$ and $\tau_2 \colon L \to J$ (thought of as functions) satisfies $\operatorname{sgn}(\tau_2 \circ \tau_1) = \operatorname{sgn}(\tau_2)\operatorname{sgn}(\tau_1)$. We will split $\tau \colon I \to J$ into such a composition where τ_1 deals with the edges in the cluster and τ_2 with the remaining edges.

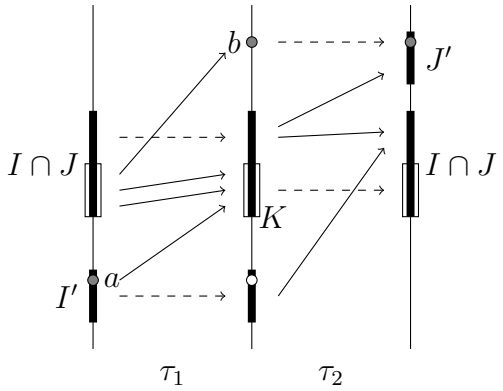
Consider an open cluster C in τ with endpoints $a \in I' = I \setminus (I \cap J)$ and $b \in J' = J \setminus (I \cap J)$:

$$\tau(a) = c_1, \quad \tau(c_1) = c_2, \quad \tau(c_2) = c_3, \quad \ldots, \quad \tau(c_{m-1}) = c_m, \quad \tau(c_m) = b.$$

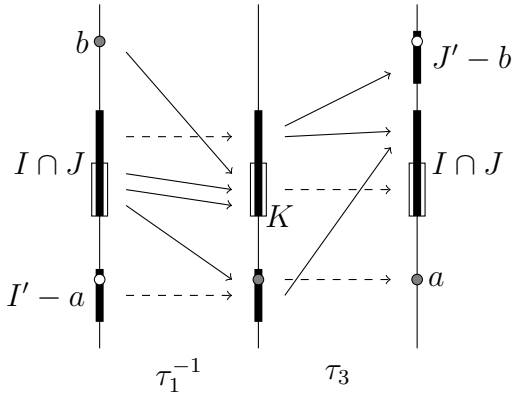
We let $K = \{c_1, c_2, \ldots, c_m\} \subset I \cap J$ and define $\tau_1 \colon I \to I + b - a$ and $\tau_2 \colon I + b - a \to J$ by

$$\tau_1(x) = \begin{cases} \tau(x), & \text{if } x \in K + a, \\ x, & \text{if } x \in I \setminus (K + a), \end{cases} \qquad \tau_2(x) = \begin{cases} x, & \text{if } x \in K + b, \\ \tau(x), & \text{if } x \in I \setminus (K + b). \end{cases} \tag{3.5}$$

(For readability, we have written $I + b - a$ instead of $(I \cup \{b\}) \setminus \{a\}$ and $K + a$ instead of $K \cup \{a\}$.) These definitions imply that $\tau = \tau_2 \circ \tau_1$. Schematically, with dashed arrows indicating the mapping $x \mapsto x$ and solid arrows indicating $x \mapsto \tau(x)$:



The result of flipping the cluster C is $f_{ij} \bullet \tau = \tau_3 \circ \tau_1^{-1}$, where $\tau_3 \colon I \to J + a - b$ differs from τ_2 in having the edge $a \to a$ instead of the edge $b \to b$:



We have $\operatorname{sgn}(\tau_3) = (-1)^{\operatorname{sep}(C)} \operatorname{sgn}(\tau_2)$, since if we imagine detaching the edge $b \to b$ from its vertices and continuously sliding it to the position $a \to a$, the crossing number of

the matching changes by one each time either the left or the right end of of that edge moves past a matched vertex, and by the definition of endpoint separation there are exactly $\mathrm{sep}(C)$ matched vertices on the levels strictly between a and b. Together with $\mathrm{sgn}(\tau_1^{-1}) = \mathrm{sgn}(\tau_1)$, this proves (3.4). □

Let us call a matching $\tau\colon I \to J$ interlacing if the sets I and J are interlacing, $I \leqslant J$ (cf. (1.3)). The orbits of the interlacing matchings (under the action of the flip group $\mathcal{G}$) will be called interlacing orbits, and all other orbits non-interlacing.

Lemma 3.9. *A matching σ belongs to an interlacing orbit if and only if every cluster C in σ has **even** endpoint separation* $\mathrm{sep}(C)$.

Proof. It is clear from the definitions that flipping a cluster does not change its endpoint separation. Closed clusters always have even endpoint separation by definition (namely, zero). If C is an open cluster in an interlacing matching $\tau\colon I \to J$, with endpoints $i_s \in I' = I \setminus (I \cap J)$ and $j_t \in J' = J \setminus (I \cap J)$, then in the case $i_s < j_t$ we find from the interlacing condition that

$$\cdots \leqslant j_{s-1} < i_s < \underbrace{j_s \leqslant \cdots \leqslant i_t}_{2(t-s)\text{ elements}} < j_t < i_{t+1} \leqslant \cdots,$$

so that $\mathrm{sep}(C) = 2(t-s)$ is even, and the case $i_s > j_t$ is similar. It follows that if σ belongs to the orbit of an interlacing matching, then every cluster in σ has even endpoint separation $\mathrm{sep}(C)$.

Conversely, given a matching $\sigma\colon A \to B$ all of whose clusters have even endpoint separation, we form interlacing sets $I \leqslant J$ by sorting the list of numbers $(a_1, \ldots, a_k, b_1, \ldots, b_k)$ in ascending order and labelling the elements of the sorted list as $(i_1, j_1, i_2, j_2, \ldots, i_k, j_k)$ (note that $I \cap J = A \cap B$). If $a \in A$ and $b \in B$ are the endpoints of an open cluster C in σ, then $(a, b) \in I \times J$ or $J \times I$ (rather than $I \times I$ or $J \times J$) because of the assumption that $\mathrm{sep}(C)$ is even. Flipping those open clusters in σ whose endpoints belong to $J \times I$ produces an interlacing matching $\tau\colon I \to J$ whose orbit σ belongs to. □

Corollary 3.10. *All matchings in a given interlacing orbit have the same sign.*

Proof. Combine Lemmas 3.8 and 3.9. □

Corollary 3.11. *In a non-interlacing orbit, there are equally many matchings of each sign.*

Proof. By Lemma 3.9, each matching in a non-interlacing orbit has at least one open cluster with odd endpoint separation; among those clusters we single out the unique one with the property that its lowest-numbered node is smaller than that of the other ones. The operation of flipping that cluster is a sign-reversing involution pairing up the matchings in the orbit. (It's sign-reversing by Lemma 3.8, and it's an involution since the lowest-numbered node will still be the lowest-numbered node after the flip.) □

Lemma 3.12. *(a) The orbit $\mathcal{G}\bullet\tau$ of a matching $\tau\colon I \to J$ contains $2^{p(I,J)}$ elements. (b) Each interlacing orbit contains exacly one interlacing matching.*

Proof. Recall that $p(I,J) = \mathrm{card}(I') = \mathrm{card}(J')$ where $I' = I \setminus (I \cap J)$ and $J' = J \setminus (I \cap J)$. Each element of I' is the endpoint of an open cluster whose other endpoint belongs to J' (by the definition of an open cluster), while all other clusters are closed. Thus there are exactly as many open clusters in τ as there are elements in I' and J', and each of these $p(I,J)$ open clusters can be flipped independently of the others. This proves part (a). To prove part (b) we consider two cases. If an interlacing matching τ has only closed clusters then by (a) $\mathrm{card}(\mathcal{G}\bullet\tau) = 1$, hence (b) holds. If, on the other hand, an interlacing matching τ has at least one open cluster then flipping any open cluster destroys the interlacing property. This follows from the fact that $I \cup J$ and $I \cap J$ are invariant under flipping, and that there is only one way of constructing interlacing sets with given union and intersection. □

Now we only have to put the pieces together:

Proof of part (b) of Theorem 1.4. Let $\mathcal{M}^* \subset \mathcal{M}$ denote the union of the interlacing orbits. Then, referring in brackets to the relevant Lemmas and Corollaries above, we have

$$\begin{aligned}
\sum_{I,J} |X_{IJ}| &= \sum_{I,J} \sum_{\tau\colon I\to J} \mathrm{sgn}(\tau) X_{(\tau)} && \text{(def. of determinant)}\\
&= \sum_{\tau\in\mathcal{M}} \mathrm{sgn}(\tau) X_{(\tau)} \\
&= \sum_{\tau\in\mathcal{M}^*} \mathrm{sgn}(\tau) X_{(\tau)} + \sum_{\tau\notin\mathcal{M}^*} \mathrm{sgn}(\tau) X_{(\tau)} \\
&= \sum_{I\leqslant J} \sum_{\tau\colon I\to J} \sum_{\sigma\in\mathcal{G}\bullet\tau} \mathrm{sgn}(\sigma) X_{(\sigma)} && \text{(def. of } \mathcal{M}^*\text{, 3.12b)}\\
&\quad + 0 && \text{(3.6, 3.11)}\\
&= \sum_{I\leqslant J} \sum_{\tau\colon I\to J} 2^{p(I,J)}\, \mathrm{sgn}(\tau) X_{(\tau)} && \text{(3.6, 3.10, 3.12a)}\\
&= \sum_{I\leqslant J} 2^{p(I,J)} \left(\sum_{\tau\colon I\to J} \mathrm{sgn}(\tau) X_{(\tau)} \right) \\
&= \sum_{I\leqslant J} 2^{p(I,J)}\, |X_{IJ}| && \text{(def. of determinant).}
\end{aligned}$$

□

A Appendix

The Canada Day Theorem first arose in the context of nonlinear partial differential equations in [8]. In this appendix we describe briefly the subject of that paper and how it resulted in the formulation of the theorem. The nonlinear partial differential equation

$$u_t - u_{xxt} + 4u^2 u_x = 3u u_x u_{xx} + u^2 u_{xxx}, \tag{A.1}$$

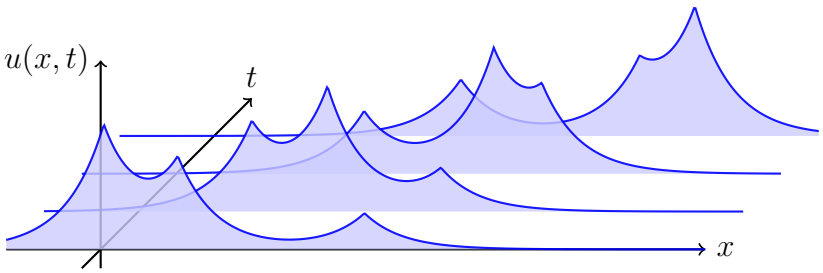


Figure 1: A three-peakon solution of the Camassa–Holm equation.

where $u = u(x,t)$, $u_x = \frac{\partial u}{\partial x}(x,t)$, $u_t = \frac{\partial u}{\partial t}(x,t)$, etc., was derived by V. Novikov [12] as part of a classification of generalized Camassa–Holm-type equations possessing infinite hierarchies of higher symmetries. The Camassa–Holm equation

$$u_t - u_{xxt} + 3uu_x = 2u_x u_{xx} + uu_{xxx}, \tag{A.2}$$

which was formulated by Camassa and Holm in [5] as an integrable shallow water equation (with u denoting a horizontal velocity component), admits a particularly interesting class of explicit weak solutions which capture phenomena such as the collision and breakdown of waves. Miraculously, many of these features turn out to be intrinsically connected to classical problems in analysis, like Stieltjes continued fractions [3], or the moment problem [4]. These weak explicit solutions are called *peakons* (short for *peaked solitons*) because of the characteristic $e^{-|x|}$ shape of the waves. Multipeakon solutions are formed by the superposition of n peakons,

$$u(x,t) = \sum_{i=1}^{n} m_i(t)\, e^{-|x - x_i(t)|}, \tag{A.3}$$

with a suitable time dependence in the amplitudes $m_i(t)$ and positions $x_i(t)$. Observe that since these solutions are not differentiable everywhere they only satisfy the PDE (A.2) in a certain weak sense. The evolution of a three-peakon solution over time is illustrated in Figure 1. One of the objectives of [8] was to find explicit formulas for the positions and amplitudes of the peaks for the multipeakon solution of Novikov's equation (A.1). After substituting the ansatz (A.3) into (A.1), taking into account the weak nature of the solutions, one obtains $2n$ ordinary differential equations which govern the time dependence of the positions and amplitudes:

$$\begin{aligned} \dot{x}_k &= \left(\sum_{i=1}^{n} m_i\, e^{-|x_k - x_i|} \right)^2, \\ \dot{m}_k &= m_k \left(\sum_{i=1}^{n} m_i\, e^{-|x_k - x_i|} \right) \left(\sum_{j=1}^{n} m_j\, \operatorname{sgn}(x_k - x_j)\, e^{-|x_k - x_j|} \right). \end{aligned} \tag{A.4}$$

(By definition, $\operatorname{sgn}(0) = 0$ here.) These equations were already stated in [9] where it was shown that they constitute a Hamiltonian system, and one of the main results of [8] was

that they are also Arnol′d–Liouville integrable [2], meaning that there exist n functionally independent and Poisson commuting constants of motion $H_1, \ldots, H_n$. To display these constants of motion we first define the matrices

$$\begin{aligned} P &= \operatorname{diag}(m_1, \ldots, m_n), \\ E &= (E_{ij})_{i,j=1}^n, \qquad \text{where } E_{ij} = e^{-|x_i - x_j|}, \\ T &= (T_{ij})_{i,j=1}^n, \qquad \text{where } T_{ij} = 1 + \operatorname{sgn}(i-j), \end{aligned} \tag{A.5}$$

where P and E depend on the variables appearing in (A.4). One can then show that for every complex λ, the polynomial $A(\lambda) = \det(I - \lambda TPEP)$ is a constant of motion for equations (A.4), which implies that the coefficient of λ^k in the polynomial $A(\lambda)$ is a constant of motion. It can also be shown that these coefficients (for $1 \leqslant k \leqslant n$) are in fact Poisson commuting and functionally independent, thereby providing the desired set of n constants of motion. By elementary linear algebra, the coefficient of λ^k can be computed (up to a sign) as the sum over all $k \times k$ principal minors of the matrix $TPEP$. However, before this result was found, direct computations for small values of n had indicated that the constants of motion ought to be given by the sums over all $k \times k$ minors of the symmetric matrix PEP. It was the attempt to reconcile these observations that led to the formulation of Theorem 1.1, and as a result, the constants of motion now have the following description:

Theorem A.1. *The Novikov peakon ODEs* (A.4) *admit n constants of motion $H_1, \ldots, H_n$, where H_k equals the sum of all $k \times k$ minors (principal and non-principal) of the $n \times n$ symmetric matrix $PEP = (m_i m_j e^{-|x_i - x_j|})_{i,j=1}^n$.*

As a final remark, let us mention that explicit expressions for the minors of PEP can be written down easily with the help of the Lindström–Gessel–Viennot lemma; see [8] for details.

References

[1] M. Aigner. *A course in enumeration*, volume 238 of *Graduate Texts in Mathematics*. Springer, Berlin, 2007.

[2] V. I. Arnol′d. *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1989. Translated from the Russian by K. Vogtmann and A. Weinstein.

[3] R. Beals, D. H. Sattinger, and J. Szmigielski. Multi-peakons and a theorem of Stieltjes. *Inverse Problems*, 15(1):L1–L4, 1999.

[4] R. Beals, D. H. Sattinger, and J. Szmigielski. Multipeakons and the classical moment problem. *Adv. in Math.*, 154:229–257, 2000.

[5] R. Camassa and D. D. Holm. An integrable shallow water equation with peaked solitons. *Phys. Rev. Lett.*, 71(11):1661–1664, 1993.

[6] F. R. Gantmacher. *The theory of matrices. Vol. 1.* AMS Chelsea Publishing, Providence, RI, 1998. Translated from the Russian by K. A. Hirsch. Reprint of the 1959 translation.

[7] I. Gessel and G. Viennot. Binomial determinants, paths, and hook length formulae. *Adv. in Math.*, 58(3):300–321, 1985.

[8] A. N. W. Hone, H. Lundmark, and J. Szmigielski. Explicit multipeakon solutions of Novikov's cubically nonlinear integrable Camassa–Holm type equation. *Dyn. Partial Differ. Equ.*, 6(3):253–289, 2009.

[9] A. N. W. Hone and J. P. Wang. Integrable peakon equations with cubic nonlinearity. *J. Phys. A*, 41(37):372002, 2008.

[10] S. Karlin and J. McGregor. Coincidence probabilities. *Pacific J. Math.*, 9:1141–1164, 1959.

[11] B. Lindström. On the vector representations of induced matroids. *Bull. London Math. Soc.*, 5:85–90, 1973.

[12] V. Novikov. Generalizations of the Camassa–Holm equation. *J. Phys. A*, 42(34):342002, 14, 2009.

[13] S. Okada. On the generating functions for certain classes of plane partitions. *J. Combin. Theory Ser. A*, 51(1):1–23, 1989.