Isotropic models of evolution with symmetries

Weronika Buczyńska, Maria Donten, and Jarosław A. Wiśniewski†

Dedicated to Andrew Sommese for his 60th birthday.

Abstract. We consider isotropic Markov models on (phylogenetic) trees whose models of evolution are symmetric, that is invariant with respect to a transitive group of permutations of letters whose evolution we consider. Transitivity of the action of the group of symmetries implies strong bounds on the space of parameters of such a model. A special consideration is given to groups of symmetries containing large abelian subgroups. We prove that only hyperbinary models have abelian groups of symmetries. Using GAP, a computer algebra program, we calculate a complete classification of symmetric isotropic models on d letters, where $d \leq 9$.

0. Introduction

The present paper is inspired by questions arising in algebraic statistics. We consider geometric models of Markov processes on phylogenetic trees or, since the biological context may be a bit too restrictive or even misleading, geometric models of hidden Markov processes on trees.

In a nutshell, the nature of the Markov process on a tree depends on three elements. Firstly, one fixes a finite set $\mathcal{A}$, elements of which are called *letters* and stand for features whose evolution we want to trace. Secondly, a graph, a *tree* $\mathcal{T}$. The vertices of $\mathcal{T}$ have assigned random variables whose states space is $\mathcal{A}$. One may think about vertices as associated to states of the process, each vertex may assume values in the set of letters. And thirdly, a *model of evolution* which describes the way the states with values in $\mathcal{A}$ possibly evolve along the edges of the graph. This boils down to defining conditional probability relating random variables associated to ends of each edge of $\mathcal{T}$ so that the model of evolution is a space parameterizing possible values of the conditional probability. The output of the process can be described by a *geometric model* of such a triple which provides information about a possible distribution of the letters over the leaves of the tree $\mathcal{T}$. More precisely, the geometric model is the locus of probability distributions arising from the given tree

1991 *Mathematics Subject Classification.* Primary 14Mxx; Secondary 14M25, 20B25, 20B30, 20B40, 62P10.

Key words and phrases. Markov process on a tree, phylogenetic tree, geometric model, model of evolution, symmetries.

† The third named author is supported by a grant of Polish MNiSzW (N N201 2653 33); the text was written up when he was an Alexander von Humboldt Fellow at Freie Universität, Berlin.

$\mathcal{T}$, with different parameter choices given by the model of evolution, in the space of all possible distributions on the leaves of $\mathcal{T}$.

A reader who finds this short paragraph above confusing rather than explanatory is requested to look into one of the standard references in the field, [**SS03**] or [**PS05**, Part I], for both proper explanation and biological context. In the present paper we concentrate on algebraic and geometrical aspects of Markov processes on trees, for which proper definitions are stated in Section 1.

We discuss the situation when such a process is *isotropic*. By this we mean that the tree is not directed (unrooted) and has a uniform distribution at the root, and the matrices describing conditional probability are symmetric. We avoid using the word "symmetric" in this context and use "isotropic" instead, since the other feature of our interest are symmetries in the set $\mathcal{A}$. That is, we postulate that the set of letters $\mathcal{A}$ admits symmetries described by a group of permutations of $\mathcal{A}$ and the model of evolution is in agreement with such symmetries.

The trick of using symmetries to reduce complexity of a problem is standard in physics and other natural sciences. In fact, in the study of phylogenetic trees related to evolution of four biologically meaningful letters A, C, G, T there are standard models with symmetries. The symmetries are implied by biological or biochemical constraints.

In the early 90's Evans and Speed, [**ES93**], as well as Székely, Steel and Erdös, [**SSE93**], studied (non-isotropic) models of evolution whose symmetries are abelian groups. The geometric models of these, so called *group-based models*, admit a particularly nice description based on the Fourier calculus for finite abelian groups. These geometric models were later described by Sturmfels and Sullivant as *toric varieties*, which made a firm connection of this subject with algebraic geometry, [**SS04**]. Recently, *equivariant models*, defined in a very general set-up, have been considered by Draisma and Kuttler, [**DK07**].

Our starting point and objectives are somewhat different from these of the above mentioned works. We assume that the models are isotropic and their groups of symmetries act transitively on the set of letters $\mathcal{A}$. We do not assume that the group of symmetries is abelian (in fact, it is hardly ever abelian, as we show).

Our interest is in pure algebraic geometry rather than in its applications in biology or algebraic statistics. In our earlier paper [**BW07**] we studied isotropic binary models and proved that geometric models of two trivalent trees with the same number of leaves can be deformed one to the other. Next, Sturmfels and Xu [**SX08**] proved that, in fact, these geometric models are specializations of somehow more general objects (spectra of Cox-Nagata rings), related to pointed rational curves. This opens a question of finding other structures defined on trees (e.g. geometric models of isotropic Markov processes) which enjoy a similar property.

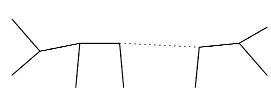
The paper is organized as follows. In Section 1 we define the notion of an *isotropic Markov process* on a phylogenetic tree $(\mathcal{T}, W, \widehat{W})$ and its *geometric model* $X(\mathcal{T}) = X(\mathcal{T}, W, \widehat{W})$. Here W is a linear space with $\mathcal{A}$ as basis and $\widehat{W} \subset S^2W$ stands for the *model of evolution*. This section is, essentially, an unpublished part of paper [**BW06**]. In Section 2 we discuss symmetries of $\mathcal{A}$. That is, we consider a finite group G of permutations of $\mathcal{A}$ and associated model $\widehat{W}_G$ which is fixed by symmetries from G. If G acts transitively on $\mathcal{A}$ then $\widehat{W}_G$ is called a *symmetric model of evolution* on $\mathcal{A}$. We introduce a notion of *saturated groups* whose conjugacy classes are in bijection with (conjugacy classes of) symmetric models. We note that geometric models of a tree for conjugate groups are isomorphic. Next, in Section 3 we examine the case when G is hyperbinary, that is when $G = \mathbb{B}^n := \mathbb{Z}_2^n$, $\mathcal{A} = \mathbb{B}^n$, so that $|\mathcal{A}| = 2^n$ and $\mathcal{A}$ can be identified with G with the regular action on itself. This is the only isotropic group-based model, see 3.8 and 3.10. Then we present the results of calculations for low dimensional cases. That is, using [**GAP**] the second author computed pairs $(G, \widehat{W}_G)$ of saturated groups of permutations and their symmetric models of evolution for $|\mathcal{A}| \leq 9$. They are presented together with respective inclusions (or nesting of models, or Felsenstein's hierarchy), c.f. [**PS05**, Sect. 4.5.1]. In the last section of the paper we discuss the situation when the group of symmetries of an isotropic model contains an abelian subgroup acting transitively on $\mathcal{A}$. This is the situation when our set-up is close to that of group-based model.

We thank the referee of this paper for remarks which helped us to put our work in the proper context and to improve the exposition. We also thank the editors of the present volume for organizing an excellent anniversary conference and preparing this publication.

1. Isotropic processes.

Notation 1.1. A tree $\mathcal{T}$ is a simply connected graph (1-dimensional CW complex) with a set of edges $\mathcal{E} = \mathcal{E}(\mathcal{T})$ and vertices $\mathcal{V} = \mathcal{V}(\mathcal{T})$ and the (unordered) boundary map $\partial : \mathcal{E} \to \mathcal{V}^{\wedge 2}$, where $\mathcal{V}^{\wedge 2}$ denotes the set of unordered pairs of distinct elements in $\mathcal{V}$. We write $\partial(e) = \{\partial_1(e), \partial_2(e)\}$, or equivalently $e = \langle \partial_1(e), \partial_2(e) \rangle$, and say v is a vertex of e, or e contains v if $v \in \{\partial_1(e), \partial_2(e)\}$, we simply write $v \in e$. The valency of a vertex v is the number of edges which contain v (the valency is positive since $\mathcal{T}$ is connected and we assume it has at least one edge). A vertex v is called a leaf if its valency is 1, otherwise it is called an inner vertex or a node. If the valency of each inner node is m then the tree will be called m-valent. The set of leaves and nodes will be denoted $\mathcal{L}$ and $\mathcal{N}$, respectively, $\mathcal{V} = \mathcal{L} \cup \mathcal{N}$. An edge which contains a leaf is called a petiole, an edge which is not a petiole is called an inner edge (or branch).

Example 1.2. A *star* is a tree with exactly one inner node. A *caterpillar* is a 3-valent tree such that there are exactly two inner nodes to which there are attached two petioles, any other inner node has exactly one petiole attached.



$\circ$

Set-Up 1.3. Let W be a (complex) vector space of dimension d, which we will call a model space of states on the tree $\mathcal{T}$, with a distinguished basis $\mathcal{A} = \{\alpha_1, \alpha_2, \ldots \alpha_d\}$, the elements of which will be called *letters*. We set $\alpha = \alpha_1 + \cdots + \alpha_d$.

By α_i^* we denote elements of the dual basis of the dual space $W^* = Hom(W, \mathbb{C})$. The pairing of W^* and W will be understood as the action of functionals on vectors, or the other way around, so that $\alpha_i(\alpha_j^*) = \alpha_j^*(\alpha_i)$ is 1 or 0 depending on whether $i = j$ or $i \neq j$. Alternatively, we can fix an inner product on W for which α_i's make an orthonormal basis, then the product allows to identify W with W^* and α_i with α_i^*.

We fix a linear map $\sigma : W \to \mathbb{C}$, such that $\sigma(\alpha_i) = 1$ for every i, that is $\sigma = \sum \alpha_i^*$. Therefore, σ is equivalent to α in terms of the above inner product. The map σ will be called *normalization*.

Let $\widehat{W}$ be a subspace of the second tensor product $W \otimes W$. We will call the pair $(W, \widehat{W})$ (or just the space $\widehat{W}$ if W is fixed) a *model of evolution* on d letters. An element $A = \sum_{i,j} a_{ij}(\alpha_i \otimes \alpha_j)$ of $\widehat{W}$ can be represented as a matrix $A = (a_{ij})$ (by abuse we use the same letter A to denote both), where a_{ij} are obtained by evaluating A on elements of the dual basis, that is $a_{ij} = A(\alpha_i^*, \alpha_j^*)$. Equivalently, the identification $W \simeq W^*$ yields $W \otimes W \simeq W \otimes W^* = End(W)$ and A can be interpreted as an endomorphism of W. Throughout the present paper we will assume that $\widehat{W}$ is contained in the symmetric product $S^2(W)$ and we will call such $\widehat{W}$ an *isotropic model of evolution*.

In addition, we will assume that $\widehat{W}(\sigma, \ \cdot \ ) \subset \mathbb{C} \cdot \alpha$ or, equivalently, that for every matrix $A = \sum_{i,j} a_{ij}(\alpha_i \otimes \alpha_j)$ in $\widehat{W}$ the sum of elements in each row (and, since the matrix is symmetric, also in each column) is the same. This means that, up to a multiplicative constant, elements of $\widehat{W}$ are *doubly stochastic* matrices.

We note that in case of transitive group action, see 2.1, which is the main case considered in the present paper, this assumption turns out to be redundant, see 2.8.

Given a tree $\mathcal{T}$ and a vector space W, and a subspace $\widehat{W} \subset W \otimes W$ we associate to any vertex v of $\mathcal{V}(\mathcal{T})$ a copy of W denoted by W_v and for any edge $e \in \mathcal{E}(\mathcal{T})$ we associate a copy of $\widehat{W}$ understood as the subspace in the tensor product $\widehat{W}^e \subset W_{\partial_1(e)} \otimes W_{\partial_2(e)}$. Note that although the pair $\{\partial_1(e), \partial_2(e)\}$ is unordered, this definition makes sense since $\widehat{W}$ consists of symmetric tensors. Elements of $\widehat{W}^e$ will be written as (symmetric) matrices $(a^e_{\alpha_i,\alpha_j}) = (a^e_{i,j})$.

In the present paper we adopt the following definition.

Definition 1.4. *The triple $(\mathcal{T}, W, \widehat{W})$ described above is called an isotropic model (or a Markov process) on a phylogenetic tree.*

Frequently however, by abuse of language, we will call the triple $(\mathcal{T}, W, \widehat{W})$ just a phylogenetic tree. Since the whole structure $(\mathcal{T}, W, \widehat{W})$ is the object of our interest (not the leaf-labelled tree $\mathcal{T}$ alone, as a combinatorial structure), this abbreviation should cause no confusion.

Discussion 1.5. The motivation for the set-up comes from statistics with compromises coming from the usage of linear algebra rather than explicit statistical language.

Roughly speaking, from the point of view of statistics, a Markov process on phylogenetic tree is a collection of random variables ξ_v with values in a set of letters associated to nodes of $\mathcal{T}$, together with collection of rules for inheritance, that is of conditional (or transition) probability, labeled by edges of $\mathcal{T}$.

In the set-up above the space W is spanned on letters from the set $\mathcal{A}$, the model space of states for variables ξ_v. The statistically meaningful domain in W is the probabilistic simplex described by the following conditions in terms of coordinates (basis dual to $\mathcal{A}$): $\mathrm{im}(\alpha_i^*) = 0$, for all i (i.e. we consider the real part of the complex vector space W), and $\alpha_i^* \geq 0$, for $i = 1, \ldots, d$, and with normalization $\sigma = 1$. Given $v \in \mathcal{V}$ the dual basis of W_v describes probability distribution of the random variable

ξ_v. That is, $P(\xi_v = \alpha_i) \sim \alpha_i^*(w_v)$, where w_v is a vector of W_v (we can call w_v the state of $\mathcal{T}$ at vertex v). Here $\sim$ stands for proportionality and this is the form σ which provides a somewhat more accurate definition $P(\xi_v = \alpha_i) = \alpha_i^*(w_v)/\sigma(w_v)$ which makes sense within the real non-negative orthant of W.

The model of evolution, that is the space $\widehat{W}$, is meant to provide the rules according to which the states are inherited along the edges of $\mathcal{T}$. That is, given $e \in \mathcal{E}$, a tensor (or matrix) $A \in \widehat{W}^e$ has entries

$$a_{ij} = A(\alpha_i^*, \alpha_j^*) \sim P(\xi_{\partial(e)_1} = \alpha_i \mid \xi_{\partial(e)_2} = \alpha_j)$$

Here, again, $\sim$ means that the actual equality makes sense when the entries of A are real and non-negative, and the sum of every row (and column) is 1, i.e. when A is doubly stochastic.

Example 1.6. Let us discuss some natural examples of isotropic models of evolution. Recall that, in the matrix representation of an element of $\widehat{W}$, the sum of the numbers in each row and each column is the same. If W is of dimension 2 this is equivalent to saying that the matrices in $\widehat{W}$ are of the form

$$\begin{bmatrix} a & b \\ b & a \end{bmatrix}$$

for some a and b in $\mathbb{C}$. Thus $\widehat{W}$ is of dimension 2 and the only interesting example of isotropic model for $d = 2$ since any proper subspace is of dimension ≤ 1 hence trivial when it comes to normalizing.

If $d = 4$, which is a case of particular interest in biology, then there are a few nontrivial choices for $\widehat{W}$. The most general case consists of the space of symmetric matrices such that the sum of the numbers in each row and each column is the same. The other two commonly known options are as follows: The Kimura 3-parameter model with $\dim \widehat{W} = 4$ and matrices of the form

$$\begin{bmatrix} a & b & c & d \\ b & a & d & c \\ c & d & a & b \\ d & c & b & a \end{bmatrix}$$

and *isotropic* (!) strand symmetric model, c.f. [**CS05**],

$$\begin{bmatrix} a & b & c & d \\ b & e & f & c \\ c & f & e & b \\ d & c & b & a \end{bmatrix}$$

where, because the matrix is proportional to doubly stochastic, it holds $a+d = e+f$, hence $\dim \widehat{W} = 5$.

$\circ$

Construction 1.7. The boundary map $\partial : \mathcal{E} \to \mathcal{V}^{\wedge 2}$ from 1.1 has its incarnation on the level of tensor products of vector spaces associated to both vertices and edges of the tree $\mathcal{T}$.

Let us consider a linear map of tensor products

$$\widehat{\Psi}_{W,\widehat{W}} : \widehat{W}^{\mathcal{E}} = \bigotimes_{e \in \mathcal{E}} \widehat{W}^e \longrightarrow W_{\mathcal{V}} = \bigotimes_{v \in \mathcal{V}} W_v$$

defined by setting its dual as follows

$$\widehat{\Psi}^*_{W,\widehat{W}}(\otimes_{v\in\mathcal{V}} \ \ \alpha_v^*) = \otimes_{e\in\mathcal{E}} \ (\alpha_{\partial_1(e)}\otimes\alpha_{\partial_2(e)})^*_{|\widehat{W}^e}$$

where α_v stands for an element of the chosen basis $\mathcal{A} = \{\alpha_1, \dots, \alpha_d\}$ of the space W_v. We will skip the subscripts in $\widehat{\Psi}_{W,\widehat{W}}$ and write just $\widehat{\Psi}$ if the model of evolution is known.

The *complete affine model* of the phylogenetic tree (or, more precisely, of the Markov process on the phylogenetic tree) $(\mathcal{T}, W, \widehat{W})$ is the closure of the image of the associated multi-linear map

$$\widetilde{\Psi} : \prod_{e\in\mathcal{E}} \widehat{W}^e \longrightarrow W_{\mathcal{V}} = \bigotimes_{v\in\mathcal{V}} W_v$$

We note that, by definition, for any function $\mathcal{V} \ni v \longrightarrow \mu(v) \in \{1, 2, \dots d\}$ and any point in the parameter space $\left(A^e = (a^e_{ij})\right)_{e\in\mathcal{E}} \in \prod_{e\in\mathcal{E}} \widehat{W}^e$ the respective coordinate of its image under the map $\widetilde{\Psi}$ in the tensor product $\bigotimes_{v\in\mathcal{V}} W_v$ is to be calculated as follows

$$(\otimes_{v\in\mathcal{V}} \ \ \alpha^*_{\mu(v)}) \left(\widetilde{\Psi}\left(A^e = (a^e_{ij})\right)_{e\in\mathcal{E}}\right) \ \ = \prod_{e=\langle u,v\rangle\in\mathcal{E}} a^e_{\mu(u)\mu(v)}$$

The induced rational map of projective varieties will be denoted by Ψ:

$$\Psi : \prod_{e\in\mathcal{E}} \mathbb{P}(\widehat{W}^e) \dashrightarrow \mathbb{P}(W_{\mathcal{V}}) = \mathbb{P}(\bigotimes_{v\in\mathcal{V}} W_v)$$

and the closure of the image of Ψ is called the *complete projective model*, or just the *complete model* of $(\mathcal{T}, W, \widehat{W})$. The maps $\widetilde{\Psi}$ and Ψ are called the parametrization of the respective model.

Given a set of vertices of the tree we can "hide" them by applying the map $\sigma = \sum_i \alpha_i^*$ to their tensor factors. In what follows we will hide all inner nodes and project to tensor product of model spaces associated to leaves. That is, we consider the map

$$\Pi_{\mathcal{L}} : W_{\mathcal{V}} = \bigotimes_{v\in\mathcal{V}} W_v \to W_{\mathcal{L}} = \bigotimes_{v\in\mathcal{L}} W_v$$

$$\Pi_{\mathcal{L}} = (\otimes_{v\in\mathcal{L}} \ id_{W_v}) \otimes (\otimes_{v\in\mathcal{N}} \ \sigma_{W_v})$$

DEFINITION 1.8. *The affine model of a phylogenetic tree (or of a Markov process on a phylogenetic tree) $(\mathcal{T}, W, \widehat{W})$ is an affine subvariety of $W_{\mathcal{L}} = \bigotimes_{v\in\mathcal{L}} W_v$ which is the closure of the image of the composition $\widetilde{\Phi} = \Pi_{\mathcal{L}} \circ \widetilde{\Psi}$. The projective model, denoted by $X(\mathcal{T}, W, \widehat{W})$ or just $X(\mathcal{T})$ if W and $\widehat{W}$ are fixed, is the underlying projective variety in $\mathbb{P}(W_{\mathcal{L}})$.*

Note that $X(\mathcal{T})$ is the closure of the image of the respective rational map

$$\Phi : \prod_{e\in\mathcal{E}} \mathbb{P}(\widehat{W}^e) \dashrightarrow \mathbb{P}\left(\bigotimes_{v\in\mathcal{L}} W_v\right)$$

which is defined by a special linear subsystem in the complete Segre linear system $|\bigotimes_{e\in\mathcal{E}} p^*_{\mathbb{P}(\widehat{W}^e)}\mathcal{O}_{\mathbb{P}(\widehat{W}^e)}(1)|$, where $p^*_{\mathbb{P}(\widehat{W}^e)}$ is the projection from the product to the respective component. We will call this map a *rational parametrization* of the model.

The coordinates of $\widetilde{\Phi}$ can be computed as follows: for any function $\mathcal{L} \ni v \longrightarrow \mu(v) \in \{1, 2, \ldots, d\}$, which describes the distribution of letters on leaves of $\mathcal{L}$, the respective coordinate of the tensor product $\bigotimes_{v\in\mathcal{L}} W_v$ is determined as follows

$$(\otimes_{v\in\mathcal{L}} \ \ \alpha^*_{\mu(v)}) \left(\widetilde{\Phi}\left(A^e = (a^e_{ij})\right)_{e\in\mathcal{E}}\right) \ = \sum_{\widehat{\mu}} \left(\prod_{e=\langle u,v\rangle\in\mathcal{E}} a^e_{\widehat{\mu}(u)\widehat{\mu}(v)}\right)$$

where the sum is taken over all functions $\widehat{\mu} : \mathcal{V} \longrightarrow \{1, 2, \ldots d\}$ which extend μ.

Discussion 1.9. In the case of a Markov process on a tree $\mathcal{T}$ we fix a root $r \in \mathcal{V}$ and this implies an order $<$ on $\mathcal{V} = \mathcal{L} \cup \mathcal{N}$. Thus, every edge $e \in \mathcal{E}$ is directed, which we denote by $e = \langle u < v\rangle$.

Random variables ξ_v determine a Markov process on $\mathcal{T}$ if the value of ξ_v depends only on the value of ξ_u, where u is the node immediately preceding v in terms of the order $<$.

This determines the distribution of variables ξ_v in terms of the initial probability distribution at the root and the relative probability along every edge. That is, for any function $\mathcal{V} \ni v \longrightarrow \mu(v) \in \{1, 2, \ldots d\}$

$$P\left(\bigcap_{v\in\mathcal{V}} (\xi_v = \alpha_{\mu(v)})\right) = P\left(\xi_r = \alpha_{\mu(r)}\right) \cdot \left(\prod_{e=\langle u<v\rangle\in\mathcal{E}} P\left(\xi_v = \alpha_{\mu(v)} \mid \xi_u = \alpha_{\mu(u)}\right)\right)$$

We note that this formula is proportional to the one describing the coordinates of the parametrization of the complete affine model of $\mathcal{T}$, with identifications described in 1.5, provided that the initial distribution at the root is uniform, that is $P(\xi_r = \alpha_i) = 1/d$ for $i = 1, \ldots, d$.

The above definition of parametrization is an unrooted, isotropic and algebraicized version of what is commonly considered in the literature, see e.g. [**SS03**, Sect. 8] or [**PS05**, Sect. 1.4.4], or [**DK07**].

2. Symmetric models of evolution.

Notation 2.1. Let $G \leq S_d$ be a subgroup of group of permutations of d elements. In the present paper these will be the letters in the set $\mathcal{A} = \{\alpha_1, \alpha_2, \ldots, \alpha_d\}$. We will sometimes confuse $\mathcal{A}$ with its set of indices $\{1, \ldots, d\}$ and write $g(\alpha_i) = a_{g(i)}$, where $g \in G$. In these terms we will write decomposition of elements of G into cycles, i.e. $g = (i_1, \ldots, i_r) \cdots$. By $\leq$ or $<$ we will denote the inclusion of groups while $\langle g\rangle \leq G$ will denote a subgroup generated by g. We will say that the subgroup $G \leq S_d$ is *transitive* if its action on $\mathcal{A}$ is transitive.

We extend the action of S_d on letters to a linear action on W, the vector space spanned on $\mathcal{A}$. That is, we consider the natural representation $\rho : S_d \longrightarrow GL(W)$ which yields a representation $\rho_G : G \longrightarrow O(W) < GL(W)$, where $O(W)$ is the group of orthogonal transformations preserving the inner product on W in which α_i's make an orthonormal basis. If no confusion is likely, we will write just g for both $g \in S_d$ as well as for the matrix $\rho_G(g) \in O(W) < GL(W)$. That is, for $w \in W$, we write $\rho_G(g)(w) = g \cdot w$, where on the right hand side w is understood as a column of coefficients of w in basis $\mathcal{A}$.

Recall that the mentioned above inner product (or the choice of the dual basis) allows us to identify W with W^*. Note that, for $g \in S_d$, we have $\alpha^*_i(\alpha_{g(j)}) = \alpha^*_{g^{-1}(i)}(\alpha_j)$. That is, the right action of G on W^*, defined as $\rho^*_G(g)(u) = u \cdot$

$g^{-1} = u \cdot g^t$, where g^t means transposition, makes the identification $W \simeq W^*$ G-equivariant.

Therefore an induced action of G on the product $W \otimes W$ and, eventually, on S^2W can be described in terms of the adjoint action Ad_G of G on $End(W)$. That is, if $A \in S^2W \subset W \otimes W$ is represented by a symmetric matrix $A \in End(W) = W \otimes W^*$ and $g \in G \subset O(d)$ then we have

$$S^2\rho_G(g)(A) = g \cdot A \cdot g^t = g \cdot A \cdot g^{-1} = Ad_G(g)(A)$$

In plain words it means that the permutation $g \in G \leq S_d$ permutes columns and rows of the matrix A as it does with the elements of the set $\mathcal{A}$. That is, if $A = (a_{ij})$ then $g(A)$'s entry in the i-th row and j-th column is equal to $a_{g^{-1}(i),g^{-1}(j)}$.

By $Fix(G)$ or $Fix(\rho_G)$ we will denote the subspace of fixed points of the action ρ_G. Similarly, by $Fix(\rho_G^*)$ or $Fix(Ad_G)$ we denote the fixed point sets of the respective representations. Clearly, $Fix(\rho_{S_d})$ contains α and its dual contains σ. In fact we have the following easy observation.

LEMMA 2.2. *A subgroup $G < S_d$ is transitive if and only if $Fix(\rho_G) = \mathbb{C} \cdot \alpha$ or, equivalently, $Fix(\rho_G^*) = \mathbb{C} \cdot \sigma$.*

We adopt the following definition.

DEFINITION 2.3. *An isotropic model of evolution $\widehat{W} \subset S^2W$ on d letters is called symmetric if $\widehat{W} = Fix(S^2(\rho_G))$ for some transitive group of symmetries $G \leq S_d$.*

In view of the discussion preceeding the definition the elements of $\widehat{W}$ can be identified with matrices whose entries are invariant with respect to permutation of rows and columns by elements of the group G.

EXAMPLE 2.4. Let $h \in S_d$ be a cyclic permutation of length d, say $h = (1, \ldots, d)$, and let $H = \langle h \rangle$ be the group generated by h. Then $\widehat{W}_H$ consists of matrices of the form

$$\begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_2 & a_1 \\ a_1 & a_0 & a_1 & \cdots & a_3 & a_2 \\ a_2 & a_1 & a_0 & \cdots & a_4 & a_3 \\ \cdots & \cdots & \cdots & & \cdots & \cdots \\ a_2 & a_3 & a_4 & \cdots & a_0 & a_1 \\ a_1 & a_2 & a_3 & \cdots & a_1 & a_0 \end{pmatrix}$$

where a_i are arbitrary numbers, hence $\dim \widehat{W}_H = (d+1)/2$ if d is odd and $d/2+1$ if d is even. We note that the above matrix can be written as the following linear combination of matrices

$$a_0 \cdot id + a_1 \cdot \left(h + h^{-1}\right) + a_2 \cdot \left(h^2 + h^{-2}\right) + \cdots$$

with h here being presented as a matrix. Every such a matrix is symmetric with respect to its center hence it is fixed by an involution ν of the type $\nu = (1, d)(2, d-1), \cdots$ such that $h^{-1} = \nu \cdot h \cdot \nu^{-1}$. Thus $\widehat{W}_H$ is fixed not only by the cyclic group H but also by the dihedral group $\langle h,\ \nu \rangle$. These observations will be generalized in section 5.

The isotropic model of evolution presented above will be called the *dihedral* model.

$\circ$

A subgroup $G \leq S_d$ is called *saturated* if for any group H, such that $G \leq H \leq S_d$ and $Fix(S^2\rho_G) = Fix(S^2\rho_H)$, it follows that $H = G$. In other words, G is saturated if it is the stabilizer of $Fix(S^2\rho_G)$.

We have yet another immediate observation and thus a subsequent definition.

LEMMA 2.5. *There is an inclusion reversing bijection between transitive saturated subgroups of S_d and isotropic symmetric models of evolution on d letter.*

DEFINITION 2.6. *If $(W, \widehat{W})$ is an isotropic symmetric model of evolution on d letters then its group of symmetries is the unique transitive saturated $G \leq S_d$ such that $Fix(S^2\rho_G) = \widehat{W}$.*

Since conjugating subgroups of S_d is just renaming its elements we can identify models associated to conjugate subgroups.

PROPOSITION 2.7. *Let $\mathcal{T}$ be a tree and let $\widehat{W}_H$, $\widehat{W}_G$ be two models of evolution on d letters with groups of symmetries H and G, respectively. The inclusion of groups $H \leq G \leq S_d$ implies an inclusion of geometric models $X(\mathcal{T}, W, \widehat{W}_G) \subset X(\mathcal{T}, W, \widehat{W}_H)$. If groups G and H are conjugate in S_d then, after some linear change of coordinates in $\mathbb{P}(W_{\mathcal{L}})$, the models $X(\mathcal{T}, W, \widehat{W}_G)$ and $X(\mathcal{T}, W, \widehat{W}_H)$ are equal.*

Thus, in what follows we will look at a classification of saturated transitive subgroups of S_d, up to conjugation.

The assumption that matrices in $\widehat{W}$ are doubly stochastic, up to a multiplicative constant, is redundant for transitive groups of symmetries.

LEMMA 2.8. *Let $G \leq S_d$ be a transitive subgroup. If a matrix $A \in S^2W$ is fixed by $S^2\rho_G$ then the sum of rows (columns) of A is constant.*

PROOF. Recall that $\alpha = \alpha_1 + \cdots + \alpha_d$, $\sigma = \alpha_1^* + \cdots + \alpha_d^*$ and we are to verify the condition that A evaluated on σ is a multiplicity of α. But for every $g \in G < O(W)$ we have $A = g \cdot A \cdot g^{-1}$ hence

$$g\,(A(\sigma)) = g \cdot g^{-1} \cdot A \cdot (g(\sigma)) = A(\sigma)$$

and we are done by 2.2 $\square$

In addition, for transitive groups $G \leq S_d$ we have a bound on the dimension of $\widehat{W}_G = Fix(S^2\rho_G)$.

LEMMA 2.9. *If $G \leq S_d$ is transitive then* $\dim \widehat{W} \leq d$. *Moreover, if d is odd then* $\dim \widehat{W} \leq (d+1)/2$.

PROOF. Let us write a general element $A \in \widehat{W}_G$ as a matrix $A = (a_{ij})$ and note that if an element appears in the first row then, because of transitivity, it has to appear in every row. That is, for $i = 1, \ldots, d$ there exists $g_i \in G$ such that $g_i(1) = i$ and, for such g_i and any $j = 1, \ldots, d$, it holds

$$a_{1,j} = a_{2,g_2(j)} = \cdots = a_{d,g_d(j)}$$

Therefore the number of linearly independent coefficients in A can not exceed the length of the row, that is d. This proves the first statement of the lemma. By the same argument all the coefficients on the diagonal of A are equal. By symmetry of A each coefficient outside the diagonal appears the same number of times above the diagonal as below the diagonal. Thus, if d is odd then every coefficient outside the diagonal appears at least $2d$ times, hence the second part follows. $\square$

The above argument can be extended to the following.

LEMMA 2.10. *Suppose that $G \leq S_d$ is transitive. Let $G_1 = G_{\alpha_1} < G$ be the subgroup fixing α_1. Then the dimension of $\widehat{W}_G$ does not exceed the number of orbits of G_{α_1} in the set $\mathcal{A}$.*

PROOF. Let $g \in G_{\alpha_1}$ then $g(\alpha_i) = \alpha_j$ implies $a_{1,i} = a_{1,j}$ in the matrix $A = (a_{i,j}) \in \widehat{W}$. Since the other rows of A are obtained by permuting the entries in the first row we get the conclusion. $\square$

The boundary cases with regard to the dimension of $\widehat{W}$ are of particular interest. We discuss the case when $\dim \widehat{W} = \dim W$ in the subsequent section, see 3.8.

At this point we note the following general observation which follows by our construction.

PROPOSITION 2.11. *Let $(\mathcal{T}, W, \widehat{W})$ be an isotropic phylogenetic tree such that $(W, \widehat{W})$ has a group of symmetries $G \leq S_d$. Then the parametrization maps Ψ and Φ are G equivariant. In particular*

$$X(\mathcal{T}, W, \widehat{W}) \subset \mathbb{P}\left(\left(\bigotimes_{v \in \mathcal{L}} W_v\right)^G\right) \subset \mathbb{P}\left(\bigotimes_{v \in \mathcal{L}} W_v\right).$$

Let us consider a subset of letters $\mathcal{B} \subset \mathcal{A}$ which spans a vector subspace $W_{\mathcal{B}} \subset W$. We have a decomposition $W = W_{\mathcal{B}} \oplus W_{\mathcal{B}}^{\perp}$ where $W_{\mathcal{B}}^{\perp}$ is spanned on the complement of $\mathcal{B}$, that is $\mathcal{A} \setminus \mathcal{B}$. We fix the projection $\pi_{\mathcal{B}} : W \longrightarrow W_{\mathcal{B}}$ the kernel of which is $W_{\mathcal{B}}^{\perp}$. This projection extends to $S^2\pi_{\mathcal{B}} : S^2W \longrightarrow S^2W_{\mathcal{B}}$. For $G \leq S_d$ take $G_{\mathcal{B}} = \{g \in G : g(W_{\mathcal{B}}) \subset W_{\mathcal{B}}\}$. Elements of $G_{\mathcal{B}}$ define symmetries of $W_{\mathcal{B}}$ with $\widehat{W}_{\mathcal{B}} = Fix(S^2\rho_{G_{\mathcal{B}}})$.

EXAMPLE 2.12. Let $g \in G \leq S_d$ be an element whose decomposition into cycles contains a cycle of length r. We may assume that the cycle concerns the first r letters, more precisely that $g = (1, \dots, r) \cdots$. We take $\mathcal{B} = \{\alpha_1, \dots, \alpha_r\}$. In terms of symmetric matrices the projection $S^2\pi_{\mathcal{B}}$ is taking the $r \times r$ upper-left corner from the $d \times d$ matrix $A \in S^2W$. In view of example 2.4 this implies constraints on the coefficients of matrices in $\widehat{W}$. Namely, for $A = (a_{i,j}) \in \widehat{W}$ we have the following constraints $a_{1,2} = a_{1,r}$, $a_{1,3} = a_{1,r-1}$, etc.

$\circ$

DEFINITION 2.13. *In the above situation we say that the subset $\mathcal{B} \subset \mathcal{A}$, or the subspace $W_{\mathcal{B}} \subset W$, is faithful if $S^2\pi$ determines an isomorphism of $\widehat{W}$ and $\widehat{W}_{\mathcal{B}}$. We say that the model of evolution $(W, \widehat{W})$ with group of symmetries G is minimal if $\mathcal{A}$ contains no proper faithful subset.*

EXAMPLE 2.14. The full symmetric group $G = S_d$ is clearly saturated, its symmetric model of evolution will be called the *Jukes-Cantor model* (we note that originally, in statistics, this name refers only to the case when $d = 4$). Then, for any

$d > 1$ any subset of $\{\alpha_1, \dots, \alpha_d\}$ consisting of more than 1 letter yields a faithful inclusion.

$\circ$

3. Hyperbinary model of evolution

Construction 3.1. Let us consider a hyperbinary group $\mathbb{B}^n = (\mathbb{Z}_2)^n$ for which we use additive notation (so its elements are 0/1 sequences of length n). It is well known that any finite group whose elements (except the unit) are of order 2 is isomorphic to some $\mathbb{B}^n$.

Define a representation $\rho^n : \mathbb{B}^n \longrightarrow GL(\mathbb{C}^{2^n})$ by induction with respect to n. For $n = 0$ we set $\rho^0 = 1$. Suppose that ρ^n is defined. Let us decompose $\mathbb{B}^{n+1} = \mathbb{B}^n \oplus \mathbb{Z}_2 \cdot e_{n+1}$ with $\mathbb{B}^n < \mathbb{B}^{n+1}$ consisting of these elements whose last coordinate is 0 and $e_{n+1} = (0, \dots, 0, 1)$. For the subset $\mathbb{B}^n < \mathbb{B}^{n+1}$ we set $\rho^{n+1}_{|\mathbb{B}^n} = \rho^n \oplus \rho^n$ and in addition

$$\rho^{n+1}(e_{n+1}) = \begin{bmatrix} 0 & I^{2^n} \\ I^{2^n} & 0 \end{bmatrix}$$

where I denotes the identity matrix of the respective dimension.

Before discussing the properties of the hyperbinary representation let us recall the following general property of representations of abelian groups. Since our abelian groups will usually be subgroups of multiplicative groups of matrices we prefer to use multiplicative notation to state this result.

Lemma 3.2. *Let G be an abelian group which acts effectively (i.e. no non-unit element of G acts trivially) and transitively on a finite set X. Then the action is equivalent to the action of G on itself. That is, choosing an element in X and identifying it with the unit of G gives a bijection $G \longrightarrow X$ such that the action $G \times X \longrightarrow X$ is identified with multiplication $G \times G \longrightarrow G$.*

The complex representation arising from such an action of G will be called the *regular representation* of G. We get an immediate corollary.

Lemma 3.3. *Let $\mathcal{A}$ be the standard basis of $W^n_{\mathbb{B}} := \mathbb{C}^{2^n}$. Then $\mathbb{B}^n$ acts effectively and transitively on $\mathcal{A}$ and thus the representation ρ^n is equivalent to the regular representation of $\mathbb{B}^n$.*

We identify, via $\rho^n_{|\mathcal{A}}$, the group $\mathbb{B}^n$ with a subgroup of S_{2^n}; that is, ρ^n is then just the restriction of the natural representation of S_{2^n}. Note that all matrices in $\rho^n(\mathbb{B}^n)$ are symmetric. This is a very special feature of $\mathbb{B}^n$ as it follows from the subsequent observation.

Lemma 3.4. *Let W be an arbitrary vector space. All matrices (except identity) in the intersection $S^2(W) \cap O(W)$ are of order 2, hence any finite subgroup of $S^2(W) \cap O(W)$ is hyperbinary.*

Let $\widehat{W}^n_{\mathbb{B}} \subset S^2(W^n_{\mathbb{B}})$ be the linear subspace spanned by $\rho^n(\mathbb{B}^n)$. The following lemma will be generalized in the last section of the paper. For the sake of clarity here we present an explicit argument.

Lemma 3.5. *The space $\widehat{W}^n_{\mathbb{B}}$ is equal to $Fix(\mathbb{B}^n)$, it is of dimension 2^n and its intersection with $GL(W)$ is a Cartan torus in $GL(W^n_{\mathbb{B}})$.*

PROOF. The fact that the space $\widehat{W}^n_{\mathbb{B}}$ is of dimension 2^n follows from a general lemma 5.1. Here, however, we note easily by induction on n that the matrices in $\rho^n(\mathbb{B}^n)$ are linearly independent. Indeed, since $\mathbb{B}^{n+1} = \mathbb{B}^n + e_{n+1} \cdot \mathbb{B}^n$ then every linear combination of matrices in $\rho_{n+1}(\mathbb{B}^{n+1})$ can be written as

$$A = \sum_{A_i \in \mathbb{B}^n} a_i \begin{bmatrix} A_i & 0 \\ 0 & A_i \end{bmatrix} + \sum_{B_i \in \mathbb{B}^n} b_i \begin{bmatrix} 0 & B_i \\ B_i & 0 \end{bmatrix}$$

which yields the inductive step.

Next, we note that $\widehat{W}^n_{\mathbb{B}} \subset Fix(\mathbb{B}^n)$. For this we are to check that $g \cdot A \cdot g^{-1} = A$ for every $g \in \rho^n(\mathbb{B}^n)$ and $A \in \widehat{W}^n_{\mathbb{B}}$. But this equality is linear with respect to A so it is enough to check it on the basis of $\widehat{W}^n_{\mathbb{B}}$ consisting of elements of $\rho^n(\mathbb{B}^n)$ for which this is obvious.

By the same argument $\widehat{W}^n_{\mathbb{B}} \cap GL(W^n_{\mathbb{B}})$ is commutative hence it is a complex torus. Its dimension is 2^n, hence it is a Cartan subgroup of $GL(W^n_{\mathbb{B}})$.

Finally we prove the equality $Fix(\rho^n(\mathbb{B}^n)) = \widehat{W}^n_{\mathbb{B}}$. One inclusion is already proved so suppose that $A \in S^2(W^n_{\mathbb{B}})$ is such that $g \cdot A \cdot g^{-1} = A$ for every $g \in \mathbb{B}$. We may assume that A is invertible. Thus the subgroup of $GL(\mathbb{B}^n)$ generated by $\rho^n(\mathbb{B}^n)$ and A is abelian and thus contained in a Cartan subgroup of $GL(W^n_{\mathbb{B}})$ which must be $\widehat{W}^n_{\mathbb{B}} \cap GL(W^n_{\mathbb{B}})$. □

LEMMA 3.6. *The group $\mathbb{B}^n < S_{2^n}$ is saturated.*

PROOF. Suppose that $h \in S_{2^n}$ preserves $\widehat{W}^n_{\mathbb{B}}$. Then, in particular, $h \cdot g \cdot h^{-1} = g$ for every $g \in \mathbb{B} < S_{2^n}$. Thus subgroup H generated by h and $\mathbb{B}^n$ is abelian. Since it acts effectively on a set of 2^n letters we get $|H| \leq 2^n$ and thus $H = \mathbb{B}^n$. □

We summarize the above results in the following.

PROPOSITION 3.7. *The pair $(W^n_{\mathbb{B}}, \widehat{W}^n_{\mathbb{B}})$ defined above is a symmetric model of evolution with group of symmetries $\mathbb{B}^n$.*

Phylogenetic trees with hyperbinary model of evolution will be called just *hyperbinary phylogentic trees*. In the above situation, if $n = 1$ then such a model is just a binary model and if $n = 2$ then it is a Kimura 3-parameter model, see 1.6.

The hyperbinary model of evolution is the unique one which admits the biggest dimension, recall 2.9.

PROPOSITION 3.8. *Let $(W, \widehat{W})$ be a symmetric model of evolution with group of symmetries G such that $\dim \widehat{W} = \dim W$. Then, up to renumbering elements of $\mathcal{A}$ (i.e. up to conjugation in the group of permutations), this model coincides with the hyperbinary model of evolution.*

PROOF. In view of the proof of 2.9, because of the discussion in the example 2.12, we see that G contains only elements of order 2. Thus G is hyperbinary and the rest follows by 3.2. □

Thus we get a corollary.

COROLLARY 3.9. *The hyperbinary model of evolution is minimal.*

And finally we prove the following theorem which follows from somewhat more general results proved in section 5.

THEOREM 3.10. *The hyperbinary group is the only abelian saturated group.*

PROOF. If G is an abelian saturated group then it satisfies assumptions of 5.3 thus it contains an element ν of order 2 such that for every $g \in G$ we have $\nu \cdot g \cdot \nu = g^{-1}$. Thus every element of G is of order 2 and G is hyperbinary. □

It follows that the hyperbinary model is the only isotropic *group-based model*, in the sense of [**SS03**, Sect. 8] or [**PS05**] and the references therein.

4. Low dimensional models of evolution

In order to have a nontrivial set of examples we determined all transitive saturated subgroups of S_d for $d = \dim W \leq 9$. Computations were done by simple functions in GAP (see [**GAP**]), which were effective only for $d \leq 9$. The GAP code of our program can be found at www.mimuw.edu.pl/˜marysia/isotrees. This section gives a brief description of the results of computations.

First, let us give the main ideas of the algorithm determining saturated subgroups of S_d. There are two parts of the algorithm:

- find $\widehat{W}_G = Fix(S^2\rho_G)$ for all $G < S_d$,
- for each $G < S_d$, decide whether it is maximal subgroup fixing $\widehat{W}_G$.

We consider only representatives of conjugacy classes of subgroups, because models of evolution for conjugate groups are essentially the same, see lemma 4.1.

The second part of the algorithm is based on functions provided by GAP. The most important is `LatticeSubgroups`, which returns the lattice of subgroups of S_d, that is the set of all subgroups and the relation of inclusion on this set (up to conjugacy). We also use `MinimalSupergroupsLattice` which, given the lattice of subgroups, calculates all minimal proper supergroup for each subgroup. Using these functions we check whether the group G is saturated by comparing $\widehat{W}_G$ to $\widehat{W}_H$ for all minimal proper supergroups H of G. However, the function `LatticeSubgroups` is not effective enough to be used in the cases $d \geq 10$. We think that this step can be done more effectively by more subtle algorithms and low-level programming.

We now turn to the first part of the algorithm, that is, to the question of determining $\widehat{W}_G$ for given $G < S_d$. Recall that a matrix $(a_{i,j})$ is a fixed point of $S^2\rho_G$ if and only if for each $g \in G$ we have $a_{i,j} = a_{g(i),g(j)}$, so the task is to find the sets of equal matrix entries. Obviously it suffices to consider only the equalities of entries for g in a generating set of G. For generators we choose the result of the GAP function `SmallGeneratingSet`. It returns a generating set which is small, not necessarily minimal, but the function is much faster than the function which computes a minimal set of generators. An important step is to find an appropriate data structure for storing information about equal matrix entries. This algorithm is implemented in the function `ModifySymmetricMatrix`. Probably much more effective algorithms to this task can be found, but our idea gives a solution which is short and easy to implement.

The program contains also a few functions which check various properties of saturated groups and models of evolution. These use similar data structures and ideas as `ModifySymmetricMatrix`, described above. For example, the function `CompareSaturatedSubgroups` tests the inclusion $\widehat{W}_G \subset \widehat{W}_H$ for given saturated subgroups $G, H < S_d$, so it solves the main problem in creating diagrams included in this section.

As stated at the beginning of this section, we identify models of evolution which differ only by a permutation of letters, i.e. models determined by conjugate subgroups of S_d. Therefore, we can reformulate lemma 2.5 as follows.

LEMMA 4.1. *There is an inclusion reversing bijection between conjugacy classes of saturated transitive subgroups of S_d and (isomorphism classes of) isotropic symmetric models of evolution on d letters.*

EXAMPLE 4.2. There are 3 possible forms of the model of evolution with $d = 4$ and dihedral group of symmetries D_8 (*Kimura 2-parameter model*). They are associated to the three conjugate subgroups of S_4 hence the choice of a cyclic permutation of length 4, c.f. example 2.4.

$$\begin{bmatrix} a & b & c & b \\ b & a & b & c \\ c & b & a & b \\ b & c & b & a \end{bmatrix} \qquad \begin{bmatrix} a & c & b & b \\ c & a & b & b \\ b & b & a & c \\ b & b & c & a \end{bmatrix} \qquad \begin{bmatrix} a & b & b & c \\ b & a & c & b \\ b & c & a & b \\ c & b & b & a \end{bmatrix}$$

○

The relation between models of evolution described by lemma 4.1, i.e. inclusion of their groups of symmetries up to conjugation, is presented in diagrams included in this section. For each conjugacy class of saturated groups a generating set of a chosen representative is given. We also find minimal models for all examples in the sense of definition 2.13.

As noted in example 1.6, the only model of evolution for $d = \dim W = 2$ is binary symmetric (or the Jukes-Cantor model, or the Cavender-Farris-Neyman model, as it is called in statistics). Also for $d = 3$ the symmetric group S_3 is the only saturated group and its model is the Jukes-Cantor model.

The smallest nontrivial example is $d = 4$. We tackled this case already in 1.6. Note however that the isotropic strand-symmetric model is not symmetric in the sense of our definition 2.6 because its group of symmetries $\mathbb{Z}_2$ does not act transitively on the set of letters.

$$\begin{pmatrix} a & b & c & d \\ b & a & d & c \\ c & d & a & b \\ d & c & b & a \end{pmatrix} \longrightarrow \begin{pmatrix} a & b & c & c \\ b & a & c & c \\ c & c & a & b \\ c & c & b & a \end{pmatrix} \longrightarrow \begin{pmatrix} a & b & b & b \\ b & a & b & b \\ b & b & a & b \\ b & b & b & a \end{pmatrix}$$

FIGURE 1. Hierarchy of models of evolution for $d = 4$

The first of the models in figure 1 is a hyperbinary model, or *Kimura 3-parameter model*, which is minimal. The second, *Kimura 2-parameter model*, must be minimal as well, because there are no models of evolution with $d < 4$ and $\dim \widehat{W} = 3$. Generators of chosen representatives are given in the table below where we also indicate the isomorphism type of the group in question. The last entry in each row indicates if the model is minimal, if it is not then the name of a minimal submodel is provided (J–C stands for the Jukes-Cantor model).

group	**type**	**generators**	**model**
$g4\_1$	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$(1,3)(2,4)$, $(1,4)(2,3)$	$\mathbb{B}^2$
$g4\_2$	D_8	$(3,4)$, $(1,3)(2,4)$	min
$g4\_3$	S_4	$(1,2,3,4)$, $(1,2)$	J–C

Dimensions $d = 5$ and $d = 7$ are not very interesting. In each of these cases there are only two models of evolution, one of them being the Jukes-Cantor model associated to the full symmetric group. The other model of evolution in each case is minimal and it is the *dihedral model* described in example 2.4.

$$\begin{tikzcd}
& \begin{pmatrix} a & b & b & c & d & e \\ b & a & b & d & e & c \\ b & b & a & e & c & d \\ c & d & e & a & b & b \\ d & e & c & b & a & b \\ e & c & d & b & b & a \end{pmatrix} \; g6_1 \arrow[d] & \\
& \begin{pmatrix} a & b & b & c & c & d \\ b & a & b & c & d & c \\ b & b & a & d & c & c \\ c & c & d & a & b & b \\ c & d & c & b & a & b \\ d & c & c & b & b & a \end{pmatrix} \; g6_2 \arrow[dl] \arrow[dr] & \\
\underset{g6_4}{\begin{pmatrix} a & b & b & c & c & c \\ b & a & b & c & c & c \\ b & b & a & c & c & c \\ c & c & c & a & b & b \\ c & c & c & b & a & b \\ c & c & c & b & b & a \end{pmatrix}} \arrow[dr] & & \underset{g6_3}{\begin{pmatrix} a & b & b & b & b & c \\ b & a & b & b & c & b \\ b & b & a & c & b & b \\ b & b & c & a & b & b \\ b & c & b & b & a & b \\ c & b & b & b & b & a \end{pmatrix}} \arrow[dl] \\
& \begin{pmatrix} a & b & b & b & b & b \\ b & a & b & b & b & b \\ b & b & a & b & b & b \\ b & b & b & a & b & b \\ b & b & b & b & a & b \\ b & b & b & b & b & a \end{pmatrix} \; g6_5 &
\end{tikzcd}$$

FIGURE 2. Hierarchy of models of evolution for $d = 6$

Figure 2 presents the case of $d = 6$. Only the model of $g6\_1$ is minimal. For the remaining models we can find faithful subspaces of dimension 2, in the case of $g6\_5$, or 4. The set $\{\alpha_2, \alpha_3, \alpha_4, \alpha_5\}$ is faithful in cases of $g6\_2$, $g6\_3$ and $g6\_4$. We give examples of generators of saturated subgroups:

group	**type**	**generators**	**model**
$g6\_1$	D_6	$(1,3,2)(4,5,6)$, $(1,6)(2,4)(3,5)$	min
$g6\_2$	D_{12}	$(1,5,3,6,2,4)$, $(1,2)(5,6)$	$g4\_1$
$g6\_3$	$\mathbb{Z}_2 \times S_4$	$1,4)(3,6)$, $(1,3,6,4)$, $(1,5)(2,6)(3,4)$	$g4\_2$
$g6\_4$	$(S_3 \times S_3) \rtimes \mathbb{Z}_2$	$(1,2,3)(5,6)$, $(1,5,3,4)(2,6)$	$g4\_2$
$g6\_5$	S_6	$(1,2)$, $(1,2,3,4,5,6)$	J–C

$$\begin{pmatrix} a & b & c & d & e & f & g & h \\ b & a & d & c & f & e & h & g \\ c & d & a & b & g & h & e & f \\ d & c & b & a & h & g & f & e \\ e & f & g & h & a & b & c & d \\ f & e & h & g & b & a & d & c \\ g & h & e & f & c & d & a & b \\ h & g & f & e & d & c & b & a \end{pmatrix} \quad g8\_1$$

$$\begin{pmatrix} a & b & c & d & e & f & g & g \\ b & a & d & c & f & e & g & g \\ c & d & a & b & g & g & e & f \\ d & c & b & a & g & g & f & e \\ e & f & g & g & a & b & d & c \\ f & e & g & g & b & a & c & d \\ g & g & e & f & d & c & a & b \\ g & g & f & e & c & d & b & a \end{pmatrix} \quad g8\_2$$

$$\begin{pmatrix} a & b & c & d & e & e & f & f \\ b & a & d & c & e & e & f & f \\ c & d & a & b & f & f & e & e \\ d & c & b & a & f & f & e & e \\ e & e & f & f & a & b & c & d \\ e & e & f & f & b & a & d & c \\ f & f & e & e & c & d & a & b \\ f & f & e & e & d & c & b & a \end{pmatrix} \quad g8\_3$$

$$\begin{pmatrix} a & b & c & c & d & d & e & e \\ b & a & c & c & d & d & e & e \\ c & c & a & b & e & e & d & d \\ c & c & b & a & e & e & d & d \\ d & d & e & e & a & b & c & c \\ d & d & e & e & b & a & c & c \\ e & e & d & d & c & c & a & b \\ e & e & d & d & c & c & b & a \end{pmatrix} \quad g8\_7$$

$$\begin{pmatrix} a & b & c & d & e & e & e & e \\ b & a & d & c & e & e & e & e \\ c & d & a & b & e & e & e & e \\ d & c & b & a & e & e & e & e \\ e & e & e & e & a & b & c & d \\ e & e & e & e & b & a & d & c \\ e & e & e & e & c & d & a & b \\ e & e & e & e & d & c & b & a \end{pmatrix} \quad g8\_5$$

$$\begin{pmatrix} a & b & c & c & d & d & e & e \\ b & a & c & c & e & e & d & d \\ c & c & a & b & d & e & d & e \\ c & c & b & a & e & d & e & d \\ d & e & d & e & a & c & c & b \\ d & e & e & d & c & a & b & c \\ e & d & d & e & c & b & a & c \\ e & d & e & d & b & c & c & a \end{pmatrix} \quad g8\_4$$

$$\begin{pmatrix} a & b & b & b & c & d & d & d \\ b & a & b & b & d & c & d & d \\ b & b & a & b & d & d & c & d \\ b & b & b & a & d & d & d & c \\ c & d & d & d & a & b & b & b \\ d & c & d & d & b & a & b & b \\ d & d & c & d & b & b & a & b \\ d & d & d & c & b & b & b & a \end{pmatrix} \quad g8\_6$$

$$\begin{pmatrix} a & b & c & c & d & d & d & d \\ b & a & c & c & d & d & d & d \\ c & c & a & b & d & d & d & d \\ c & c & b & a & d & d & d & d \\ d & d & d & d & a & b & c & c \\ d & d & d & d & b & a & c & c \\ d & d & d & d & c & c & a & b \\ d & d & d & d & c & c & b & a \end{pmatrix} \quad g8\_8$$

$$\begin{pmatrix} a & b & c & c & c & c & c & c \\ b & a & c & c & c & c & c & c \\ c & c & a & b & c & c & c & c \\ c & c & b & a & c & c & c & c \\ c & c & c & c & a & b & c & c \\ c & c & c & c & b & a & c & c \\ c & c & c & c & c & c & a & b \\ c & c & c & c & c & c & b & a \end{pmatrix} \quad g8\_9$$

$$\begin{pmatrix} a & b & b & b & c & c & c & c \\ b & a & b & b & c & c & c & c \\ b & b & a & b & c & c & c & c \\ b & b & b & a & c & c & c & c \\ c & c & c & c & a & b & b & b \\ c & c & c & c & b & a & b & b \\ c & c & c & c & b & b & a & b \\ c & c & c & c & b & b & b & a \end{pmatrix} \quad g8\_10$$

$$\begin{pmatrix} a & b & b & b & b & b & b & b \\ b & a & b & b & b & b & b & b \\ b & b & a & b & b & b & b & b \\ b & b & b & a & b & b & b & b \\ b & b & b & b & a & b & b & b \\ b & b & b & b & b & a & b & b \\ b & b & b & b & b & b & a & b \\ b & b & b & b & b & b & b & a \end{pmatrix} \quad g8\_11$$

FIGURE 3. Hierarchy of models of evolution for $d = 8$

In dimension $d = 8$ the relation between models of evolution (see figure 3) is much more complex than in previous examples (in the subsequent table we skip the description of the isomorphism type of the group in question if it is too complicated). It can be seen from the following table that only 4 of 11 models of evolution are not minimal. Thus the situation is much different from the cases $d = 6$ (one minimal model) and $d = 9$ (no minimal models).

group	**type**	**generators**	**model**
$g8\_1$	$\mathbb{Z}_2^3$	$(1,2)(3,4)(5,6)(7,8),\ (1,3)(2,4)(5,7)(6,8),$ $(1,5)(2,6)(3,7)(4,8)$	$\mathbb{B}^3$
$g8\_2$	D_8	$(1,4)(2,3)(5,8)(6,7),\ (1,7,2,8)(3,6,4,5)$	min
$g8\_3$	$\mathbb{Z}_2 \times D_8$	$(5,6)(7,8),\ (1,3)(2,4)(5,7)(6,8),$ $(1,5)(2,6)(3,7)(4,8)$	min
$g8\_4$	D_{16}	$(1,2)(5,7)(6,8),\ (1,8)(2,5)(3,7)(4,6)$	min
$g8\_5$	$\mathbb{Z}_2^4 \rtimes \mathbb{Z}_2$	$(1,8,4,5)(2,7,3,6),\ (1,8,3,6)(2,7,4,5),$ $(1,8)(2,7)(3,6)(4,5)$	min
$g8\_6$	$\mathbb{Z}_2 \times S_4$	$(1,3,4,2)(5,7,8,6),\ (1,7,2,8)(3,6,4,5)$	$g4\_1$
$g8\_7$	—	$(1,5,2,6)(3,7,4,8),\ (1,5)(2,6)(3,8,4,7),$ $(1,7)(2,8)(3,5,4,6)$	min
$g8\_8$	$(D_8 \times D_8) \rtimes \mathbb{Z}_2$	$(1,4,2,3)(5,8)(6,7),\ (1,5)(2,6)(3,7,4,8),$ $(1,8,4,6)(2,7,3,5)$	min
$g8\_9$	—	$(1,2)(3,8,4,7)(5,6),\ (1,5,8,3)(2,6,7,4)$	$g4\_2$
$g8\_10$	$(S_4 \times S_4) \rtimes \mathbb{Z}_2$	$(1,4,3,2)(5,8)(6,7),\ (1,7,2,5,3,6)(4,8)$	$g4\_2$
$g8\_11$	S_8	$(1,2),\ (1,2,3,4,5,6,7,8)$	J–C

In case of $d = 9$ there are 6 different models of evolution, presented at figure 4. It turns out that there are no minimal models of evolution. For all models we can find faithful subspaces of dimensions 6, 4 or 2. In all cases a 6-dimensional faithful subspace is spanned by the set of the first 6 basis vectors. For $g9\_3$, $g9\_4$ and $g9\_5$ there are 4-dimensional faithful subspaces contained in the subspace spanned by the first 6 letters. We give examples of generating sets of saturated groups.

group	**type**	**generators**	**model**
$g9\_1$	$(\mathbb{Z}_3 \times \mathbb{Z}_3) \rtimes \mathbb{Z}_2$	$(2,3)(4,7)(5,9)(6,8),\ (1,2,3)(4,5,6)(7,8,9),$ $(1,4,7)(2,5,8)(3,6,9)$	$g6\_1$
$g9\_2$	D_{18}	$(1,6)(2,5)(3,4)(7,8),\ (1,7)(2,9)(3,8)(5,6)$	$g6\_1$
$g9\_3$	$S_3 \times S_3$	$(1,2)(4,5)(7,8),\ (1,2)(4,8)(5,7)(6,9),$ $(1,8,3,7,2,9)(4,5,6)$	$g4\_1$
$g9\_4$	$(S_3 \times S_3) \rtimes \mathbb{Z}_2$	$(1,5)(3,8)(6,7),\ (1,7,8,2)(3,4,9,5)$	$g4\_2$
$g9\_5$	—	$(4,5),\ (1,5,3,4)(2,6)(7,8,9),$ $(1,7,3,9)(2,8)(4,5,6)$	$g4\_2$
$g9\_6$	S_9	$(1,2),\ (1,2,3,4,5,6,7,8,9)$	J–C

These low-dimensional examples suggest that there are more models of evolution (or classes of saturated subgroups) in even dimensions than in odd dimensions.

$$\begin{pmatrix} a & b & b & c & d & e & c & e & d \\ b & a & b & e & c & d & d & c & e \\ b & b & a & d & e & c & e & d & c \\ c & e & d & a & b & b & c & d & e \\ d & c & e & b & a & b & e & c & d \\ e & d & c & b & b & a & d & e & c \\ c & d & e & c & e & d & a & b & b \\ e & c & d & d & c & e & b & a & b \\ d & e & c & e & d & c & b & b & a \end{pmatrix} g9\_1 \qquad \begin{pmatrix} a & b & b & c & d & e & e & d & c \\ b & a & b & e & c & d & c & e & d \\ b & b & a & d & e & c & d & c & e \\ c & e & d & a & b & b & c & d & e \\ d & c & e & b & a & b & e & c & d \\ e & d & c & b & b & a & d & e & c \\ e & c & d & c & e & d & a & b & b \\ d & e & c & d & c & e & b & a & b \\ c & d & e & e & d & c & b & b & a \end{pmatrix} g9\_2$$

$$\begin{pmatrix} a & b & b & c & d & d & c & d & d \\ b & a & b & d & c & d & d & c & d \\ b & b & a & d & d & c & d & d & c \\ c & d & d & a & b & b & c & d & d \\ d & c & d & b & a & b & d & c & d \\ d & d & c & b & b & a & d & d & c \\ c & d & d & c & d & d & a & b & b \\ d & c & d & d & c & d & b & a & b \\ d & d & c & d & d & c & b & b & a \end{pmatrix} g9\_3$$

$$\begin{pmatrix} a & b & b & b & c & c & b & c & c \\ b & a & b & c & b & c & c & b & c \\ b & b & a & c & c & b & c & c & b \\ b & c & c & a & b & b & b & c & c \\ c & b & c & b & a & b & c & b & c \\ c & c & b & b & b & a & c & c & b \\ b & c & c & b & c & c & a & b & b \\ c & b & c & c & b & c & b & a & b \\ c & c & b & c & c & b & b & b & a \end{pmatrix} g9\_4 \qquad \begin{pmatrix} a & b & b & c & c & c & c & c & c \\ b & a & b & c & c & c & c & c & c \\ b & b & a & c & c & c & c & c & c \\ c & c & c & a & b & b & c & c & c \\ c & c & c & b & a & b & c & c & c \\ c & c & c & b & b & a & c & c & c \\ c & c & c & c & c & c & a & b & b \\ c & c & c & c & c & c & b & a & b \\ c & c & c & c & c & c & b & b & a \end{pmatrix} g9\_5$$

$$\begin{pmatrix} a & b & b & b & b & b & b & b & b \\ b & a & b & b & b & b & b & b & b \\ b & b & a & b & b & b & b & b & b \\ b & b & b & a & b & b & b & b & b \\ b & b & b & b & a & b & b & b & b \\ b & b & b & b & b & a & b & b & b \\ b & b & b & b & b & b & a & b & b \\ b & b & b & b & b & b & b & a & b \\ b & b & b & b & b & b & b & b & a \end{pmatrix} g9\_6$$

FIGURE 4. Hierarchy of models of evolution for $d = 9$

It is also possible that minimal models appear more often in even dimensions, and most often in dimensions $d = 2^k$. Our examples also yield an observation regarding the family of hyperbinary models: up to dimension 9 there are no other abelian saturated groups.

5. Abelian groups of symmetries

The present section concerns the case when the group of symmetries of an isotropic model contains an abelian subgroup acting transitively on the set of letters. Let us begin by recalling trivialities regarding actions of abelian groups. The set $\mathcal{A}$, as usual, consists of d letters. Let H be an abelian group acting effectively and transitively on the set $\mathcal{A}$, which yields the regular representation of H on the vector space W spanned by $\mathcal{A}$.

LEMMA 5.1. *The regular representation $\rho_H : H \longrightarrow GL(W)$ can be diagonalized in terms of characters of H. That is, ρ_H is equivalent to $\rho_H^\chi : H \longrightarrow GL(W)$ such that for every $h \in H$ it holds $\rho_H^\chi(h) = Diag(\chi_i(h))$, where $Diag$ stands for a diagonal matrix and χ_i runs over all different characters in the dual group $\widehat{H} = Hom(H, \mathbb{C}^*)$.*

In this situation, we identify H with a subgroup of $GL(W)$ and argue similarly as in 3.5. Let us consider a linear span $W_H = \sum_{h \in H} \mathbb{C} \cdot h \subset End(W)$. Then $dim W_H = d$ (because characters are linearly independent) and H acts by multiplications on W_H as the regular representation. Letus set $T_H := W_H \cap GL(W)$. Then T_H is a connected abelian algebraic subgroup of $GL(W)$ of dimension d, hence a Cartan torus in $GL(W)$. Thus W_H is the fixed point set of the adjoint action of H on $End(W)$. By the above lemma, the lattice of (algebraic) characters of T_H, $M_H = Hom(T_H, \mathbb{C}^*)$, has a distinguised basis consisting of characters of H, that is $\widehat{H} = Hom(H, \mathbb{C}^*)$.

Now we turn to the situation which is our principal interest.

SET-UP 5.2. Let $(W, \widehat{W})$ be an isotropic symmetric model of evolution on the set $\mathcal{A}$ of d letters with the (saturated) group of symmetries $G \leq S_d$. Assume that there exists $H \leq G$, an abelian subgroup which acts transitively on $\mathcal{A}$.

The following result generalizes our observation from example 2.4.

PROPOSITION 5.3. *In the situation of 5.2 there exists an involution $\nu \in G$, $\nu^2 = id$, such that for every $h \in H$ it holds $\nu \cdot h \cdot \nu = h^{-1}$.*

The proof of the above proposition is divided into some steps. Elements of H can be identified with the letters, so that the action of H on $\mathcal{A}$ is equivalent to the action of H on itself. From now on we use this identification. Also, we will use additive notation for H, as for an abelian group, while for permutations and matrices (the group G and H treated as its subgroup) we will use multiplicative notation.

By classification of finite groups we can write H as a product of cyclic groups, that is $H = \mathbb{Z}_{p_1} \times \ldots \times \mathbb{Z}_{p_k}$, for suitable choice of numbers p_i. We have an inclusion of *sets* $H \subset \mathbb{Z}^k$, coming from the natural inclusion $\mathbb{Z}_{p_i} = \{0, 1, \ldots, p_i - 1\} \subset \mathbb{Z}$. This leads to a linear order on H, which is the restriction of the lexicographical order on $\mathbb{Z}^k$ to H. That is, $h_i = (i_1, \ldots, i_k)$ is the i-th element of H with respect to this order, if

$$i = i_1 \cdot p_2 \cdots p_k + i_2 \cdot p_3 \cdots p_k + \cdots + i_k.$$

Now we can write $H = \{h_0, \ldots, h_{d-1}\}$.

We use the map $\nu : \mathbb{Z}^k \to \mathbb{Z}^k$, defined by

$$\nu((i_1, \ldots, i_k)) = (p_1 - 1 - i_1, \ldots, p_k - 1 - i_k).$$

Note that ν is an involution, $\nu^2 = id$, and it can be restricted to the subset H. The restriction, also denoted by ν, is a permutation of elements of H (hence of $\mathcal{A}$) such that $\nu \cdot h \cdot \nu = h^{-1}$ for every $h \in H$. Indeed, in terms of operations within H we have the following indentities

$$\nu((i_1, \ldots, i_k)) = (-i_1 - 1, \ldots, -i_k - 1) = -(i_1, \ldots, i_k) - (1, \ldots, 1)$$

from which we get $\nu \cdot h \cdot \nu = h^{-1}$. Similarly we note that for every $h \in H$ we have $h + \nu(h) = -(1, 1, \dots, 1)$.

It turns out that the involution ν is compatible with the chosen order on H in the sense of the following lemma.

LEMMA 5.4. *The map ν satisfies $\nu(h_i) = h_{d-1-i}$.*

PROOF. Let $h_i = (i_1, \dots, i_k)$. Then, by the definition of the chosen order on H, $i = \sum_{m=1}^{k} i_m \cdot p_{m+1} \cdots p_k$. Let $h_j = \nu(h_i) = (p_1 - 1 - i_1, \dots, p_k - 1 - i_k)$, then

$$j = (p_1 - 1 - i_1) \cdot p_2 \cdots p_k + (p_2 - 1 - i_2) \cdot p_3 \cdots p_k + \ldots + p_k - 1 - i_k =$$
$$= d - 1 - \sum_{m=1}^{k} i_m \cdot p_{m+1} \cdots p_k = d - 1 - i.$$

□

Now the proposition 5.3 is obtained by the following lemma.

LEMMA 5.5. *Let $H = \{h_0, \dots, h_{d-1}\}$ be as above, with its regular represention (on the set $\mathcal{A} = H$ with the order defined above) denoted by ρ_H. Let $A = (a_{i,j})$ be a symmetric matrix fixed by the induced action $S^2\rho_H$. Then A is also fixed by the involution ν.*

PROOF. To prove the lemma we need to show that for any $i, j \in \{0, \dots d-1\}$ we have $a_{i,j} = a_{\nu(i),\nu(j)}$. Because A is symmetric we have $a_{i,j} = a_{j,i}$ so this is equivalent to proving $a_{j,i} = a_{\nu(i),\nu(j)}$. However, we noted that $h_i + \nu(h_i) = h_j + \nu(h_j) = -(1, \dots, 1)$ and thus we can take $h \in H$ such that $h = \nu(h_i) - h_j = \nu(h_j) - h_i$ (the $\pm$ operations are in group H). This implies that $h + h_j = \nu(h_i)$ and $h + h_i = \nu(h_j)$ hence, in terms of the action of H on itself, $h(j) = \nu(i)$ and $h(i) = \nu(j)$. Thus, because A is fixed by H, we get

$$a_{j,i} = a_{h(j),h(i)} = a_{\nu(i),\nu(j)}$$

□

The above argument can be reversed.

LEMMA 5.6. *Let H and ρ_H be as above. Suppose that a matrix $A = (a_{i,j}) \in End(W)$ is fixed by Ad_H and $Ad(\nu)$. Then A is symmetric.*

PROOF. If A is fixed by H then, as above, $a_{j,i} = a_{\nu(i),\nu(j)}$ and since it is fixed by ν it follows $a_{j,i} = a_{i,j}$. □

As a result we get the following.

PROPOSITION 5.7. *Assume that $(W, \widehat{W})$ is a symmetric model of evolution with G, the saturated group of symmetries, satisfying 5.2. Then any matrix $A \in End(W)$ fixed by Ad_G is symmetric hence $\widehat{W} = Fix(Ad_G)$.*

PROOF. By 5.5 the involution ν is in G hence the result follows by 5.6. □

From the above it follows that, in the situation of 5.2, the space $\widehat{W}$ is the centralizer of G in $End(W)$ or, more precisely, the closure of the centralizer of $\rho(G)$ in $GL(W) \subset End(W)$.

References

[BW06] Weronika Buczyńska and Jarosław A. Wiśniewski. On phylogenetic trees - a geometer's view, http://arxiv.org/abs/math/0601357, 2006.

[BW07] Weronika Buczyńska and Jarosław A. Wiśniewski. On geometry of binary symmetric models of phylogenetic trees. *J. Eur. Math. Soc. (JEMS)*, 9(3):609–635, 2007.

[CS05] Marta Casanellas and Seth Sullivant. The strand symmetric model. In *Algebraic statistics for computational biology*, pages 305–321. Cambridge Univ. Press, New York, 2005.

[DK07] Jan Draisma and Jochen Kuttler. On the ideals of equivariant tree models, http://arxiv.org/abs/0712.3230, 2007.

[ES93] Steven N. Evans and T. P. Speed. Invariants of some probability models used in phylogenetic inference. *Ann. Statist.*, 21(1):355–377, 1993.

[GAP] GAP. Groups, algorithms, programming - a system for computational discrete algebra, at http://www.gap-system.org/.

[PS05] Lior Pachter and Bernd Sturmfels. *Algebraic statistics for computational biology.* Cambridge: Cambridge University Press. xii, 420 p., 2005.

[SS03] Charles Semple and Mike Steel. *Phylogenetics*, volume 24 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2003.

[SS04] Bernd Sturmfels and Seth Sullivant. Toric ideals of phylogenetic invariants, http://arxiv.org/q-bio/abs/0402015, 2004.

[SSE93] L. A. Székely, M. A. Steel, and P. L. Erdős. Fourier calculus on evolutionary trees. *Adv. in Appl. Math.*, 14(2):200–210, 1993.

[SX08] Bernd Sturmfels and Zhiqiang Xu. Sagbi bases of Cox-Nagata ring, http://arxiv.org/math/abs/0803.0892, 2008.

Instytut Matematyki UW, Banacha 2, PL-02097 Warszawa

E-mail address: `[wkrych],[marysia],[jarekw]@mimuw.edu.pl`