Having Fun in the Trees

青山綠水, 白草紅葉黃花

with Xinmao Wang (王新茂), Ziqing Xiang (向子卿),

Zeying Xu (徐澤瀛) and Yinfeng Zhu (祝隱峰)

Yaokun Wu (吳耀琨) Shanghai Jiao Tong University (上海交大)

Algebraic Combinatorics Workshop at USTC (中國科大代數組合工作誉)

Plan (計劃)

Phase space (相空間)
 Permutation (置換)
 Primitivity (本原性)
 Phylogenetics (系統發生學)
 Preorder (准序)
 Partition (劃分)



Figure: http://www.
artpromotivate.
com/2013/05/
christine-keech-my-he
html

1. Phase space (相空間)

2. Phylogenetics (系統發生學)

Discrete dynamical system (離散動力係統)

Let *S* be a set and let \mathcal{F} be a family of maps from *S* to *S*.

Viewing the maps in \mathcal{F} as a set of time-evolution laws and *S* the set of possible states, the pair (S, \mathcal{F}) forms a discrete dynamical system, where the dynamics are given by iterating the maps in \mathcal{F} .

The phase space of the discrete dynamical system (S, \mathcal{F}) , denoted by $\mathcal{PS}_{S,\mathcal{F}}$ or simply $\mathcal{PS}_{\mathcal{F}}$, is the digraph with vertex set *S* and arc set $\{s \to f(s) : s \in S, f \in \mathcal{F}\}$.

When \mathcal{F} is a singleton set $\{f\}$, we call $\mathcal{PS}_{\mathcal{F}}$ the phase space of a single map f and often write it as \mathcal{PS}_f . The digraph \mathcal{PS}_f has constant out-degree 1 and so each weakly connected component of it is a cycle with a directed tree (known as its transient there) attached to each vertex in the cycle. The wiring diagram:



Let
$$f_1 = \neg x_2, f_2 = x_4 \lor (x_1 \land x_3), f_3 = x_4 \land x_2, f_4 = x_2 \lor x_3.$$

The phase space of *f*:



Figure: http://www.samsi.info/sites/default/files/ abdul_jarrah_122008.pdf 1.1 Permutation (置換)
 1.2 Primitivity (本原性)

Rooted labelled tree (有根標號樹)

- A rooted labelled *n*-vertex tree is a rooted tree *T* with *n* vertices together with a bijection ℓ from V(T) to Z_n.
- ▶ We say that this labelled tree $T = (T, \ell)$ has tree type *T*.

Trees and permutations (樹與置換)

We define two maps p_n and q_n from the set of all rooted labelled *n*-vertex trees to itself, which involve both one-line notation and cycle notation of permutations.

The map p_n (映射 p_n)



Multiplication from right to left (cycle notation): (2, 1)(1, 3)(6, 4)(5, 4)(4, 3) = (3, 5, 6, 4, 2, 1)

Change the labelling (one-line notation):

ℓ	3	5	6	4	2	1
\downarrow	\downarrow	\downarrow	\downarrow	\downarrow	\downarrow	↓
$p_6(\ell)$	3	4	5	6	1	2

The map q_n (映射 q_n)



Multiplication from right to left (cycle notation): (2, 1)(1, 3)(6, 4)(5, 4)(4, 3) = (3, 5, 6, 4, 2, 1)

Change the labelling (one-line notation):

ℓ	3	5	6	4	2	1
\downarrow	\downarrow	\downarrow	↓	\downarrow	\downarrow	\downarrow
$q_6(\ell)$	3	2	1	6	5	4

A quiz (小游戲)

The phase spaces of p_n and q_n are denoted \mathcal{P}_n and Q_n , respectively.

What is the shape of \mathcal{P}_n and Q_n ? Or, what is the dynamical behavior of p_n and q_n ?

Fixed points of p_n (p_n 的不動點)

Theorem (W., Xu, Zhu). The set of labelled *n*-vertex trees of star type coincides with the set of fixed points of p_n .



Figure: \mathbb{T}_4 is a fixed point of p_6 .

Fixed points of q_n (q_n 的不動點)

Theorem (W., Xu, Zhu). A rooted *n*-vertex tree *T* with i_T inner vertices admits exactly $\alpha(T)$ labellings ℓ such that (T, ℓ) is a fixed point of q_n , where $\alpha(T) = 2^{i_T}$ if the root of *T* is a leaf and the maximum degree of *T* is at most 3, and $\alpha(T) = 0$ otherwise.



Figure: Both \mathbb{T}_5 and \mathbb{T}_6 are fixed points of q_6 .



Figure: Typical weakly connected components of \mathcal{P}_6 .



Figure: Typical weakly connected components of Q_6 .

Theorem (W., Xu, Zhu). Each vertex of $\mathcal{P}_n(Q_n)$ is either a loop vertex or on a 2-cycle or has its unique out-neighbor in a 2-cycle.

Theorem (W., Xu, Zhu). For every rooted tree *T* with *n* vertices, the number of labellings ℓ such that (T, ℓ) is on a cycle of $\mathcal{P}_n(Q_n)$ is $n \prod_{v \in V(T)} \deg_T(v)!$.



Perturbation of the rule (規則微擾)

- ▶ In general, \mathcal{P}_n and Q_n are not isomorphic to each other.
- ► Let σ_n be the map from \mathbb{Z}_n to \mathbb{Z}_n that sends *i* to *i* + 1. Instead of using the map p_n and q_n , we can use $\sigma_n^k \circ p_n$ and $\sigma_n^k \circ q_n$ for some fixed *k* and get new dynamical systems. It is observed that cycles of various lengths can happen in the phase spaces of these more general dynamical systems.

1.1 Permutation (置換)
 1.2 Primitivity (本原性)

Matrix as a map on nonempty subsets (矩陣與映射)

Let *k* be a positive integer and let Set_k denote $2^{[k]} \setminus \{\emptyset\}$.

A map f from Set_k to Set_k is essential provided

•
$$f(A) \cup f(B) = f(A \cup B)$$
, and

▶
$$f([k]) = [k].$$

The digraph of f, denoted Γ_f , is the digraph with vertex set [k] such that $i \to j$ is an arc of Γ_f if and only if $j \in f(i)$.

An essential map from Set_k to Set_k is the combinatorial counterpart of a k by k matrix without zero lines. Indeed, such a map f can be thought of as any k by k matrix M whose ith column has f(i) as its support for all $i \in [k]$.

Primitive matrix set (本原矩陣族)

Let \mathcal{F} be a family of essential maps on Set_k. Let the primitive index of \mathcal{F} , which we denote by $g(\mathcal{F})$, be the longest possible length of a walk in $\mathcal{PS}_{\mathcal{F}}$ without using the arc $[k] \rightarrow [k]$.

We say that \mathcal{F} is primitive provided $g(\mathcal{F})$ is finite, namely whenever the only cycle in $\mathcal{PS}_{\mathcal{F}}$ is the loop at [k].

It is clear that \mathcal{F} is primitive if and only if $g(\mathcal{F}) \leq 2^k - 2$ and if and only if $\mathcal{PS}_{\mathcal{F}}$ is acyclic after deleting the loop edge at [*k*].

Wielandt-type matrices (維蘭特型矩陣)

Take a positive integer $k \ge 2$ and choose $i \in [k - 1]$ satisfying gcd(i, k) = 1. A Wielandt-type matrix $W_{k;i}$ is the essential map (matrix) *A* from Set_k to Set_k such that $A(1) = \{2\}, ..., A(k - 1) = \{k\}, A(k) = \{1, 1 + i\}.$

$$W_{4;1} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, W_{4;3} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}, W_{5;4} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

$\mathcal{PS}_{W_{4;1}}$



 $g(W_{4;1}) = 10 = (4-1)^2 + 1.$

Wielandt (1959) shows that the primitive index of a primitive (0,1) matrix of order *k* is at most $(k - 1)^2 + 1$ and the bound is attained if and only if the matrix is permutation similar to $W_{k;1}$.

$\mathcal{PS}_{W_{4;3}}$



 $g(W_{4;3}) = 6 = 2 \times 4 - 2$

$\mathcal{PS}_{W_{5;4}}$



 $g(W_{5:4}) = 8 = 2 \times 5 - 2$

Extremal behavior of Wielandt-type matrices (極值表現)

Lemma (Wang, W., Xiang). Let $k \ge 2$ and A be a $k \times k$ primitive (0,1) matrix. Then \mathcal{PS}_A has at least 2^{k-2} vertices of in-degree zero, where the equality holds if and only if A is permutation similar to a Wielandt-type matrix.

Primitive index of matrix set of fixed size and order

Let $g_{k,t}$ be the maximum possible value of $g(\mathcal{A})$ where \mathcal{A} is a primitive matrix set consisting of t(0, 1) matrices of order k.

Besides Wielandt's bound of $g_{k,1} = (k-1)^2 + 1$ and the trivial bound of $g_{k,t} \le 2^k - 2$, very little is known about $g_{k,t}$.

$$\begin{cases} g_{2,2} = 2 = 2^2 - 2 = (2 - 1)^2 + 1\\ g_{3,2} = 6 = 2^3 - 2\\ g_{4,2} = 12 = 2^4 - 2^2\\ g_{5,2} \ge 23 < 24 = 2^5 - 2^3\\ g_{6,2} \ge 39 < 48 = 2^6 - 2^4 \end{cases}$$

Cohen-Sellers (1982) suggest to estimate the parameter γ_k , where $\gamma_k = \min\{t : g_{k,t} = 2^k - 2\}$.

Theorem (Wang, W., Xiang). It holds for all positive integers k that $\gamma_k \leq 1 + \binom{k-2}{\lfloor (k-2)/2 \rfloor}$.

Is the inequality in the theorem indeed an equality?

▶
$$\gamma_1 = \gamma_2 = 1, \gamma_3 = 2, \gamma_4 = 3, \gamma_5 \in \{3, 4\}.$$

- ► The proof of the theorem makes very heavy use of the structure analysis of the phase space of W_{k;k-1} as well as the chain decomposition of the Boolean algebra.
- Is there good understanding of the phase space of the general Wielandt-type matrices?

A pair of chain decompositions of $2^{[k]}$



Period (周期)

Let *A* be an essential map on Set_k such that Γ_A is strongly connected. For every $S \subseteq [k]$, we say that a positive integer *i* is a period of *A* at *S* provided $A^i(S) \supseteq S$ and we write Per_A(*S*) for the least period of *A* at *S*.



and

$$25 \rightarrow 3 \rightarrow 4 \rightarrow 15 \rightarrow 23 \rightarrow 34 \rightarrow 145 \rightarrow 1235$$

is a walk of length 7 in \mathcal{PS}_A .

Let $Prim_k$ be the set of all primitive essential maps on Set_k .

The *k*th primitive essential map complex is the simplicial complex on the ground set Prim_k such that $\mathcal{F} \subseteq \operatorname{Prim}_k$ is a face of the complex if and only if \mathcal{F} is primitive.

Can we see/smell a linear map? (蘭有秀兮菊有芳)



Figure: http://media.sjtu.edu.cn/photo!list.do?cid=23

- ► Take an essential map *A* on Set_k such that Γ_A is strongly connected. What is the shape of \mathcal{PS}_A ? Especially, how to get an upper bound for $\max_{S \in Set_k} \operatorname{Per}_A(S)$?
- Is the primitive matrix set complex a pure simplicial complex? Namely, is it true that all maximal faces of it are of the same dimension?

1. Phase space (相空間)

2. Phylogenetics (系統發生學)

An evolutionary tree of life (生命演化樹)



Figure: https://en.wikipedia.org/wiki/Phylogenetics

2.1 Preorder (准序) 2.2 Partition (劃分)



The leaf set of the tree *T* is $X = \{a, b, c, d\}$. The tree *T* induces the tree metric $D_T \in \mathbb{Z}^{X \times X}$ given by $D_T(a, a) = D_T(b, b) = D_T(c, c) = D_T(d, d) = 0, D_T(a, b) = D_T(c, d) =$ $2, D_T(a, c) = D_T(a, d) = D_T(b, c) = D_T(b, d) = 3.$

It also defines a total preorder p_T on $\binom{X}{2}$:

$$ab = cd < ac = ad = bc = bd$$
,

which comes from

$$D_T(a,b) = D_T(c,d) < D_T(a,c) = D_T(a,d) = D_T(b,c) = D_T(b,d).$$

We call p_T a tree preorder.

- The tree metrics have the famous 4-point condition characterization.
- Is there any characterization for the tree preorders?

Are there two different trees sharing the same leaf set and the same preorder?

Are there two different trees sharing the same leaf set and the same preorder?

Of course, subdividing edges appropriately will give us many such examples.



 $\mathbf{p}_T = \mathbf{p}_{T'}: \ ab = cd < ac = ad = bc = bd$

To see a world in a grain of sand (一花一世界 一葉一如來)

Conjecture (W., Xiang, Xu). Let *T* and *T'* be two trees with the same leaf set and without degree two vertices. If $p_T = p_{T'}$, then there is a graph isomorphism from *T* to *T'* which fixes every leaf vertex of the tree.

Since the tree metric D_T uniquely determines T, the above conjecture basically says that we can uniquely reconstruct the tree metric from the tree preorder provided the tree has no degree two vertices.

A regular tree is a tree all of whose inner vertices have the same degree and that degree is at least 3.

Theorem (W., Xiang, Xu). Let *T* and *T'* be two regular trees sharing the same leaf set. If $p_T = p_{T'}$, then there is a graph isomorphism from *T* to *T'* which fixes every leaf vertex.

Tell a tree from its distance set (從距離集看樹)

- ► The distance set of a tree *T*, denoted by DS_T, is the set of numbers which appear as distances on *T* between pairs of leaves.
- ► A set *I* of positive integers is called an avoidable tree distance set if for every number *k* there exists a tree *T* without degree 2 vertices such that its diameter is greater than *k* and $I \cap DS_T = \emptyset$.
- ▶ We say that a positive integer *k* is a distance jump of a tree *T* provided there exists $j \in DS_T$ such that

 $k = \min\{i : i > 0, i + j \in DS_T\}.$

Symbolic dynamics (符號動力系統)

Theorem (W., Xu, Zhu). There is an algorithm to decide, for any finite set *I* of positive integers, whether or not *I* is an avoidable tree distance set.

Our algorithm is to transform the problem to that of deciding whether or not a shift of finite type is an empty shift space.

A tree without leaf distance 6 (葉間距不出現六的樹)



Figure: $DS_T = \{2, 3, 4, 5, 7\}$

Some low-hanging fruits (舉手之勞)

Theorem (W., Xu, Zhu).

- A positive integer is an unavoidable tree distance if and only if it is one of 2, 4 and 6.
- A tree without degree 2 vertices and with diameter at least 6 can miss leaf distance 6 if and only if it is obtained from the tree in the previous slides by adding leaves to those vertices which is already adjacent to a leaf.
- The set {2k − 1, 2k} is an unavoidable tree distance set if and only if k ≤ 6.

▶

Let *O* be the set of positive odd integers. An even tree is a tree such that $DS_T \cap O = \emptyset$. It is easy to construct even trees without degree 2 vertices and with arbitrarily large distance jumps.

For any given positive number k, is there always a set I of consecutive positive integers of size bigger than k such that $I \cup O$ is avoidable?

2.1 Preorder (准序)
 2.2 Partition (劃分)

Edge boundary partition (邊集界線劃分)

For any graph *G* and any $U \subseteq V(G)$, let

 $E_G(U) = \{uv \in E(G) : u \in U, v \notin U\}$

denote the edge boundary of U in G.

An edge boundary partition of a graph G is a collection Π of subsets of V(G) such that

- ▶ G[U] is connected for all $U \in \Pi$ and
- ▶ { $E_G(U)$: $U \in \Pi$ } form a partition of E(G).

Every edge boundary partition Π of *G* determines a vertex covering multiplicity vector χ_{Π} which maps $v \in V(G)$ to the size of the multiset { $S : v \in S \in \Pi$ }.

King I (王一)



A <u>nested</u> edge boundary partition: $\Pi_1 = \{\{v_1\}, \{v_2\}, \{v_3\}, \{v_4\}, \{v_5\}\}\$ $\chi_{\Pi_1}(v_1) = 1, \chi_{\Pi_1}(X) = 0, \chi_{\Pi_1}(u) = 0$

King II (王二)



A <u>nested</u> edge boundary partition: $\Pi_2 = \{\{v_1, v_2, v_3, u\}, \{u\}, \{v_4\}, \{v_5\}\}$ $\chi_{\Pi_2}(v_1) = 1, \chi_{\Pi_2}(X) = 0, \chi_{\Pi_2}(u) = 2$

King III (王三)



A nonnested edge boundary partition:

 $\Pi_3 = \{\{v_1, v_3, u\}, \{u, v_2\}, \{v_4\}, \{v_5\}\}\$ $\chi_{\Pi_3} = \chi_{\Pi_2}$

- An edge boundary partition Π should in general be viewed as a multiset. But for a connected graph G, the only possible elements appeared in Π with multiplicity greater than 1 are only V(G) and Ø.
- It is easy to see that a graph has an edge boundary partition if and only if it is a bipartite graph.

Representing partitions on trees (允許樹表現之劃分系統)

Due to the phylogenetics background, Huber-Moulton-Semple-Wu (2014) initiate the study of those partitions of a set X which can generate a weighted split system of X represented on a tree.

A reformulation of their main concern using the concept of edge boundary partition is as follows:

Let *T* be a tree with leaf set *X*. An edge boundary partition Π of *T* is normal if χ_{Π} takes value zero on *X*. What is the global structure of all the normal edge boundary partitions of *T*?

Two operations (兩種變換)

Let *G* be a bipartite graph and let \mathcal{EBP}_G be the set of all edge boundary partitions of *G*. For any $U \subseteq V(G)$, let $f_G(U)$ be the set of connected components of G[U].

We define two natural operations which are self-maps of \mathcal{EBP}_G . Let $\Pi \in \mathcal{EBP}_G$.

- ▶ Operation I (decreasing vertex covering multiplicity): Take $A, B \in \Pi$ such that $A \subseteq B$, and set $\Pi' = (\Pi \{A, B\}) \cup f_G(B A);$
- ▶ Operation II (increasing nestedness): Take $A, B \in \Pi$ and set $\Pi' = (\Pi \{A, B\}) \cup \{A \cup B\} \cup f_G(A \cap B)$.

Two operations (兩種變換)

Let *G* be a bipartite graph and let \mathcal{EBP}_G be the set of all edge boundary partitions of *G*. For any $U \subseteq V(G)$, let $f_G(U)$ be the set of connected components of G[U].

We define two natural operations which are self-maps of \mathcal{EBP}_G . Let $\Pi \in \mathcal{EBP}_G$.

- ▶ Operation I (decreasing vertex covering multiplicity): Take $A, B \in \Pi$ such that $A \subseteq B$, and set $\Pi' = (\Pi \{A, B\}) \cup f_G(B A);$
- ▶ Operation II (increasing nestedness): Take $A, B \in \Pi$ and set $\Pi' = (\Pi \{A, B\}) \cup \{A \cup B\} \cup f_G(A \cap B)$.

Operation II : $\Pi_3 \rightarrow \Pi_2$

Operation I: $\Pi_2 \rightarrow \Pi_1$

Poset structure of the edge boundary partitions (偏序結構)

For any $\Pi_1, \Pi_2 \in \mathcal{EBP}_G$, we write

 $\Pi_1 \Subset \Pi_2$

whenever we can transform Π_2 into Π_1 via a sequence of Operations I and II.

Lemma (W., Xu). The binary relation \Subset gives a partial order on \mathcal{EBP}_G .

Operation II and partitions with the same multiplicity (變換 二與覆蓋重數一致的邊集界線劃分)

Let \mathcal{NEBP}_G be the set of nested edge boundary partitions of G. For every $\Pi \in \mathcal{NEBP}_G$, let L_{Π} consist of those elements Π' of \mathcal{EBP}_G with $\chi_{\Pi} = \chi_{\Pi'}$.

Lemma (W., Xu). Take $\Pi \in \mathcal{NEBP}_G$. A map $\Pi' \in \mathcal{NEBP}_G$ is equal to Π if and only if $\chi_{\Pi} = \chi_{\Pi'}$. Operation II sends L_{Π} to itself and every element in L_{Π} can be transformed to Π by a sequence of Operation II.

Problem (W., Xu). Take a tree *G* and a map $\Pi \in \mathcal{NEBP}_G$. Is it true that the subposet of $(\mathcal{EBP}_G, \Subset)$ induced by L_{Π} is a ranked poset?

The problem has a positive answer when G is a path. In that case, that subposet restricted to normal edge boundary partitions corresponds to an interval in the strong Bruhat order of the symmetric group.

Distributive lattice of nested edge boundary partition (嵌套 邊集界線劃分構成的分佈格)

Let *G* be a connected bipartite graph with partite sets V_0 and V_1 . For $i \in \{0, 1\}$, let $\mathcal{NEBP}_G(i)$ represent the elements $\Pi \in \mathcal{NEBP}_G$ such that χ_{Π} takes odd value on V_i and takes even value on V_{1-i} .

Theorem (W., Xu). For any $i \in \{0, 1\}$, the subposet of $(\mathcal{EBP}_G, \Subset)$ induced by $\mathcal{NEBP}_G(i)$ is a (ranked) distributive lattice.

When *G* is the path of even length 2n, the distributive lattice induced by the set of normal nested edge boundary partitions is just the strong Bruhat order restricted on 312-avoiding permutations and its size is the Catalan number $C_n = \frac{\binom{2n}{n+1}}{\binom{2n}{n+1}}$ (Barcucci-Bernini-Ferrari-Poneti, 2005).

Operation I and Bruhat order (變換一與布呂阿序)

 $\Pi_1 = \{ \{c\}, \{b, c, d\}, \{f\}, \{a, b, c, d, e, f, g\} \}$ $\Pi_2 = \{ \{b\}, \{d\}, \{f\}, \{a, b, c, d, e, f, g\} \}$



Normal nested edge boundary partitions of a path vs the Bruhat order (路的正規邊集界線劃分與布呂阿序)



Normal nested edge boundary partitions of a tree (樹的正 規嵌套邊集界線劃分)



Figure: All vertex covering multiplicity vectors take value 0 on $\{a, b, c, d, e, f\}$ and value 1 on $\{g, h, i, j, k, \ell\}$. We thus only record their values on $\{m, n, o, p, q\}$.

A possible generalization of Catalan number (卡特蘭數的 可能推廣)

Problem (W., Xu). What is the size of the lattice of all normal nested edge boundary partitions of a tree?

Once there was a tree... (從前有顆樹...)



https://allpoetry.com/poem/
8538991-The-Giving-Tree-by-Shel-Silverstein