Variétés toriques : phylogénie et catégorie dérivées

Mateusz Michalek

UNIVERSITÉ DE GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

préparée dans le cadre d'une cotutelle entre l'Université de Grenoble et l'Académie Polonaise des Sciences (PAN)

Spécialité : **Mathématiques**

Arrêté ministériel : le 6 janvier 2005 -7 août 2006

Présentée par

Mateusz MICHALEK

Thèse dirigée par **Laurent MANIVEL** et **Jaroslaw WIŚNIEWSKI**

préparée au sein de l' **Institut Fourier**

dans **les École Doctorale MSTII**

Variétés toriques: phylogénie et catégories dérivées

Thèse soutenue publiquement le **29 mars 2012**,
devant le jury composé de :

M. Andrzej BIAŁYNICKI-BIRULA
Professeur, Université de Varsovie, Rapporteur
M. Zbigniew JELONEK
Professeur, Académie Polonaise des Sciences (PAN), Varsovie, Membre du Jury
M. Laurent MANIVEL
Directeur de recherche, Université Grenoble 1, Directeur de thèse
M. Andrzej SKOWROŃSKI
Professeur, Université de Nicolaus Copernicus de Torun, Membre du Jury
M. Bernd STURMFELS
Professeur, Université de Californie de Berkeley, Rapporteur
M. Tomasz SZEMBERG
Professeur, Université de pédagogie de Cracovie, Rapporteur
M. Henryk TORUŃCZYK
Professeur, Université de Varsovie, Président du Jury
M. Jaroslaw WIŚNIEWSKI
Professeur, Université de Varsovie, Directeur de thèse

Université Joseph Fourier / Université Pierre Mendès France / Université Stendhal / Université de Savoie / Grenoble INP

Contents

TORIC VARIETIES: PHYLOGENETICS AND DERIVED CATEGORIES

MATEUSZ MICHAŁEK

Streszczenie pracy w języku polskim

Rozmaitości toryczne: filogenetyka i kategorie pochodne

Celem niniejszej pracy doktorskiej jest badanie specjalnych własności rozmaitości torycznych. Praca jest podzielona na trzy części. Pierwsze dwie z nich są silnie ze sobą powiązane.

W pierwszej części zajmujemy się głównie badaniem rozmaitości algebraicznych związanych z procesami Markowa na drzewach. Z każdym procesem Markowa na drzewie można stowarzyszyć rozmaitość algebraiczną. W związku z motywacjami biologicznymi, skupiamy się na procesach Markowa określonych poprzez działanie grupy. Badamy warunki, kiedy uzyskane rozmaitości są toryczne oraz podajemy ich opis, Twierdzenie 5.63. Przedstawiamy twierdzenia, podające warunki wystarczające do tego, aby otrzymane rozmaitości były normalne, 5.73, jak również podajemy przykłady, gdy nie są one normalne 5.74, 5.75. Jednym z głównych używanych narzędzi jest uogólnienie pojęć wtyków i sieci, wprowadzonych w [BW07], do dowolnych grup abelowych. W naszej definicji sieci tworzą grupę, Definicja 5.24, która działa na rozmaitości. Ponadto, przestrzeń w której zanurzona jest rozmaitość jest regularną reprezentacją tej grupy.

Głównym otwartym problemem do którego odnosimy się w tej części jest hipoteza Sturmfelsa i Sullivanta [SS05, Hipoteza 2]. Stwierdza ona, że ideał afinicznej rozmaitości skojarzonej z modelem 3-Kimury jest generowany w stopniu 4. Nasz najsilniejszy wynik dowodzi, że schemat rzutowy związany z tym modelem może być opisany poprzez ideał generowany w stopniu 4, Twierdzenie 12.1. Wraz z Marią Donten–Bury przedstawiamy sposób generowania wielomianów należących do ideału stowarzyszonego z rozmaitością dla dowolnego modelu. Dowodzimy, że nasza metoda generuje cały ideał dla wielu modeli wtedy i tylko wtedy, gdy zachodzi hipoteza Sturmfelsa i Sullivanta [SS05, Hipoteza

1], Twierdzenie 7.8. Prezentujemy kilka zastosowań, na przykład do problemu identyfikowalności w biologii.

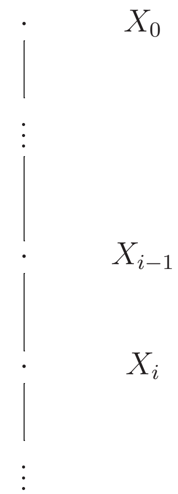
Druga część pracy dotyczy rozmaitości algebraicznych związanych z trójwalentnymi grafemi oraz modelem binarnym Jukesa-Cantora. Jest to wspólna praca z Weroniką Buczyńską, Jarosławem Buczyńskim i Kaie Kubjas. W przypadku grafu, stowarzyszona rozmaitość może być reprezentowana przez półgrupę z gradacją. Badamy związki pomiędzy własnościami grafu i otrzymanej półgrupy. Główne twierdzenie 14.1 dowodzi, iż pierwsza liczba Bettiego grafu plus jeden jest górnym oszacowaniem na stopień w którym generowana jest półgrupa.

W ostatniej części badamy kategorie pochodne gładkich, zupełnych rozmaitości torycznych. We wspólnej pracy z Michałem Lasoniem [LM11] konstruujemy pełne, silnie wyjątkowe kolekcje wiązek liniowych dla szerokiej klasy gładkich, zupełnych rozmaitości torycznych o liczbie Picarda równej trzy. Wiele pytań dotyczących jakiego rodzaju kolekcji można oczekiwać na rozmaitościach torycznych pozostaje otwartych. Jeden z otrzymanych wyników pokazuje, że $\mathbb{P}^n$ rozdmuchane w dwóch punktach nie posiada pełnej, silnie wyjątkowej kolekcji złożonej z wiązek liniowych dla wystarczająco dużego n. Otrzymujemy nieskończoną rodzinę kontrprzykładów do hipotezy Kinga 19.2. Pierwszy taki kontrprzykład został skonstruowany przez Hille i Perlinga [HP06]. Ostatnio Efimov podał także kontrprzykłady dla rozmaitości Fano [Efi].

Pracujemy nad ciałem liczb zespolonych $\mathbb{C}$. Wszystkie rozmaitości są rozmaitościami algebraicznymi w sensie [Har77].

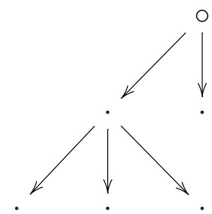
Streszczenie i wprowadzenie do części pierwszej

Motywacją dla konstrukcji rozpatrywanych w pierwszej części pracy jest matematyka stosowana. Zacznijmy od przypomnienia podstawowych własności łańcuchów Markowa oraz procesów Markowa na drzewach. Łańcuch Markowa to ciąg zmiennych losowych (X_i) spełniający określone warunki. Przy ustalonym stanie zmiennej X_{i-1} zmienna losowa X_i jest niezależna od wszystkich zmiennych losowych X_{i-j} dla $j > 1$. Zazwyczaj łańcuch Markowa jest przedstawiany jako ścieżka. Każdy wierzchołek odpowiada zmiennej losowej. Zmienne X_i oraz X_{i-1} są połączone, jak na rysunku poniżej.



Dla danego łańcucha Markowa wprowadza się prawdopodobieństwa warunkowe, które określają wszystkie własności łańcucha. Załóżmy, że zmienna X_i może być w $a_i < \infty$ stanach. Każdej krawędzi łączącej X_{i-1} z X_i możemy przypisać macierz o wymiarach $a_{i-1} \times a_i$. Kolumny i rzędy tej macierzy są oznaczone odpowiednio stanami zmiennych X_{i-1} oraz X_i. Odpowiednie wpisy w macierzy określają prawdopodobieństwa warunkowe. Konkretnie, wpis w p-tym rzędzie i i q-tej kolumnie odpowiada prawdopodobieństwu, że X_i jest w stanie p pod warunkiem, że X_{i-1} jest w stanie q. Otrzymane macierze nazywamy macierzami przejścia. Jeśli znamy rozkład zmiennej losowej X_0 oraz macierze przejścia, to możemy łatwo obliczyć rozkłady wszystkich zmiennych losowych występujących w danym łańcuchu Markowa.

Konstrukcję tę możemy bezpośrednio uogólnić do drzew ukorzenionych. Drzewem ukorzenionym określamy spójny graf, bez cykli, z wyróżnionym wierzchołkiem. Liście drzewa to wierzchołki, które posiadają tylko jednego sąsiada. Węzły to wierzchołki, które nie są liśćmi. W pracy czasami utożsamiamy liście z krawędziami z którymi są one połączone. Dla uproszczenia języka przyjmujemy, że drzewo jest grafem skierowanym i wszystkie krawędzie są skierowane od korzenia. W poniższym przykładzie korzeń został oznaczony jako $\circ$.

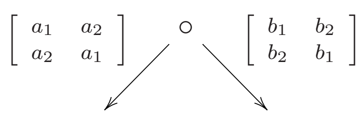


Tak jak w przypadku łańcuchów Markowa, każdemu wierzchołkowi przyporządkowujemy zmienną losową. Mówimy, iż wierzchołek v_1 jest bezpośrednim przodkiem v_2 jeśli istnieje krawędź skierowana od v_1 do v_2. Zauważmy, iż każdy wierzchołek ma dokładnie jednego bezpośredniego przodka, poza korzeniem, który nie posiada przodków. Potomkami wierzchołka v nazywamy wszystkie wierzchołki do których istnieje ścieżka skierowana, zaczynająca się w v. Własność Markowa stwierdza, iż zmienna X jest niezależna od wszystkich zmiennych które nie są jej potomkami, przy ustalonym stanie bezpośredniego przodka.

Procesy Markowa na drzewach są dobrymi modelami dla wielu zjawisk przyrodniczych. Sztandarowym przykładem jest tutaj proces ewolucji. Jednym ze znanych założeń jest fakt, iż DNA danego gatunku zależy tylko od stanu bezpośredniego poprzednika. Filogenetyka jest nauką badającą zmiany ewolucyjne. Jej głównym zadaniem jest opis procesu Markowa modelującego ewolucję gatunków. Przy tym modelu zakłada się, że zmienne mogą mieć cztery stany odpowiadające zasadom azotowym wchodzącym w skład DNA: adeninie, cytozynie, guaninie oraz tyminie. Stany te oznacza się literami A, C, G, T. Oczywiście, a priori, nie znamy parametrów macierzy przejścia, ani kształtu drzewa. Jednakże badając żyjące gatunki możemy poznać rozkład zmiennych losowych przypisanych liściom odpowiadającym tym gatunkom. Biologia teoretyczna przedstawia również możliwe typy macierzy przejścia. W zależności od modelu teoretycznego który wybierzemy, macierze przejścia mogą należeć do różnych przestrzeni liniowych. Różne modele biologiczne są przedstawione w Rozdziale 4. Bardzo interesujący jest fakt, iż modele zaproponowane przez biologów teoretycznych często posiadają własności ciekawe z matematycznego punktu widzenia. Dokładnie rzecz ujmując pewne przestrzenie macierzy przejścia są zadane jako macierze niezmiennicze ze względu na działanie grupy.

Przedstawmy jeden z możliwych sposobów rozwiązania problemów filogenetycznych, korzystający z geometrii algebraicznej. Ustalmy drzewo T, o którym podejrzewamy, że może właściwie opisywać proces ewolucji. Rozważmy macierze przejścia z wolnymi parametrami, które zależą jedynie od wybranego przez nas modelu biologicznego. Do przestrzeni parametryzującej dodajemy również parametry rozkładu zmiennej losowej stowarzyszonej z korzeniem. Dla danych parametrów obliczamy rozkład zmiennych losowych stowarzyszonych z liśćmi. Otrzymujemy odwzorowanie[1] $\pi \circ \widehat{\psi}$. Jego dziedzina to parametry macierzy przejścia oraz zmiennej losowej przypisanej korzeniowi. Obraz odwzorowania to wszystkie możliwe rozkłady zmiennych losowych przypisanych liściom.

Przykład *W tym przykładzie zakładamy, że każda zmienna może mieć dwa stany oznaczone poprzez* 0 *oraz* 1. *Korzeń ma dwóch potomków. Zmienna losowa przyjmuje wartość* 0 *z prawdopodobieństwem* λ_0 *oraz* 1 *z prawdopodobieństwem* λ_1. *Macierze przejścia mają następującą postać.*



Mamy 6 parametrów. Zmienne stowarzyszone z liśćmi mogą być w 4 *stanach:*

1) obie w stanie 0*,*
2) lewa w stanie 0*, prawa w stanie* 1*,*
3) prawa w stanie 1*, lewa w stanie* 0*,*
4) obie w stanie 1*.*

Otrzymujemy odwzorowanie:

$$\pi \circ \widehat{\psi} : (\lambda_0, \lambda_1, a_1, a_2, b_1, b_2) \to$$
$$(\lambda_0 a_1 b_1 + \lambda_1 a_2 b_2, \lambda_0 a_1 b_2 + \lambda_1 a_2 b_1, \lambda_0 a_2 b_1 + \lambda_1 a_1 b_2, \lambda_0 a_2 b_2 + \lambda_1 a_1 b_1).$$

Niech P będzie punktem, wyznaczonym na podstawie badań biologicznych, reprezentującym rozkład zmiennych losowych przypisanych liściom. Pragniemy stwierdzić czy punkt P należy do obrazu odwzorowania $\pi \circ \widehat{\psi}$. Jeśli punkt nie należy do obrazu, to możemy stwierdzić, iż wybrany model biologiczny jest błędny lub rozważane drzewo nie opisuje ewolucji w sposób prawidłowy. Jeśli punkt P należy do

[1] Wybór notacji zostanie uzasadniony w kolejnych rozdziałach.

obrazu, możemy pytać o włókno odwzorowania $\pi \circ \widehat{\psi}$ nad punktem P. Niestety stwierdzenie czy punkt należy do obrazu jest w ogólności bardzo trudne. Jedna z metod wykorzystuje fakt, iż odwzorowanie $\pi \circ \widehat{\psi}$ jest algebraiczne. Rozważa się domknięcie obrazu w topologii Zariskiego. Jest to afiniczna rozmaitość algebraiczna. Problem sprowadza się wtedy do opisu ideału tej rozmaitości i stwierdzeniu czy jego generatory zerują się na punkcie P. Elementy wyżej wymienionego ideału nazywane są *niezmiennikami filogenetycznymi*.

Podejście przez nas przedstawione może nie być efektywne. Opis ideału rozmaitości zadanej przez parametryzację nie jest prostym zadaniem. Jednakże odwzorowania, które rozpatrujemy często posiadają specjalne własności. Jak zauważyli Evans i Speed [ES93] rozmaitości stowarzyszone z niektórymi modelami ewolucji są toryczne. Dokładniej, istnieje układ współrzędnych w którym odwzorowanie parametryzujące rozmaitość jest zadane jednomianami. Pozwala to na zastosowanie metod geometrii torycznej przy wyznaczaniu ideału rozmaitości.

W całej pracy zakładamy, że *zmienna losowa stowarzyszona z wierzchołkiem posiada rozkład jednorodny.* Założenie to nie jest motywowane przez biologię. Otrzymujemy jednak dzięki niemu lepszy opis matematyczny. Z tego powodu zakładamy, że przestrzeń parametryzująca rozmaitość składa się tylko z parametrów macierzy przejścia.

Jednym z głównych celów pracy jest ustalenie przy jakich warunkach z danym modelem jest stowarzyszona rozmaitość toryczna oraz podanie jej opisu. Otrzymane wyniki przedstawiają bardzo ogólną konstrukcję 5.63. Wszystkie definicje obiektów występujących w twierdzeniu pojawią się w późniejszych rozdziałach.

Twierdzenie 5.63 *Niech H będzie normalną, abelową podgrupą grupy $G \subset S_n$. Załóżmy, że H działa w sposób tranzytywny i wolny na zbiorze S o n elementach. Rozważmy macierze przejścia należące do przestrzeni $\widehat{W}$, które są niezmiennicze ze względu na działanie grupy G. Niech W będzie przestrzenią wektorową rozpiętą przez wektory bazowe utożsamiane z elementami zbioru S. Model filogenetyczny dla dowolnego drzewa T związany z przestrzeniami W oraz $\widehat{W}$ zadaje toryczną rozmaitość algebraiczną.*

W szczególności rozpatrywane przez nas modele zawierają wszystkie modele biologiczne o których wiadomo, że są stowarzyszone z rozmaitościami torycznymi. Badamy również własności otrzymanych rozmaitości. Dowodzimy, że rozmaitości stowarzyszone z pewnymi modelami są normalne 5.73.

Twierdzenie *Modele filogenetyczne związane z dowolnym drzewem trójwalentnym oraz jedną z grup:* $\mathbb{Z}_2$, $\mathbb{Z}_2 \times \mathbb{Z}_2$, $\mathbb{Z}_3$ *oraz* $\mathbb{Z}_4$ *zadają rozmaitość normalną.*

Podajemy również przykłady rozmaitości, które nie są normalne 5.75. Następnie badamy dla jakich modeli rozmaitości stowarzyszone z drzewami trójwalentnymi o ustalonej liczbie liści należą do jednej rodziny płaskiej. Dla modelu binarnego Jukesa-Cantora fakt ten został udowodniony w pracy [BW07]. Dla 3-Kimury nie jest on prawdziwy, co wykazano w pracy [Kub10]. Obliczając wielomiany Hilberta wielu rozmaitości stwierdziliśmy, że większość rozważanych modeli nie ma tej własności.

Kolejny, bardzo istotny problem badany w doktoracie dotyczy niezmienników filogenetycznych.

Definicja (Drzewo gwieździste) Drzewo gwieździste $K_{n,1}$ to drzewo posiadające jeden węzeł i n liści.

Dla wielu modeli, w szczególności tych które są głównym przedmiotem tej pracy, wyznaczanie niezmienników filogenetycznych zostało zredukowane do przypadku drzewa gwieździstego [SS05], [AR08], [DK09]. Jednakże wyznaczenie ich nawet w tym szczególnym przypadku jest bardzo trudnym zadaniem. Nie wiemy nawet w jakim stopniu ideał stowarzyszonej rozmaitości jest generowany. Znana hipoteza Sturmfelsa i Sullivanta [SS05, Conjecture 1] podaje dokładne górne ograniczenie na ten stopień. Ciekawą obserwacją jest fakt, iż wyżej wymieniona hipoteza implikuje opis ideału jako sumy prostszych ideałów. Prezentujemy metodę generowania wielu niezmienników filogenetycznych dla dowolnego modelu, dla drzewa gwieździstego 7.2. Stawiamy hipotezę, iż nasza metoda pozwala w pełni opisać ideał. Dowodzimy, iż w wielu przypadkach nasza hipoteza jest równoważna hipotezie Sturmfelsa i Sullivanta – Twierdzenie 7.8. Nasz najsilniejszy wynik 12.1 dotyczący tego tematu dowodzi słabszej, teorio zbiorowej wersji [SS05, Hipoteza 2], co jest wystarczające z punktu widzenia zastosowań.

Streszczenie i wprowadzenie do części drugiej

Niech $\mathcal{G}$ będzie grafem trójwalentnym. Niech d będzie liczbą naturalną. Głównym przedmiotem naszych badań jest podzbiór $\tau(\mathcal{G})_d$ zbioru wszystkich numerowań krawędzi grafu $\mathcal{G}$ za pomocą liczb całkowitych. Dane numerowanie należy do $\tau(\mathcal{G})_d$, gdy spełnione są następujące wrunki:

- [♡♡] *(parzystość)* suma liczb przyporządkowanych krawędziom zawierającym dany wierzchołek jest parzysta;
- [+] *(dodatniość)* liczby przypisane krawędziom są nieujemne;
- [△] *(nierówności trójkąta)* trzy liczby przypisane krawędziom zawierającym dany wierzchołek spełniają warunek trójkąta;
- [°] *(ograniczenie stopnia)* dla dowolnego wierzchołka suma liczb przypisanych krawędziom, które go zawierają nie przekracza $2d$.

Szczegóły konstrukcji oraz formalne definicje znajdują się w Rozdziale 15. Badamy obiekt $\tau(\mathcal{G}) = \bigsqcup_{d\in\mathbb{N}} \tau(\mathcal{G})_d$, który poprzez dodawanie liczb przypisanym krawędziom ma strukturę monoidu. Nazywamy go **monoidem filogenetycznym grafu** $\mathcal{G}$. Główny wynik tej części to następujące twierdzenie.

Twierdzenie *Niech $\mathcal{G}$ będzie dowolnym grafem trójwalentnym o pierwszej liczbie Bettiego g. Monoid $\tau(\mathcal{G})$ jest generowany w stopniu co najwyżej $g+1$. Ponadto dla każdego g parzystego istnieją grafy dla których podane oszacowanie jest dokładne.*

Dla $g = 1$ oraz $g = 3$ także istnieją grafy, które nie są generowane w stopniu g. Konstruujemy również przykłady grafów o nieparzystej liczbie Bettiego g, które nie są generowane w stopniu $g-1$. Otwartym problemem pozostanie pytanie czy istnieją grafy o nieparzystym $g \geq 5$, które nie są generowane w stopniu g. Podajemy także dokładne stopnie w których generowane są monoidy stowarzyszone z grafami typu gąsienica z pętelkami przedstawionymi poniżej.

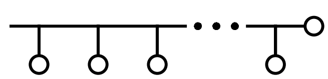


Figure 1: Gąsienica z pętelkami

Streszczenie i wprowadzenie do części trzeciej

W tej części wszystkie rozważane rozmaitości algebraiczne są gładkie. Czytelnikowi zainteresowanemu konstrukcją kategorii pochodnej snopów koherentnych na rozmaitości X polecamy pierwsze rozdziały [Huy06] lub artykuł [Căl05]. Dłuższym, klasycznym źródłem informacji na ten temat jest również książka [GM03].

Struktura i własności kategorii pochodnej mogą być bardzo skomplikowane i są przedmiotem licznych badań. Jeden ze sposobów opisu tej kategorii używa pojęć obiektów wyjątkowych. Przedstawy następujące definicje (patrz również [GR87]):

Definicja

(i) Snop koherentny F na X jest nazywany *wyjątkowym* jeśli $\mathrm{Hom}(F, F) = \mathbb{K}$ oraz $\mathrm{Ext}^{\,i}_{\mathcal{O}_X}(F, F) = 0$ dla $i \geq 1$.

(ii) Ciąg $(F_0, F_1, \ldots, F_m)$ snopów koherentnych na X nazywamy *kolekcją wyjątkową* jeśli każdy snop F_i jest wyjątkowy oraz $\mathrm{Ext}^{\,i}_{\mathcal{O}_X}(F_k, F_j) = 0$ dla $j < k$ oraz $i \geq 0$.

(iii) Kolekcja wyjątkowa $(F_0, F_1, \ldots, F_m)$ snopów koherentnych na X jest *silnie wyjątkową kolekcją* jeśli $\mathrm{Ext}^{i}_{\mathcal{O}_X}(F_j, F_k) = 0$ dla $j \leq k$ oraz $i \geq 1$.

(iv) (Silnie) wyjątkowa kolekcja $(F_0, F_1, \ldots, F_m)$ snopów koherentnych na X jest *pełną, (silnie) wyjątkową kolekcją* jeśli generuje ograniczoną kategorię pochodną $D^b(X)$ rozmaitości X, tzn. najmniejsza triangulowalna kategoria zawierająca $\{F_0, F_1, \ldots, F_n\}$ jest równoważna z $D^b(X)$.

W tej części doktoratu badamy pełne, silnie wyjątkowe kolekcje wiązek liniowych na gładkich, zupełnych rozmaitościach torycznych o liczbie Picarda 3. Wiadomo, że dla każdej gładkiej rzutowej rozmaitości torycznej istnieje pełna, wyjątkowa kolekcja snopów koherentnych – [Kaw06]. Jednakże wiele pytań w tej dziedzinie pozostaje otwartych. W szczególności nie wiadomo czy istnieje pełna, silnie wyjątkowa kolekcja snopów koherentnych lub czy istnieje pełna, wyjątkowa kolekcja złożona z wiązek liniowych. Wiadomo jednak, iż istnieją gładkie rzutowe rozmaitości toryczne nie posiadające pełnej, silnie wyjątkowej kolekcji złożonej z wiązek liniowych, co pierwotnie sugerowała hipoteza Kinga. Pierwszy kontrprzykład został podany w pracy [HP06]. W tej części pracy pokazujemy, iż $\mathbb{P}^n$ rozdmuchane w dwóch punktach nie posiada pełnej, silnie wyjątkowej kolekcji złożonej z wiązek liniowych dla dostatecznie dużych n – Twierdzenie 19.72.

Twierdzenie *Niech $n > 20$. Dowolna silnie wyjątkowa kolekcja wiązek liniowych na $\mathbb{P}^n$ rozdmuchanym w dwóch punktach ma długość* *co najwyżej $3n-2$. Ranga grupy Grothendiecka wynosi $3n-1$, więc kolekcja ta nie może być pełna.*

We wspólnej pracy z Michałem Lasoniem konstruujemy również takie kolekcje dla szerokiej klasy gładkich, zupełnych rozmaitości torycznych o liczbie Picarda 3. Rozmaitości te zostały sklasyfikowane przez Batyreva [Bat91] w terminach *kolekcji prymitywnych*. Są to minimalne kolekcje promieni wachlarza, które nie tworzą stożka. Takich kolekcji może być 3 lub 5. Przypadek, gdy występują tylko 3 kolekcje jest dobrze zbadany. Zajmowaliśmy się głównie przypadkiem 5 kolekcji. Wachlarze takie można dokładnie sklasyfikować. Definicje terminów występujących w klasyfikacji znajdują się w ostatniej części doktoratu.

Twierdzenie [Bat91, Theorem 6.6]
Niech $Y_i = X_i \cup X_{i+1}$, dla $i \in \mathbb{Z}/5\mathbb{Z}$,

$$X_0 = \{v_1, \ldots, v_{p_0}\}, \quad X_1 = \{y_1, \ldots, y_{p_1}\}, \quad X_2 = \{z_1, \ldots, z_{p_2}\},$$

$$X_3 = \{t_1, \ldots, t_{p_3}\}, \quad X_4 = \{u_1, \ldots, u_{p_4}\},$$

gdzie $p_0 + p_1 + p_2 + p_3 + p_4 = n + 3$. Dowolny n-wymiarowy wachlarz Σ o zbiorze generatorów promieni $\bigcup X_i$ oraz pięciu kolekcjach prymitywnych Y_i może być opisany z dokładnością do symetrii pięciokąta za pomocą następujących kolekcji prymitywnych o współczynnikach naturalnych $c_2, \ldots, c_{p_2}, b_1, \ldots, b_{p_3}$:

$$v_1+\cdots+v_{p_0}+y_1+\cdots+y_{p_1}-c_2z_2-\cdots-c_{p_2}z_{p_2}-(b_1+1)t_1-\cdots-(b_{p_3}+1)t_{p_3} = 0,$$

$$y_1 + \cdots + y_{p_1} + z_1 + \cdots + z_{p_2} - u_1 - \cdots - u_{p_4} = 0,$$

$$z_1 + \cdots + z_{p_2} + t_1 + \cdots + t_{p_3} = 0,$$

$$t_1 + \cdots + t_{p_3} + u_1 + \cdots + u_{p_4} - y_1 - \cdots - y_{p_1} = 0,$$

$$u_1+\cdots+u_{p_4}+v_1+\cdots+v_{p_0}-c_2z_2-\cdots-c_{p_2}z_{p_2}-b_1t_1-\cdots-b_{p_3}t_{p_3} = 0.$$

□

W celu odnalezienia pełnych, silnie wyjątkowych kolekcji używaliśmy metody pochodzącej od Bondala. Polega ona na rozważaniu rozpadu pchnięcia wiązki trywialnej przez odpowiednio wysoki toryczny morfizm Frobeniusa. Kolkecja taka nie musi być silnie wyjątkowa. Może nawet nie zawierać takiej kolekcji, co wykazaliśmy razem z Michałem Lasoniem. Jednakże dla bardzo wielu rozmaitości otrzymane w ten sposób wiązki liniowe stanowią dobry punkt wyjścia. W nowej pracy Efimov [Efi] wykazał, że istnieją gładkie, zupełne rozmaitości toryczne typu Fano o liczbie Picarda 3, nie posiadające pełnej, silnie wyjątkowej kolekcji wiązek liniowych.

Résumé en Francais

Variétés toriques: phylogénie et catégories dérivées

L'objectif de cette thèse est d'étudier les propriétés de variétés toriques particulières. La thèse est divisée en trois parties, les deux premières étant fortement liées.

Dans la première partie, nous étudions des variétés algébriques associées aux processus de Markov sur les arbres. A chaque processus de Markov sur un arbre on peut associer une variété algébrique. Motivé par la biologie, nous nous concentrons sur les processus de Markov définis par une action de groupe. Nous étudions les conditions pour que la variété obtenue soit torique, le théorème 5.63. Nous donnons un résultat où les variétés obtenues sont normales (cf proposition 5.73), ainsi que des exemples où elles ne le sont pas (cf proposition 5.74 et calcul 5.75). L'une des principales méthodes que nous utilisons est la généralisation des notions de prises et de réseaux introduites dans [BW07] à des groupes abéliens arbitraires. Dans notre contexte, les réseaux forment un groupe décrit à la définition 5.24 qui agit sur la variété. Par ailleurs, l'espace ambiant de la variété est la représentation régulière de ce groupe.

Le principal problème ouvert que nous essayons de résoudre dans cette partie est une conjecture de Sturmfels et Sullivant [SS05, Conjecture 2] indiquant que le schéma affine associé au modèle 3-Kimura est défini par un idéal engendré en degré 4. Notre meilleur résultat dit que le schéma projectif associé peut être défini par un idéal engendré en degré 4 (cf théorème 12.1). Avec Maria Donten–Bury, nous proposons une méthode pour engendrer l'idéal associé à la variété pour tous les modèles. Nous montrons que notre méthode fonctionne pour de nombreux modèles ainsi que pour les arbres si et seulement si la conjecture de Sturmfels et Sullivant est vraie (cf proposition 7.8). Nous présentons quelques applications, par exemple au problème d'identifiabilité en biologie.

La deuxième partie concerne les variétés algébriques associées aux graphes trivalents pour le modèle de Jukes-Cantor binaire. Il s'agit d'un travail en commun avec Weronika Buczyńska, Jarosław Buczyński et Kaie Kubjas. La variété associée á un graphe peut être représentée par un semi-groupe gradué. Nous étudions les liens entre les propriétés du graphe et le semigroupe. Le théorème principal 14.1 borne le degré en lequel le semi-groupe est engendré par le premier nombre de Betti du graphe, plus un.

Dans la dernière partie, nous étudions la structure de la catégorie dérivée des faisceaux cohérents des variétés toriques lisses. Dans un travail commun avec Michał Lasoń [LM11], nous construisons une collection fortement exceptionnelle complète de fibrés en droites pour une grande classe de variétés toriques complètes lisses dont le nombre de Picard est égal á trois. De nombreuses questions concernant le type de collections auxquelles on peut s'attendre sur les variétés toriques de certains types sont encore ouvertes. A ce titre, nous prouvons que $\mathbb{P}^n$ éclaté en deux points ne possède pas de collection fortement exceptionnelle complète de fibrés en droites pour n assez grand. Ceci fournit une collection infinie de contre-exemples à la conjecture de King 19.2. Le premier contre-exemple est dû à Hille et Perling [HP06]. Récemment, des contre-exemples ont également été trouvés par Efimov [Efi] dans le cadre des variétés de Fano.

Nous allons travailler sur le corps des nombres complexes $\mathbb{C}$. Toutes les variétés considérées sont des variétés algébriques dans le sens de [Har77].

General Introduction

The aim of this thesis is to investigate the properties of special toric varieties. The thesis is divided into three parts. The first two of them are strongly related to each other.

In the first, main part we study algebraic varieties associated to Markov processes on trees. To each Markov process on a tree one can associate an algebraic variety. Motivated by biology, we focus on Markov processes defined by a group action. We investigate under which conditions the obtained variety is toric, Theorem 5.63. We provide conditions ensuring that the obtained varieties are normal, 5.73, as well as give examples when they are not 5.74, 5.75. One of the main tools we use is the generalization of the notions of sockets and networks introduced in [BW07] to arbitrary abelian groups. In our setting the networks form a group, Definition 5.24, that acts on the variety. Moreover the ambient space of the variety is the regular representation of this group.

The main open problem that we address in this part is a conjecture of Sturmfels and Sullivant [SS05, Conjecture 2] stating that the affine scheme associated to the 3-Kimura model is defined by an ideal generated in degree 4. Our strongest result states that the associated projective scheme can be generated in degree 4, Theorem 12.1. Together with Maria Donten–Bury we also propose a method for generating the ideal defining the variety for any model. We prove that our method works for many models and trees if and only if the conjecture of Sturmfels and Sullivant holds, Proposition 7.8. We present some applications, for example to the identifiability problem in biology.

The second part concerns algebraic varieties associated to trivalent graphs for the binary Jukes-Cantor model. It is a joint work with Weronika Buczyńska, Jarosław Buczyński and Kaie Kubjas. In case of the graph, the associated variety can be represented by a graded semigroup. We investigate the connections between properties of the graph and the semigroup. The main theorem 14.1 bounds the degree in which the semigroup is generated by the first Betti number of the graph plus one. Due to connections with the first part much of the terminology that we use is either a specialization or generalization of previous definitions. From the one hand, as we are working with graphs with possible loops the notions of leaves, nodes and valency are more subtile than for trees. From the other hand, as we are dealing only with the binary Jukes-Cantor model, sockets and networks have got a very special form.

In the last part we study the structure of the derived category of coherent sheaves for smooth toric varieties. As a result of a joint work with Michał Lasoń [LM11] we construct a full, strongly exceptional collection of line bundles for a large class of smooth, complete toric varieties with Picard number three. Many questions concerning what kind of collections should be expected on toric varieties of certain types are still open. As a contribution we prove that $\mathbb{P}^n$ blown up in two points does not have a full, strongly exceptional collection of line bundles for n large enough. This provides an infinite collection of counterexamples to King's conjecture 19.2. The first such counterexample is due to Hille and Perling [HP06]. Recently also counterexamples in the Fano case were found by Efimov [Efi].

We will work over the field of complex numbers $\mathbb{C}$. All the varieties considered are algebraic varieties in the sense of [Har77].

Acknowledgements

First and foremost I want to thank my supervisors Laurent Manivel and Jarosław Wiśniewski. It was a great pleasure and opportunity to learn from them. Throughout the thesis they provided me with encouragement and ideas. I am sincerely grateful for their effort.

It is difficult to overstate my gratitude to Rosa-Maria Miró-Roig, who provided me with a lot of help. I am heartily thankful to Laura Costa, who together with Rosa-Maria, taught me a lot and presented inspiring and interesting problems.

I offer my sincerest gratitude to Laurent Bonavero, with whom I have wrote my master thesis and started the PhD studies. I thank also Sławomir Cynk, who with Laurent Bonavero, introduced me to algebraic geometry.

I am very grateful to Elizabeth Allman, Marta Casanellas, Bernd Sturmfels and Seth Sullivant for their explanations, answers to my questions and important remarks. Their expert opinions have helped a lot to improve the results of the thesis.

I sincerely acknowledge the administrative help of Emmanuel Peyre and Anna Poczmańska, who made the PhD studies possible.

I am grateful to Jan Draisma and Markus Perling for their crucial remarks.

Writing this thesis would not be possible without the support of my family and girlfriend. I thank a lot Gabriela Bebło, Martyna, Małgorzata and Krzysztof Michałek for their love and support.

I thank a lot my friends: Roland Abuaf, Piotr Achinger, Agnieszka Bodzenta, Weronika Buczyńska, Jarosław Buczyński, Maria Donten–Bury, Nathan Ilten, Thomas Kahle, Grzegorz Kapustka, Kaie Kubjas, Karine Kuyumzhiyan, Michał Kapustka, Kevin Langlois, Michał Lasoń, Christopher Manon, Karol Pałka, Clélia Pech, Ronan Terpereau for interesting talks, questions, answers and a great working atmosphere.

I acknowledge that during the work on my thesis I have been a scholarship fellow of the "Doctus – Małopolski fundusz stypendialny dla doktorantów" project cofunded by EU funds within European Social Fund. I was also supported by the grant of Polish Ministry of Science and Higher Education N N201 413 539.

1. Notation

We present the list of symbols used in the thesis. The definitions presented here are not formal and should be only treated as indications. Precise definitions are given later – we provide the references.

A_g	a transition matrix associated to the action of g on W, Definition 5.1
add	the morphism summing elements associated to edges at each vertex, Definition 5.16
$\mathfrak{add}$	extension of add to a lattice, Definition 6.3
add'	the group morphism summing elements associated to leaves, Definition 5.19
bi	a bijection between sockets and networks, Definition 5.29
C	a monoid in a lattice
GM	the category of G-models, Definition 10.3
GM^{ab}	the category of general group-based models
deg_e	the function summing up coordinates in M_e, Definition 5.39
$\deg_v(\omega)$	Definition 15.5
$D(X)$	the derived category of X, Subsection 18.1
$D^b(X)$	the bounded derived category of X, Subsection 18.1
E	the set of edges of a tree, Definition 4.2
$\mathcal{E}$	the set of edges of a graph
fo	a morphism that forgets coordinates, Definition 5.29
f_o	Definition 5.55
$\phi(G), \phi(G,n)$	phylogenetic complexity of a group, Subsection 7.1
G	a group
$\mathcal{G}$	a trivalent graph
G_N	Definition 5.13
H	an abelian group
$K_{n,1}$	the claw tree with n leaves, Definition 3.2
L	the set of leaves of a tree, Definition 4.2
Lab	a finite set of labels
l_χ	the basis element of $\widehat{W}$ indexed by a character of an abelian group, Definition 5.6
M	a lattice of characters

M_{deg}	a sublattice of M_E, Definition 6.1
M_e	the lattice with basis elements indexed by characters, Definition 5.32
$M_{E,G}$	the lattice with basis elements indexed by pairs of an edge and an orbit, Definition 5.64
M_E	the lattice with basis elements indexed by pairs of an edge and a character, Definition 5.32
$M_{E,0}$	sublattice of M_E, Definition 5.40
$\widehat{M}_{E,0}$	sublattice of $\widehat{M}_E$, Definition 5.40
$\widehat{M}_E$	a sublattice of M_E generated by points of P, Definition 5.38
M^{gr}	a lattice, Definition 15.2
M_S	the lattice with basis elements indexed by sockets, Definition 5.32
$M_{S,0}$	sublattice of M_S with coordinates summing up to zero, Definition 5.40
N	lattice of one parameter subgroups or the set of nodes of a tree
$\mathcal{N}$	the set of inner vertices of a graph
$\mathfrak{N}$	the group of networks, Definition 5.24
O	the set of orbits (usually of the adjunction action of a group G on H^*)
w_χ	a basis element of W indexed by a character of an abelian group, Definition 5.2
P	an integral polytope, (often representing the variety associated to a model, Definition 5.34)
$\mathbb{P}(X)_P$	an projective toric variety, Definition 2.7
π	Definition 4.9
$Poly$	the category of polytopes in lattices, Definition 10.4
$\widehat{\psi}$	Definition 4.7
$\check{\psi}$	the rational map induced by $\pi \circ \widehat{\psi}$, after Definition 4.9
$\widetilde{\psi}$	the morphism of lattices induced by $\widehat{\psi}$, Definition 5.33
p_v	a projection onto the vertex v, Definition 5.16
$\mathbb{P}(X(T, W, \widehat{W}))$	the projective variety associated to the tree T with a model distinguished by $\widehat{W}$
S	a finite set of states
$\mathfrak{S}$	the group of sockets, Definition 5.24
Σ	a fan, Definition 2.25
T	a rooted tree
$\mathcal{T}$	a trivalent tree

$\mathbb{T}$	an algebraic torus
$\mathbf{T}$	a real (topological) torus
$\tau(\mathcal{G})$	phylogenetic monoid, Definition 15.6
V	the set of vertices of a tree, Definition 4.2
$\mathcal{V}$	the set of vertices of a graph
W	a vector space with basis elements corresponding to states, Definition 4.1
$\widehat{W}_e$	Definition 4.4
$\widehat{W}_E$	Definition 4.6
$\widetilde{W}_E$	the space isomorphic to the ambient space of the variety representing the model, Definition 5.31
W_L	Definition 4.6
$\widetilde{W}_L$	the ambient space of the variety representing the model, Definition 5.31
W_v	Definition 4.4
W_V	Definition 4.6
$\widehat{W}$	the space of transition matrices, Definition 4.3
$X(T, W, \widehat{W})$	the affine variety associated to a tree T and a model distinguished by $\widehat{W}$
Y_i	primitive collection, Theorem 19.7
$\widehat{Y}_i$	collection of indices in Y_i
$\mathbb{Z}\mathcal{E}$	a lattice with basis elements indexed by edges, Definition 15.2
$\mathbb{Z}_2\mathcal{E}$	the group of networks for the binary Jukes-Cantor model, Definition 15.9

2. Toric varieties – the setting

The study of toric varieties is a relatively new subject. However its origins can be traced back even to Newton who had an idea to represent a polynomial by lattice points. To a monomial in n variables $x_1^{a_1} \cdots x_n^{a_n} =: x^a$ one associates a point $(a_1, \dots, a_n) \in \mathbb{Z}^n$. The following definition will not be used throughout the thesis. However we include it to give a reader not familiar with toric geometry first foundations.

Definition 2.1 (Newton polytope)**.** *Let $f = \sum_{a \in \mathbb{N}^n} \alpha_a x^a$ be a polynomial in n variables. The Newton polytope of f is the convex hull of points associated to monomials x^a, such that $\alpha_a \neq 0$. The definition can be easily extended to Laurent polynomials.*

To find much more information on Newton polytopes we advise the reader to consult [Stu98]. One of the first papers where toric varieties were studied in a systematic way is [KKMSD73]. The authors call toric varieties "toroidal embeddings" and view them as special compactifications of the algebraic torus $(\mathbb{C}^*)^n$. Classical reference for toric varieties are [Oda87] and [Ful93]. The latter book focuses more on the torus action. Recently a new, very modern, user friendly book appeared [CLS]. The point of view on toric varieties presented there is closest to the one from this thesis. The reasons why toric varieties have recently become so popular are numerous. A few most important are for sure:

(i) toric varieties are strongly related to combinatorial objects, what makes a lot of computations possible or at least easier,
(ii) toric varieties are simple, but fertile enough to provide a good testing ground for conjectures, proofs, theorems, examples,
(iii) toric varieties appear naturally as simplifications of other varieties,
(iv) toric varieties appear in applied mathematics.

This section contains well known results. We present the proofs, trying to find the easiest and most direct. We hope that, with little effort, the section can be read by people not familiar with toric geometry. Details that are skipped can be considered as exercises. We avoid referring to any general theorems, as the theory is, on this level, easy enough to develop from scratch. Many ideas presented in this part come from [CLS] and [Stu96]. We will use the setting presented in this section throughout the thesis. We encourage the reader familiar with toric geometry to take a look, as often our approach is different from the standard one.

In modern algebraic geometry a variety is locally described as a spectrum of an algebra. Thus the most important object connected to an affine algebraic variety is its ideal containing all polynomials vanishing on it. Note however that many varieties can be constructed in a different way. Given k polynomials $f_1, \dots, f_k$ in n variables one can consider the map $(f_1, \dots, f_k) : \mathbb{C}^n \to \mathbb{C}^k$. The Zariski closure of the image is an algebraic variety. Furthermore we can generalize this construction assuming that f_i are Laurent polynomials. In this case the domain of the map is $(\mathbb{C}^*)^n$. Let us start the discussion of toric geometry by introducing affine toric varieties. In most simple terms the study of affine toric varieties is the study of the case where all f_i are monomials.

Definition 2.2 (Affine toric variety). *Consider k Laurent monomials in n variables $f_i = x^{a_i}$, where $a_i \in \mathbb{Z}^n$. An affine toric variety is the Zariski closure of the image of the map $(f_1, \dots, f_k) : (\mathbb{C}^*)^n \to \mathbb{C}^k$.*

Note that we do not require the affine toric variety to be normal. This issue will be addressed later. Moreover the affine toric varieties come with an embedding in the affine space. Recalling Newton's idea the map (f_i) can be represented by k points $a_i \in \mathbb{Z}^n$. The geometry of these points is strongly related to the geometry of the affine toric variety. We will say that the variety is associated to the set of points $\{a_i\}$.

Proposition 2.3. *The ideal of the affine toric variety is generated by binomials. Suppose that the parametrization of the variety is given by k monomials f_i in n variables x_i. Let $P_i \in \mathbb{Z}^k$ be a point associated to f_i. A binomial $y_1^{b_1} \cdots y_k^{b_k} - y_1^{c_1} \cdots y_k^{c_k}$ for $b_i, c_i \in \mathbb{N}$ is in the ideal if and only if $\sum_i b_i P_i = \sum_i c_i P_i$.*

Proof. The binomials of the given form vanish on the image of the map $(f_1, \dots, f_k)$, hence also on the Zariski closure. We will prove that they not only generate the ideal, but span it as a vector space. Fix any order on the monomials. Suppose that the ideal is not spanned by the binomials of the given form. Let g be such a polynomial in the variables y_i that:

- is in the ideal of the variety,
- is not spanned by binomials of the given form,
- its leading coefficient is least possible.

Let $\alpha m(y_1, \dots, y_k)$ be the leading coefficient of g where m is a monomial. As g is in the ideal, by substituting y_i by f_i we get a Laurent polynomial that is zero on $(\mathbb{C}^*)^n$. Hence it has to be equal to zero. In particular the term $\alpha m(f_1, \dots, f_k)$ has to reduce with the term induced by some different monomial $\beta m'(f_1, \dots, f_k)$ appearing in g. Thus the monomials m and m' induce an integer relation between the points P_i. In particular $m - m'$ is a binomial of the chosen form. By subtracting $\alpha(m - m')$ from g we get a polynomial in the ideal with a strictly smaller leading coefficient which gives a contradiction. □

The above proposition allows us to describe the algebra of an affine toric variety.

Definition 2.4 (Semigroup algebra)**.** *Let $(C, \oplus)$ be a monoid. The monoid algebra $\mathbb{C}[C]$ as a vector space is spanned freely by the elements of C. The multiplication for $c_1, c_2 \in C \subset \mathbb{C}[C]$ is defined as $c_1 c_2 := c_1 \oplus c_2$ and extended to $\mathbb{C}[C]$ using the axioms of $\mathbb{C}$-algebra.*

Example 2.5. For the monoid $\mathbb{N}^n$ we obtain the algebra of polynomials in n variables. For the group $\mathbb{Z}^n$ we obtain the algebra of Laurent polynomials.

Corollary 2.6 (From Proposition 2.3)**.** *Consider the affine toric variety parameterized by monomials f_i in n variables. Let $P_i \in \mathbb{Z}^n$ be the* *point representing f_i. Let C be the monoid generated by points P_i. The algebra of the affine toric variety is $\mathbb{C}[C]$.* □

We will be often working with projective toric varieties.

Definition 2.7 (Projective toric variety)**.** *Consider $k+1$ Laurent monomials f_i in n variables. A projective toric variety is the Zariski closure of the map $(f_1, \dots, f_{k+1}) : (\mathbb{C}^*)^n \to \mathbb{P}^k$.*

If $P \subset \mathbb{Z}^n$ is the set of points representing the monomials f_i, we will say that the closure of the image of (f_i) in $\mathbb{P}^k$ is a projective toric variety associated to P and we will denote it by $\mathbb{P}(X)_P$. We can adapt Proposition 2.3 and Corollary 2.6. First let us consider an affine cone over a projective toric variety. Its parametrization is as follows:

$$(\lambda f_1, \dots, \lambda f_{k+1}) : (\mathbb{C}^*)^{n+1} \to \mathbb{C}^{k+1}.$$

Notice that we have added a nonzero parameter λ, as we passed to affine space. Of course λf_i is still a monomial. If f_i is represented by a point $P_i \in \mathbb{Z}^n$ then λf_i is represented by $P_i \times \{1\} \in \mathbb{Z}^{n+1}$. Thus in the projective case it is more natural to consider the points P_i in the lattice of dimension one bigger and put the last coordinate equal to 1. The monoid generated by $P_i \times \{1\}$ gives rise to a monoid algebra of the cone over the projective variety. Moreover the last coordinate gives the grading of this algebra. The projective toric variety is the Proj of this graded algebra. Thus affine toric varieties correspond to finitely generated monoids in $\mathbb{Z}^n$. Projective toric varieties correspond to finitely generated monoids in $\mathbb{Z}^{n+1}$ with generators with last coefficient equal to 1. A reader interested in this topic may extend these results to varieties embedded in weighted projective spaces as an exercise.

Usually one assumes that a toric variety is normal. Let us explain why. We start by recalling basic definitions.

Definition 2.8 (Normal algebraic variety)**.** *An affine algebraic variety is normal if and only if its algebra is integrally closed in its field of fractions. An abstract algebraic variety is normal if and only if it can be covered by normal affine algebraic varieties.*

The concept of normality is very important for a number of reasons. Let us recall that smoothness implies normality. Moreover the singular locus of a normal variety has codimension at least 2. Most toric geometers work with normal varieties, as this allows for a nice combinatorial description of the variety [Oda87, Theorem 1.4].

Definition 2.9 (Lattice)**.** *A lattice is a finitely generated abelian group with no torsion. In other words a lattice is an abelian group isomorphic to $\mathbb{Z}^n$.*

Consider a subset of points P in a lattice $M \simeq \mathbb{Z}^n$. As in the Definition 2.7 the set P defines a projective toric variety $\mathbb{P}(X)_P$ together with an embedding. Let X be the affine cone over $\mathbb{P}(X)_P$. Let C be the monoid generated by the points of $P \times \{1\} \subset M \times \mathbb{Z}$. We know that $X = \operatorname{Spec} \mathbb{C}[C]$. Let $\widetilde{M} \subset M \times \mathbb{Z}$ be the sublattice generated by $P \times \{1\}$.

Definition 2.10 (Projective normality)**.** *We call the projective variety $\mathbb{P}(X)$ projectively normal if and only if the affine cone X over this variety is normal.*

Of course each projectively normal variety is normal. In the toric setting both normality and projective normality can be described in combinatorial language.

Definition 2.11 (Saturated monoid, saturation, saturated set of points)**.** *Let C be a monoid contained in a lattice $\widetilde{M}$. We say that C is saturated if and only if for any $x \in \widetilde{M}$ and any positive integer k the element $kx \in C$ if and only if $x \in C$.*

For any monoid C one can define its saturation $\widetilde{C}$ that is the smallest saturated monoid containing C. In other words $x \in \widetilde{C}$ if and only if for some positive integer k we have $kx \in C$.

We say that a set of points is saturated in a lattice M if and only if it generates a saturated monoid. We say that a set of points is saturated if it is saturated in the lattice that it generates.

Definition 2.12 (Integral polytope)**.** *An integral polytope is a convex hull of a finite number of points in the lattice. As we will be dealing only with lattice polytopes we will often call them just polytopes.*

Definition 2.13 (Normal polytope)**.** *We say that a polytope $P \subset M$ is normal in the lattice M if and only if the set $P \times \{1\}$ is saturated in $M \times \mathbb{Z}$. We say that a polytope P is normal if and only if it is normal in the lattice that it generates.*

In other words a polytpe P is normal in the lattice M if and only if for any $k \in \mathbb{N}$ any point $Q \in kP \cap M$ is a sum of k points from P.

Note that it is very important to specify the lattice. Consider the polytope $P \subset M := \mathbb{Z}^3$. Let P have got four integral points: $(0,0,0)$, $(1,1,0)$, $(0,1,1)$, $(1,0,1)$. This is a normal polytope. Note however that it is not normal in M. Indeed $(1,1,1) \in 2P$ and $(1,1,1)$ is not the sum of two integral points of the polytope.

Note that if the set $P \times \{1\}$ is saturated then P must be a polytope in the lattice that it generates. Indeed suppose that $P \times \{1\}$ is a saturated set of points. Let M be the lattice spanned by P. Let $D \in M$ be a linear combination of points from P with positive coefficients summing up to 1. From the linear algebra it follows that we can assume that the coefficients are rational. Hence some multiple of $D \times \{1\}$ is in the monoid generated by $P \times \{1\}$. As $P \times \{1\}$ is saturated it must contain $D \times \{1\}$. Thus the convex hull of P intersected with M equals P. Hence P is a polytope.

Fact 2.14. *The variety $\mathbb{P}(X)_P$, defined by a set of points P, is projectively normal if and only if the set of points $P \times \{1\}$ is saturated. Equivalently P must be a normal polytope.*

Fact 2.15. *Let D be any point of the set $P \times \{1\}$. Let P_D be the set $P \times \{1\} - D$, where the minus is the lattice operation. The variety $\mathbb{P}(X)_P$ associated to $P \times \{1\}$ is normal if and only if for any $D \in P \times \{1\}$ the set P_D is saturated. In such a case P does not have to be normal.*

Proof. Both facts are a direct consequences of Proposition 2.22. For the first, the algebra of the cone over the variety equals the monoid algebra for the monoid C spanned by $P \times \{1\}$. The monoid C is saturated, if and only if P is normal.

For the second, one can notice that points of $P \times \{1\}$ correspond to variables of the ambient projective space. Consider the affine subvariety of $\mathbb{P}(X)$ corresponding to setting one variable, corresponding to a point D, to 1. The algebra of this affine variety is the monoid algebra associated to the monoid spanned by P_D. □

Definition 2.16 (Cone, cone over a polytope)**.** *A cone is a finitely generated, saturated monoid of a lattice.*

In the literature it is often called a convex polyhedral cone. More precisely in this thesis we identify lattice points of the polyhedral cone with the cone.

Let P be a polytope that spans the lattice M. The cone over P is the saturation of the monoid spanned by $P \times \{1\} \subset M \times \mathbb{Z}$.

We will see in Proposition 2.22 that normal affine toric varieties are associated to finitely generated cones. Projectively normal projective toric varieties are associated to cones over normal polytopes.

There is one important case where even in the projective case one can consider the set of points P instead of $P \times \{1\}$. Suppose that P is contained in a hyperplane given by an equation $\sum a_i x_i = b$ for $b \neq 0$. In this case the monoid generated by P is isomorphic to the monoid generated by $P \times \{1\}$. In the first part of the thesis we will be considering such polytopes.

We would like now to explain the name toric variety. It is connected to the algebraic torus $\mathbb{T} = (\mathbb{C}^*)^n = \operatorname{Spec} \mathbb{C}[x_i^{\pm 1}]$. Using coordinatewise multiplication $\mathbb{T}$ is an algebraic group. On the level of algebras the action is given by a morphism $\mathbb{C}[x_i^{\pm 1}] \to \mathbb{C}[x_i^{\pm 1}] \otimes \mathbb{C}[x_i^{\pm 1}]$ that associates to a generator x_i the tensor product $x_i \otimes x_i$. Note that an arbitrary Laurent polynomial f is not sent to $f \otimes f$. This is true only for monomials. Let us consider algebraic morphisms $\mathbb{T} \to \mathbb{C}^*$ that preserve the abelian group structure. These are called *characters*. Such a map is in particular a regular function on $\mathbb{T}$ hence must be given by a Laurent polynomial. Due to the fact that it must preserve the group structure one can prove that it must be a monomial. By identifying a monomial with a lattice point we see that characters form a lattice $\mathbb{Z}^n$. Intrinsically, one defines the sum of characters f and g by $(f + g)(x) = f(x)g(x)$.

Definition 2.17 (Lattice of characters M)**.** *The lattice of characters M of a torus $\mathbb{T}$ consists of morphisms of algebraic groups $\mathbb{T} \to \mathbb{C}^*$ with addition defined by $(f + g)(x) = f(x)g(x)$.*

Dually one defines one parameter subgroups as morphisms of algebraic groups $\mathbb{C}^* \to \mathbb{T}$. By projecting onto coordinates we see that each such morphism is of a form $t \to (t^{a_1}, \dots, t^{a_n})$ for $a_i \in \mathbb{Z}$. It can be identified with a point $(a_1, \dots, a_n) \in \mathbb{Z}^n$. Hence one parameter subgroups also form a lattice.

Definition 2.18 (Lattice of one parameter subgroups N)**.** *The lattice of one parameter subgroups of a torus $\mathbb{T}$ consists of morphisms of algebraic groups $\mathbb{C}^* \to \mathbb{T}$ with addition defined by $(\lambda + \delta)(t) = \lambda(t)\delta(t)$.*

It is well known that lattices M and N are dual. The pairing can be described as follows. Fix $f \in M$ and $\lambda \in N$. The composition $f \circ \lambda$ is a morphism of one dimensional tori. Hence it is a form $t \to t^a$. We define the product of f and λ to be equal to a. After using the identification of M and N with $\mathbb{Z}^n$ this is the standard scalar product.

As we have seen the characters correspond exactly to monomials in the algebra of the torus. Hence $\mathbb{T}$ is the spectrum of the monoid algebra $\mathbb{C}[M]$. Points of $\mathbb{T}$ correspond to maximal ideals of this algebra or to surjective morphisms of algebras $f : \mathbb{C}[M] \to \mathbb{C}$. Of course to determine such a morphism it is enough to define it on M. As M is a group its image has to be contained in $\mathbb{C}^*$. Moreover due to the fact that f is a map of algebras the map $M \to \mathbb{C}^*$ must preserve the group structure. Hence the points of $\mathbb{T}$ correspond to maps $M \to \mathbb{C}^*$ that preserve the group structure. Precisely for a point P we associate to a character χ its value on P.

Definition 2.19 (Abstract toric variety)**.** *A toric variety X is an algebraic variety, finitely generated over $\mathbb{C}$, containing $\mathbb{T}$ as a dense open subset. Moreover we require that the action of $\mathbb{T}$ on itself extends to an algebraic action on X.*

A crucial fact is that an abstract toric variety that is affine is an affine toric variety in the sense of Definition 2.2. This fact is usually proved using the following, very important lemmas.

Lemma 2.20. *Suppose that a torus $\mathbb{T}$ acts on a vector space V. Then there exists a basis of V such that the action is diagonal.*

Proof. For $t \in \mathbb{T}$ and $v \in V$ we have:

$$tv = \sum \chi(t) A_\chi(v),$$

where the sum is over a finite collection of characters of $\mathbb{T}$. One can notice that A_χ are projections to subspaces on which $\mathbb{T}$ acts by multiplication by a value of the corresponding character. □

Lemma 2.21. *The algebra of an abstract toric variety X that is affine is a monoid algebra associated to a monoid contained in the character lattice of the torus associated to the variety.*

We propose an approach that proves the this lemma directly.

Proof. As $\mathbb{T}$ is Zariski dense in X we know that the algebra A of X is a subalgebra of $\mathbb{C}[M]$. Fix $f \in A$. We know that $f = \sum_{i=1}^k a_i \chi_i$ for some $\chi_i \in M$ and $a_i \neq 0$. Let W be a vector space spanned by characters χ_i for $i = 1, \dots, k$. Consider the vector subspace $V := A \cap W$. Our first aim is to prove that $V = W$. Suppose that V is contained in a proper vector subspace. Let $(b_1, \dots, b_k)$ be such that if $\sum_{i=1}^k d_i \chi_i \in V$, then $\sum_{i=1}^k d_i b_i = 0$. By the assumptions $\mathbb{T}$ acts on X, hence on A. An action of a point $c \in \mathbb{T}$ on χ_i is given by $\chi_i(c)\chi_i$. Hence the action of c on f gives $\sum_{i=1}^k a_i \chi_i(c) \chi_i \in V$. Thus for any $c \in \mathbb{T}$ we must have $\sum_{i=1}^k b_i a_i \chi_i(c) = 0$. Hence $\sum_{i=1}^k b_i a_i \chi_i$ must be identically zero on $\mathbb{T}$. This is possible only if all $b_i = 0$ what gives a contradiction.

Hence the algebra A is spanned as a vector space by characters of M. Obviously these characters must form a monoid. □

As we have seen the algebra of an abstract toric variety X that is affine is equal to $\mathbb{C}[C]$ for a monoid $C \subset M$. As the algebra is finitely generated, so is the monoid C. Let $\chi_1, \dots, \chi_k$ be generators of C. Consider the embedding of the torus acting on X by $(\chi_1, \dots, \chi_k)$. Due to Corollary 2.6 its Zariski closure in $\mathbb{C}^k$ is isomorphic to X.

Proposition 2.22. *Let X be an affine toric variety. Let C be a monoid in the character lattice M of the torus acting on X. The variety X is normal if and only if C is a cone.*

Proof. First let us prove that if X is normal then C is saturated. Consider any point $kc \in C$ for $c \in M$. We want to prove that $c \in C$. For $m \in M$ let χ_m be a corresponding character. Consider a polynomial $f(X) = X^k - \chi_{kc}$ with coefficients in the algebra of X. Clearly χ_c satisfies the equation f. Moreover as C spans M, the character χ_c is in the quotient field of the algebra of X. Due to the normality of X we know that χ_c is also in the algebra. Hence $c \in C$.

Now we want to prove that if C is saturated, then $\mathbb{C}[C]$ is normal. First note that the quotient field of $\mathbb{C}[C]$ is equal to the quotient field of $\mathbb{C}[M]$. As the torus is smooth, its algebra is normal. One can also prove it by noticing that its algebra is a UFD (as it is a localization of the polynomial ring). Consider any monic polynomial $f \in \mathbb{C}[C][x]$. Suppose that g is in the quotient field and satisfies the equation $f(g) = 0$. From the normality of $\mathbb{C}[M]$ we know that $g \in \mathbb{C}[M]$. One can repeat the argument of Lemma 2.21. Namely we can act on the equation $f(g)$ by any point P of the torus. The action of P on f gives a monic polynomial with coefficients in $\mathbb{C}[C]$. Hence the action of P on g gives polynomials that are in the normalization of $\mathbb{C}[C]$. By the same arguments as in Lemma 2.21 we see that every character appearing in g with nonzero coefficient must be in the normalization of $\mathbb{C}[C]$. Thus we can assume that $g \in M$.

Suppose that f is of degree d. Notice that $f(g) = 0$ implies that $dg = d'g + c_0$ for some integer $0 \leq d' < d$ and $c_0 \in C$, as the character χ_{dg} must reduce with some other character. Thus $(d - d')g \in C$ and by normality $g \in C$. □

It is also worth mentioning how we can recover the torus of an affine toric variety given by a parametrization. There are a few equivalent ways to do this. Note that our construction of an affine or projective variety defines them with an embedding in an affine or projective space *with a distinguished system of coordinates*. These coordinates are in bijection with the points in the lattice that define the variety. The construction also distinguishes a dense torus in the embedding space. It contains all points with nonzero coordinates.

Fact 2.23. *Consider a parametrization $f = (f_1, \ldots, f_k) : \mathbb{T}' := (\mathbb{C}^*)^n \to \mathbb{C}^k$, where f_i are Laurent monomials in n variables. Let X be the Zariski closure of the image of this map. Let $\mathbb{T}'' = (\mathbb{C}^*)^k \subset \mathbb{C}^k$ be the torus containing all points with all coordinates different from zero, with* *the action given by coordinatewise multiplication. Let M' and M'' be the character lattices respectively of the tori $\mathbb{T}'$ and $\mathbb{T}''$. Then:*

(i) On the level of algebras the parametrization map f is induced by group homomorphism $\tilde{f} : M'' \to M'$,
(ii) The image $\mathbb{T}$ of $\mathbb{T}'$ in $\mathbb{T}''$ is Zariski closed, isomorphic to a torus, with the group action induced from $\mathbb{T}''$,
(iii) The character lattice of $\mathbb{T}$ is equal to the image of $\tilde{f}$ or equivalently to the quotient of M'' by the kernel of $\tilde{f}$,
(iv) The variety X contains $\mathbb{T}$ as a dense open subset and the action of $\mathbb{T}$ extends to X.

□

One can identify the torus $\mathbb{T}$ that acts on the projective toric variety $\mathbb{P}(X)_P$. As in the affine case it is the image of the parameterizing torus. It is also equal to the intersection of $\mathbb{P}(X)_P$ with a torus $\mathbb{T}''$ containing all points of the projective space with all coordinates different from zero. The action of $\mathbb{T}$ is induced from the action of $\mathbb{T}''$ on the projective space. Using the basis it is given by the coordinatewise multiplication.

We will be often comparing a projective variety with its affine cone. The following discussion concerns the ambient spaces. There is a natural morphism $m : \mathbb{C}^{n+1} \setminus \{0\} \to \mathbb{P}^n$. A system of coordinates distinguishes a torus $\mathbb{T}'$ in $\mathbb{C}^{n+1}$ consisting of the points with all coordinates different from zero. Let M' be the character lattice of $\mathbb{T}'$. Choose a coordinate system on $\mathbb{P}^n$ compatible with the one on $\mathbb{C}^{n+1}$ by the morphism m. The image of $\mathbb{T}'$ is a torus $\mathbb{T}''$ consisting of the points with all coordinates different from zero. Let M'' be the character lattice of $\mathbb{T}''$. Note that $\mathbb{C}^{n+1}$ is a toric variety, with the action of $\mathbb{T}'$ given by coordinatewise multiplication. So is $\mathbb{P}^n$ with the action of $\mathbb{T}''$. Each coordinate of $\mathbb{C}^{n+1}$ is a character of M'. All coordinates distinguish a basis of M'. The morphism m can be restricted to $\mathbb{T}'$ and can be considered as morphism of tori, preserving the group action. It induces a map of character lattices $\tilde{m} : M'' \to M'$. As m is a surjective morphism of tori, the morphism $\tilde{m}$ is injective. Hence M'' is a sublattice of M'. Using the basis of M' we can give a precise description of elements that belong to M''. Namely an element of M' belongs to M'' if and only if its sum of coordinates in M' is zero.

Definition 2.24 (Face of a cone)**.** *Let C be any cone in a lattice M. Let $v \in M^* = \mathrm{Hom}(M, \mathbb{Z})$. Suppose that for any $c \in C$ we have $v(c) \geq 0$. Let $v^\perp$ be a hyperplane of M consisting of elements x such that $v(x) = 0$. A face of the cone C is any subset that is given by $v^\perp \cap C$*

for some v satisfying the conditions above. Notice that a face of a cone is also a cone.

Equivalently a face F of C can be defined as a submonoid satisfying the following condition:

- *For any $c_1, c_2 \in C$ such that $c_1 + c_2 \in F$ we have $c_1, c_2 \in F$.*

For an affine toric variety corresponding to a cone C the faces of C correspond to orbits of the torus acting on it. Let us present this correspondence in details. We fix a finitely generated monoid C in a lattice M and its generators $\chi_1, \dots, \chi_k \in C$. As in Definition 2.2 the closure of the embedding in $\mathbb{C}^k$ of the torus $\operatorname{Spec} \mathbb{C}[M]$ by characters χ_i is the affine toric variety $X := \operatorname{Spec} \mathbb{C}[C]$. Note that we distinguished a basis in $\mathbb{C}^k$, but not on the torus $\mathbb{C}[M]$. Due to Fact 2.23 we know that:

- the dense torus orbit of X contains precisely those points that have all coordinates different from zero,
- the character lattice of the torus acting on X is equal to the sublattice of M spanned by C.

We will generalize this to other orbits. Assume that C is a cone. Each orbit will be indexed by a face F of the cone. The face F distinguishes a subset I of indices from $\{1, \dots, k\}$ such that $i \in I$ if and only if $\chi_i \in F$. The orbit corresponding to F can be characterized as follows:

1) the orbit contains precisely those points that have got coordinates corresponding to $i \in I$ different from zero and all other equal to zero,
2) the orbit is a torus with a character lattice spanned by elements of F,
3) the closure of the orbit is a toric variety given by the cone F,
4) each point of the orbit is a projection of the dense torus orbit onto the subspace spanned by basis elements indexed by indices from I,
5) the inclusion of the orbit in the variety is given by a morphism of algebras $\mathbb{C}[C] \to \mathbb{C}[F]$. This morphism is an identity on $F \subset \mathbb{C}[C]$ and zero on $C \setminus F$.

Note that each orbit will contain a unique distinguished point given by the projection of the point $(1, \dots, 1) \in \mathbb{C}^k$. We will only present a sketch of a proof of these observations.

Proof. As in case of the torus we can identify the points of X with monoid morphisms $C \to (\mathbb{C}, \cdot)$. Fix any point $x \in X$. The characters $\chi \in C$ such that $\chi(x) \neq 0$ must form a face of F. Hence x distinguishes a subset of indices $I \subset \{1, \dots, k\}$. Of course the set of points with nonzero coordinates indexed by I and other coordinates equal to zero in X is invariant with respect to the action of the torus acting on X. So to prove 1) it is enough to prove that all these points are in one orbit. The point x represents a morphism $C \to (\mathbb{C}, \cdot)$ that is nonzero on F and zero on $C \setminus F$. Consider the restriction of this morphism to F. As it is nonzero it can be extended to a morphism $M' \to \mathbb{C}^*$, where M' is a sublattice generated by F. Next we can extend this morphism to the lattice M'' generated by C. Thus we obtain a morphism $f : M'' \to \mathbb{C}^*$ that agrees with the one representing x on F. Note that f represents a point p in the torus acting on X. By the action of p^{-1} on x we obtain a point given by a morphism that associates one to elements from F and zero to elements from $C \setminus F$. Thus we have proved 1). Moreover we showed that each orbit contains the distinguished point. Point 2) follows, as morphism that are nonzero on F and zero on $C \setminus F$ are identified with morphisms from M' to $\mathbb{C}^*$. Point 3) is a consequence of 2) and previous discussion on affine toric varieties. Indeed, we already know that the orbit is a torus with the lattice generated by F. This torus is the image of the torus $\operatorname{Spec} \mathbb{C}[M]$ in $\mathbb{C}^k$ by characters from I and all other coordinates equal to zero. Let A be the affine space spanned by basis elements indexed by indices in I. The orbit corresponding to F is contained in A. In fact, by the construction it is the image of $\operatorname{Spec} \mathbb{C}[M]$ by characters χ_i, such that $i \in I$. The closure of this torus is exactly given by $\operatorname{Spec} \mathbb{C}[F]$, as generators of the monoid C contained in F are generators of F. Point 4) is obvious, as the point p constructed in the first part of the proof projects to x. □

We finish this section by stating some results about normal abstract toric varieties.

Definition 2.25 (Fan). *A fan Σ is a finite collection of cones in a lattice that satisfy the following conditions:*
1) if a cone C is in the fan then all its faces are also in the fan,
2) an intersection of any two cones from the fan is a face of both,
3) for any cone $C \in \Sigma$ if $x \in C$, then $-x \notin C$.

A general, normal toric variety can be represented by a fan in the one parameter subgroups lattice N.

Definition 2.26 (Dual cone). *Let L and L' be dual lattices with the pairing given by $(\cdot, \cdot)$. Let $\delta \subset L$ be a cone in L. We define the dual cone $\delta^* \subset L'$ as:*

$$\delta^* = \{x \in L' : \text{ for any } y \in \delta \text{ we have } (x, y) \geq 0\}.$$

A toric variety X is constructed from a fan Σ by gluing together affine schemes $\operatorname{Spec}(\mathbb{C}[\sigma_i^*])$, where $\sigma_i^* \subset M$ is a cone dual to $\sigma_i \in \Sigma$.

One dimensional cones in Σ are called rays. The generators of these monoids are called ray generators.

Many properties of the variety X can be described using the fan Σ. For example X is smooth if and only if for every cone σ_i the set of its ray generators can be extended to a basis of N. Moreover to each ray generator v we may associate a unique T invariant Weil divisor denoted by D_v. For fans containing maximal dimensional cones there is a well known exact sequence:

$$0 \to M \to Div_T \to \; Cl(X) \to 0, \tag{2.1}$$

where Div_T is the group of T invariant Weil divisors and $Cl(X)$ is the class group. The map $M \to Div_T$ is given by:

$$m \to \sum m(v_i) D_{v_i},$$

where the sum is taken over all ray generators v_i.

So far we have defined objects of the category of toric varieties. Not every algebraic morphism is a morphism in this category. Indeed, as toric varieties are endowed with the torus action, it is natural to distinguish those morphisms that respect this action.

Definition 2.27 (Toric morphism)**.** *Let $f : X \to Y$ be a morphism of toric varieties. Let $\mathbb{T}_X \subset X$, $\mathbb{T}_Y \subset Y$ be the tori acting respectively on X and Y. We call f a toric morphism if $f(\mathbb{T}_X) \subset \mathbb{T}_Y$ and for any points $p, q \in \mathbb{T}_X$ we have:*

$$f(pq) = f(p)f(q).$$

Notice that, as the tori are Zariski dense in the varieties, this immediately implies that for any $p \in \mathbb{T}_X$ and $q \in X$ the same equality holds.

As the restriction of the toric morphism is a morphism of algebraic tori, it induces a map of character lattices $\tilde{f} : M_Y \to M_X$. By dualizing, this gives a map of one parameter subgroups $\tilde{f}^* : N_X \to N_Y$. In fact one can easily characterize which morphisms of one parameter subgroups give rise to toric morphisms. For each cone δ in the fan representing X there must be a cone δ' in the fan representing Y such that $\tilde{f}^*(\delta) \subset \delta'$.

Much more information on the topic can be found in [CLS], [Ful93].

Part 1. Algebraic varieties associated to Markov processes on trees

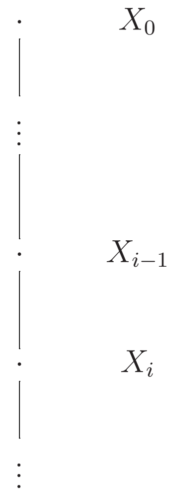
Dans la première partie, nous étudions des variétés algébriques associées aux processus de Markov sur les arbres. A chaque processus de Markov sur un arbre on peut associer une variété algébrique. Motivé par la biologie, nous nous concentrons sur les processus de Markov définis par une action de groupe. Nous étudions les conditions pour que la variété obtenue soit torique, le théorème 5.63. Nous donnons un résultat où les variétés obtenues sont normales (cf proposition 5.73), ainsi que des exemples où elles ne le sont pas (cf proposition 5.74 et calcul 5.75). L'une des principales méthodes que nous utilisons est la généralisation des notions de prises et de réseaux introduites dans [BW07] à des groupes abéliens arbitraires. Dans notre contexte, les réseaux forment un groupe décrit à la définition 5.24 qui agit sur la variété. Par ailleurs, l'espace ambiant de la variété est la représentation régulière de ce groupe.

Le principal problème ouvert que nous essayons de résoudre dans cette partie est une conjecture de Sturmfels et Sullivant [SS05, Conjecture 2] indiquant que le schéma affine associé au modèle 3-Kimura est défini par un idéal engendré en degré 4. Notre meilleur résultat dit que le schéma projectif associé peut être défini par un idéal engendré en degré 4 (cf théorème 12.1). Avec Maria Donten–Bury, nous proposons une méthode pour engendrer l'idéal associé à la variété pour tous les modèles. Nous montrons que notre méthode fonctionne pour de nombreux modèles ainsi que pour les arbres si et seulement si la conjecture de Sturmfels et Sullivant est vraie (cf proposition 7.8). Nous présentons quelques applications, par exemple au problème d'identifiabilité en biologie.

3. Introduction

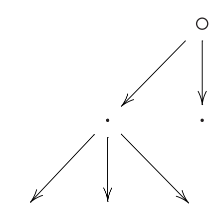
The motivation for the constructions in the first part of the thesis comes from applied mathematics. Let us recall basic properties of Markov chains and Markov processes on trees. A Markov chain is a sequence of random variables $\{X_i\}$ that satisfy specific conditions. For a fixed state of a variable X_{i-1} the variable X_i is independent from the set of all the variables X_{i-j} for $j > 1$. Typically, this chain is depicted vertically by associating a vertex to each variable and joining X_i with

X_{i-1}.



For a Markov chain we usually introduce conditional probabilities that specify all the properties of the chain. Suppose that each variable X_i can be in $a_i < \infty$ states. Then to each edge joining X_{i-1} and X_i we can associate an $a_{i-1} \times a_i$ matrix. The columns and rows of the matrix are indexed respectively by states of X_{i-1} and X_i. The given entries correspond to conditional probabilities. Namely, an entry indexed by a pair of states (p, q) equals the probability that X_i is in the state q under the condition that X_{i-1} is in the state p. These matrices are called transition matrices. If we know the distribution of X_0 and the transition matrices we can easily calculate the distributions of all other variables.

This construction can be directly generalized to rooted trees. By a rooted tree we will always mean a connected graph with one distinguished vertex and no cycles. By leaves we mean vertices of valency one. Nodes are vertices that are not leaves. In the thesis we will sometimes identify leaves with edges adjacent to them. To simplify the language we assume that the tree is a directed graph and all the edges are directed away from the root. In the example below the root is denoted by $\circ$.



As before to each vertex we associate a random variable. We say that a node v_1 is a direct ancestor of v_2 if there is an edge directed from

v_1 to v_2. Note that there is always one direct ancestor, except for the root that does not have ancestors. The descendants of a vertex are all the vertices that can be reached from it by a directed path. The Markov property ensures that a variable X is independent from all other variables that are not its descendants once the state of the direct ancestor is fixed.

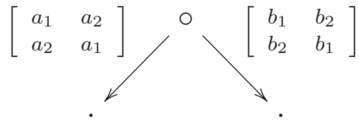
Markov processes on trees are good models for many empirical phenomena. For example evolution processes are often modeled it this way. It is intuitively plausible that the DNA of a species depends only on the state of its direct ancestor. The science that models the evolutionary changes is called phylogenetics. For more information about mathematical and computational methods in phylogenetics the reader is advised to consult [SS03] and [Fel04]. The main aim of phylogenetics is to establish the Markov process that models evolution of species. In this situation we assume that the random variables have four states corresponding to four nucleobases that form the DNA. These are called adenine, cytosine, guanine, thymine and are denoted respectively by A, C, G and T. A priori we do not know the transition matrices and the shape of the tree. However, by examining the living species, we know the distribution of random variables associated to leaves. Theoretical biology also provides us with possible types of transition matrices. According to the theoretical model we choose the transition matrices may belong to different linear subspaces. Different biological models are discussed in Section 4. A very interesting fact is that the models proposed by theoretical biologists often have very nice mathematical properties. Precisely certain subspaces of possible transition matrices are given as invariants under a group action.

One of the possible approaches to solve the problems in phylogenetics using algebraic geometry is as follows. We fix a rooted tree T that we suspect is a correct model of evolution. We consider any transition matrices with entries that are free parameters, that possibly depend only on the biological model that we choose. To the space of parameters we add also possible distributions of the variable associated to the root. We calculate the distribution of random variables associated to leaves. More precisely we get a map[2] $\pi \circ \widehat{\psi}$. Its domain parameterizes entries of transition matrices and possible distributions of the random variable associated to the root. Its image parameterizes all possible distributions of the random variables associated to leaves.

Example 3.1. In this example we suppose that each variable can be in two states denoted by 0 and 1. There is one root with two descendants.

[2]The reason for choosing this notation will become clear in the following sections

The variable associated to the root attains the value 0 and 1 with the probability given respectively by λ_0 and λ_1. The transition matrices are as follows.



Hence there are 6 parameters. The leaves can be in 4 states. We order them as follows:

1) both leaves are in state 0,
2) the left leaf is in state 0 and right in state 1,
3) the left leaf is in state 1 and right in state 0,
4) both leaves are in state 1.

We obtain the map:

$$\pi \circ \widehat{\psi} : (\lambda_0, \lambda_1, a_1, a_2, b_1, b_2) \to$$

$$(\lambda_0 a_1 b_1 + \lambda_1 a_2 b_2, \lambda_0 a_1 b_2 + \lambda_1 a_2 b_1, \lambda_0 a_2 b_1 + \lambda_1 a_1 b_2, \lambda_0 a_2 b_2 + \lambda_1 a_1 b_1).$$

Let P be the point, established empirically, that represents the distribution of random variables associated to leaves. We would like to check if P belongs to the image of $\pi \circ \widehat{\psi}$. If it is not in the image, then we know that either the biological model we used is wrong, or the tree T is not the right one. If the point P is in the image, we can ask for a description of the fiber. However determining if P belongs to the image is hard in general. One of the methods bases on the fact that $\pi \circ \widehat{\psi}$ is an algebraic map. We can consider the Zariski closure of its image. This is an affine algebraic variety. One would like to describe its ideal and check weather the generators vanish at P. The elements of this ideal are called phylogenetic invariants.

This approach may be not very effective. The description of the ideal of a variety given by a parametrization is not an easy task. However the maps we get are not arbitrary. As it was observed first by Evans and Speed [ES93] for certain models of evolution the variety we consider is toric. More precisely there are coordinates in which the parametrization map is given by monomials. This allows to apply methods of toric geometry in order to determine the ideal of the variety.

Throughout the thesis we assume that *the random variable associated to the root has got a uniform distribution.* This assumption is not motivated by biology. We use it only to obtain nicer results from the mathematical point of view. Hence in our study the parameter space contains only coefficients of transition matrices.

One of the main aims of this thesis is to determine under what conditions the model of evolution gives rise to toric varieties. Our results give the most general known criterion 5.63. In particular we believe that our approach covers all biological models of interest that were known to give rise to toric varieties. Further we investigate properties of the obtained toric varieties. We prove that varieties associated to certain biological models are normal 5.73. However we give also examples where the obtained varieties are not normal 5.75. Next we address the question for which models the varieties associated to trivalent[3] trees belong to the same flat family. For the binary Jukes-Cantor this fact was known to be true by [BW07], while for 3-Kimura it does not hold due to [Kub10]. By calculating Hilbert polynomials of many varieties we found out that most considered models do not have this property.

Another very important task concerns phylogenetic invariants.

Definition 3.2 (Claw tree)**.** *A claw tree $K_{n,1}$ is a tree with exactly one inner vertex and n leaves.*

For many models, in particular those that are most important for us, the study of phylogenetic invariants of any tree was reduced to the case of the claw tree [SS05], [AR08], [DK09]. However establishing phylogenetic invariants in this special case turned out to be very difficult. We do not even know the degree in which the ideal of phylogenetic invariants is generated. There is a well-known conjecture due to Sturmfels and Sullivant [SS05, Conjecture 1] that gives a precise upper bound for this degree. The conjecture is astonishingly similar to an old theorem of Noether. The theorem bounds the degree in which the ring of invariants of the group action on the polynomials is generated. However, as we will see in Section 6 it is hard to give a description of the whole algebra of the phylogenetic variety as a ring of invariants. Moreover, even if some description is possible, the order of the group is big – Corollary 6.6. One of interesting observations is that the conjecture implies a description of the ideal as a sum of more simple ideals. In fact we propose a method for obtaining many phylogenetic invariants for any model for the claw tree 7.2. We conjecture that our method gives a description of the whole ideal. We show that in many cases our conjecture is equivalent to the one made by Sturmfels and Sullivant 7.8. Our strongest result 12.1 in this topic proves a weaker, set-theoretic version of [SS05, Conjecture 2], that is sufficient for applications.

[3]The valency of all vertices is either one or three.

4. Basic definitions

The section introduces objects that will be studied in the first part of the thesis. The subsection 4.1 is the most important. Other parts can be treated as motivations and examples.

We will be dealing with algebraic varieties associated to phylogenetic models. These varieties are always given as closures of the image of a parametrization map – details will be presented in Section 4.1. A short, algebraic introduction to the topic can be found in [ERSS04].

Let S be a finite set, called the set of states. In the biological setting S is often supposed to have four elements. These elements correspond to four nucleobases. The set S is the codomain of random variables in the Markov process. Let $\Delta \subset \mathbb{R}^{|S|}$ be the probabilistic simplex that contains all the points with nonnegative coordinates summing up to one. The points of Δ parameterize all possible distributions of random variables with the set of states equal to S. In algebraic geometry instead of considering the simplex Δ one considers the whole complex vector space $\mathbb{C}^{|S|}$.

Definition 4.1 (Space W)**.** *We define W to be a complex vector space spanned freely by elements of S. More precisely $W = \oplus_{a \in S}\mathbb{C}_a$, where $\mathbb{C}_a$ is a field of complex numbers corresponding to one dimensional vector space spanned by $a \in S$.*

Suppose that we are given a rooted tree T with edges directed from the root.

Definition 4.2 (Sets L, V, N and E)**.** *Let L, V, N and E be respectively the set of leaves, vertices, nodes and edges of the tree T. We have $V = L \cup N$ and $L \cap N = \emptyset$. We identify leaves with edges adjacent to them.*

The objects that we study are derived from Markov processes on a tree. To each vertex one can associate a random variable with the set of states equal to S. The Markov property ensures that the variable in a vertex depends only on the variable associated to its first ancestor. Formally let X_i be a variable associated to a vertex v_i. Suppose that there is an edge directed from v_1 to v_2. Consider any set of vertices $v_3, \dots, v_j$ that are not descendants of v_2. Then $P(X_2 = x_2 | X_1 = x_1, X_3 = x_3, \dots, X_j = x_j) = P(X_2 = x_2 | X_1 = x_1)$, where x_i are some states. This mathematical model is applied for example in phylogenetics. The nodes of the tree correspond to species and the Markov property describes the fact that evolutionary changes depend only on the direct ancestor. More information on Markov processes can be found for example in [Ibe09]. The reader interested in phylogenetics in advised to look in [PS05]. There one can also find a detailed explanation of the relationship between Markov processes on trees and models that we consider.

To define a model we need to distinguish a subspace $\widehat{W} \subseteq \mathrm{End}(W)$.

Definition 4.3 (Transition matrix)**.** *Any element of the space $\widehat{W}$ represented as a matrix in the basis corresponding to S is called a transition matrix.*

The entries of a transition matrix correspond in biology to probabilities of mutation. Most often a model is distinguished by specifying the type of transition matrices.

Let us present some of the models.

(i) **The Cavender-Farris-Neyman model** also called **2-state Jukes-Cantor model**[4]. This is the most simple model. It was first introduced in [Ney71]. In most of biological articles it is called the Cavender-Farris-Neyman model or just the Neyman model. However recently, especially in algebraic phylogenetics, it is called the 2-state Jukes-Cantor model or the binary model [SS05], [BW07], [ERSS04]. In this model S has got two elements and the transition matrices are of the following type:

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

This model has got a lot of nice properties. One of the most interesting is the fact that the algebraic varieties arising from trivalent trees with the same number of leaves are deformation equivalent – see [BW07] for the original, algebraic proof and [Ilt10] for a combinatoric one. It is a general group-based model for the group $G = \mathbb{Z}_2$ – the definition of general group-based models will be introduced in subsection 5.1.

(ii) **3-Kimura model**. This is a four state model. It was introduced in [Kim81]. It is a general group-based model for the natural action of the group $G = \mathbb{Z}_2 \times \mathbb{Z}_2$ on the nucleobases A, C, G, T [ES93]. The transition matrices are of the type

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

[4]We would like to thank Elizabeth Allman for the information on the ambiguity.

(iii) **2-Kimura model**. This is a model for four states. It was introduced in [Kim80]. The transition matrices are of the type:

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

(iv) **Jukes-Cantor model**. This is the most simple model for four states. It was introduced in [JC69]. The transition matrices are of the type:

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

(v) **General Markov model**. This model can be considered on any number of states, but for biological reasons it is typically considered for four states. The space $\widehat{W}$ is equal to the whole space of endomorphisms $\operatorname{End} W$. Hence for four states the transition matrices are arbitrary:

$$\begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{bmatrix}.$$

4.1. **A variety associated to a model.** We will associate an algebraic variety to a tree T and a space $\widehat{W} \subset \operatorname{End} W$. This is a standard construction. In the literature one can find a lot of generalizations of the approach presented here – see for example [DK09].

Definition 4.4 (Spaces W_v and $\widehat{W_e}$)**.** *To each vertex v of the tree we attach a complex vector space W_v with a fixed isomorphism $iso_v : W \simeq W_v$. The images of the basis elements of W corresponding to states S by iso_v give a basis of W_v. The elements of this basis are denoted by $\{\alpha_v\}$. We also consider a vector space $\widehat{W} \subset \operatorname{End}(W)$, determined by the model we choose. To each edge e of the given rooted tree T we associate a vector space $\widehat{W_e}$ isomorphic to $\widehat{W}$.*

Remark 4.5. The natural basis on W induces an isomorphism $W \cong W^*$. Hence $\operatorname{End}(W) \cong W^* \otimes W \cong W \otimes W$. We may regard $\widehat{W}$ and respectively each $\widehat{W_e}$ as subspaces of $W \otimes W$.

Definition 4.6 (Spaces W_V, $\widehat{W}_E$, W_L)**.** *We recall that V, L and E are respectively the set of vertices, leaves and edges of a tree. We define the three following spaces:*

$$W_V = \bigotimes_{v\in V} W_v, \quad W_L = \bigotimes_{l\in L} W_l, \quad \widehat{W}_E = \bigotimes_{e\in E} \widehat{W}_e.$$

We call W_V the space of all possible states of the tree, W_L the space of states of leaves and $\widehat{W}_E$ the parameter space.

Definition 4.7 (The map $\widehat{\psi}$, Construction 1.5 [BW07])**.** *Let $\widehat{\psi} : \widehat{W}_E \to W_V$, be a map whose dual is defined as:*

$$\widehat{\psi}^*(\otimes_{v\in V}\alpha_v^*) = \otimes_{e\in E}(\alpha_{v_1(e)} \otimes \alpha_{v_2(e)})^*_{|\widehat{W}_e}.$$

Here the edge e is directed from the vertex $v_1(e)$ to $v_2(e)$.

The map $\widehat{\psi}$ is just a map well known to biologists that to a given choice of matrices associates the probability distribution on the set of all possible states of vertices of the tree.

Example 4.8. Let us consider the binary Jukes–Cantor model. Fix the tree with one root r and two leaves l_1 and l_2. The spaces W and $\widehat{W}$ are two dimensional. Hence the spaces W_V and $\widehat{W}_E$ are respectively 8 and 4 dimensional. The basis elements of W_V correspond to states of the variables associated to nodes of trees. Hence they can be indexed by triples (p, q, s) for $p, q, s = 0, 1$. Assume that the first element of the triple is associated to the state of r. The elements of $\widehat{W}$ are matrices of the type

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

Fix a simple tensor in $\widehat{W}_E$ represented by a pair of such matrices:

$$\begin{bmatrix} a_1 & b_1 \\ b_1 & a_1 \end{bmatrix}, \begin{bmatrix} a_2 & b_2 \\ b_2 & a_2 \end{bmatrix}.$$

To this element the morphism $\widehat{\psi}$ associates an element of W_V given as:

$$a_1a_2(0,0,0) + a_1a_2(1,1,1) + a_1b_2(0,0,1) + a_1b_2(1,1,0)$$
$$+b_1a_2(0,1,0) + b_1a_2(1,0,1) + b_1b_2(0,1,1) + b_1b_2(1,0,0).$$

Thus the map $\widehat{\psi}$ associates to a given choice of transition matrices the "probability dirstribution" on the set of all possible states of the tree. This is up to a scalar, as we assume that the root has got uniform distribution. Moreover, as we work over complex numbers and there are no probabilistic restrictions on elements of $\widehat{W}$ the map $\widehat{\psi}$ is obtained by the rule for Markov processes, but in general the elements of the image have no probabilistic meaning.

Recall that $N = V \setminus L$ is the set of nodes of a tree. We consider the map $\delta = \sum \alpha_i^* \in W^*$ that sums up all the coordinates.

Definition 4.9 (π)**.** *Let* $\pi : W_V \to W_L$ *be a map defined as* $\pi = (\otimes_{v \in L} id_{W_v}) \otimes (\otimes_{v \in N} \delta_{W_v})$. *The map* π *sums the probabilities of all the states of vertices that differ only on nodes.*

If we compose the map $\widehat{\psi}$ with π we obtain a map from $\widehat{W_E}$ to W_L. This induces a rational map:

$$\check{\psi} : \prod_{e \in E} \mathbb{P}(\widehat{W_e}) \dashrightarrow \mathbb{P}(W_L).$$

The closure of the image of this map is denoted by $\mathbb{P}(X(T, W, \widehat{W}))$**. This is the algebraic projective variety associated to the model that is the main object of study of this section.** We will also consider the affine model $X(T, W, \widehat{W})$ that is the affine cone over this variety.

5. Group-based models

The aim of this subsection is to investigate the properties of certain models. The space of transition matrices will be given as a subspace invariant under a group action. We will see under what conditions we obtain a toric variety. We will also study the properties of so obtained varieties and their connections with trees and groups. We have to point out that in this section we do not assume that a toric variety has to be normal. We only assume that a torus acts on a variety and one of the orbits is dense. This setting is most common when dealing with applications. Much information can be found in [Stu96]. The main drawback of this approach is that the varieties we consider will not be given by a fan. However, still they can be represented by polytopes, that do not have to be normal. For this reason we will often work with the character lattice M instead of the one parameter subgroup lattice N.

We will be defining objects that will depend on a tree T and a group G. For any object O if we want to stress its dependence on either T or G we write them in the indices: O_G^T. For the vector spaces on which a group G acts we use the standard notation for the subspace of invariants, by putting G in the upper index.

5.1. **General group-based models.** In our study we are mainly interested in specific models. We set the notation for general group-based models. We generalize the notions of "sockets" and "networks" introduced in [BW07]. This enables us to extend some of the results from $\mathbb{Z}_2$ to arbitrary abelian groups. We believe that these notions give a nice, unified description of the variety associated to the model.

The inspiration for this section comes from the work [ES93] of Evans and Speed who recognized a natural action of an abelian group G on S in biological case. Namely the group $G = \mathbb{Z}_2 \times \mathbb{Z}_2$ acts on $\{A, C, G, T\}$ transitively and freely. Hence from now on we assume that we have a *transitive and free action of an abelian group G on S*. In such a situation S if often called a G-torsor. The action of G on S extends naturally to the action of G on W. The fact that general group-based models give toric varieties was already observed in [ES93], [SSE93].

Definition 5.1 (A_g)**.** *For $g \in G$ let A_g be the transition matrix (equivalently the linear map) corresponding to the action of g on W.*

By choosing one element of the set S and associating it to the neutral element of G we obtain an action preserving bijection between the elements of S and G. The element associated to $a \in S$ will be denoted by g_a. Canonically the rows and columns of the transition matrix are labeled by elements of S. After fixing a bijection we can also label them with group elements, but this is not canonical. The choice of a bijection allows us also to find another basis of W, indexed by characters of G. This is done by the discrete Fourier transform.

Definition 5.2 (w_χ)**.** *Let $\chi \in G^*$ be any character of the group G. We define a vector $w_\chi \in W$ by:*

$$w_\chi = \sum_{a \in S} \chi(g_a)a.$$

Due to the orthogonality of characters the elements w_χ form a basis of W. Let us notice that although the choice of the bijection between S and G is not canonical, the one dimensional spaces spanned by w_χ are. Changing the bijection just multiplies each vector w_χ by $\chi(g)$ for some $g \in G$. In the language of representation theory W is the regular representation of G. The one dimensional spaces spanned by w_χ are of course unique irreducible one dimensional representations corresponding to all characters of G.

The group structure distinguishes also naturally a specific model, namely the vector space $\widehat{W}$. This is done as follows. We have a natural action of G on $W \otimes W$ – the action of g is just $g \otimes g$:

$$g(\sum \lambda a_1 \otimes a_2) = \sum \lambda g(a_1) \otimes g(a_2).$$

Definition 5.3 ($\widehat{W}$)**.** *Let G be an abelian group acting on the set S transitively and freely. Due to Remark 4.5 we identify* $\mathrm{End}(W)$ *with $W \otimes W$. For a general group based model we define $\widehat{W}$ as the set of fixed points of the G action on* $\mathrm{End}(W) \cong W \otimes W$.

Remark 5.4. In other words we take only such transition matrices that satisfy the following condition for any $g \in G$:
If we permute the columns and rows of a matrix with a permutation corresponding to g, then we obtain the same matrix.

Hence the parameters in the transition matrices depend only on the difference of group elements labelling the row and column of a given entry. In particular the dimension of $\widehat{W}$ is equal to $|G|$.

In general in the thesis we assume that the tree is rooted and directed away from the root. However the construction from subsection 4.1 can be easily generalized to other orientations of the edges of the tree. The reason why we make the assumption is that it simplifies the language.

Remark 5.5. One can see that if $A \in \widehat{W}$, then $A^T \in \widehat{W}$. This means that if we consider a tree T with two different orientations then the associated varieties are exactly the same. If a point is the image of some element of the parameter space with respect to a given orientation than it is also the image of an element of the parameter space with respect to the second orientation. We just have to transpose matrices that are associated to edges with different orientation.

The following elements are invariant with respect to the G action hence belong to $\widehat{W}$.

Definition 5.6 (Elements $l_\chi \in \widehat{W}$)**.** *Let χ be a character of G. We define*

$$l_\chi(w_{\chi'}) := \begin{cases} w_\chi & \chi = \chi' \\ 0 & \chi \neq \chi' \end{cases}.$$

It follows that $(l_\chi)_{\chi \in G^*}$ is a base of $\widehat{W}$. Moreover $\widehat{W}$ is equal to the space of diagonal matrices in the basis $(w_\chi)_{\chi \in G^*}$. The following Proposition gives the description of l_χ in terms of the basis associated to elements of S. We omit the proof, as it relies on basic computation.

Proposition 5.7.

$$l_\chi(a_0) = \frac{1}{|G|}\chi(g_{a_0}^{-1})w_\chi = \frac{1}{|G|}\sum_{a\in S}\chi(g_{a_0}^{-1}g_a)a.$$

□

The vectors l_χ are independent from the choice of the bijection between S and G. The element $g_{a_0}^{-1}g_a$ is a unique element of G that sends a_0 to a, hence does not depend on the bijection. The map l_χ is a projection onto the (canonical) one dimensional subspace spanned by w_χ.

Using this basis we will see that the map $\widehat{\psi}$ is injective. Hence the induced algebraic map $\prod_{e\in E}\mathbb{P}(W_e)\to\mathbb{P}(W_V)$ is given by the full Segre system. The algebraic map $\pi\circ\widehat{\psi}$ will be given by a subsystem of the Segre system. We will describe it using the notions of "sockets" and "networks". Let us start with a few lemmas. The action of G on W extends to the action of G on W_V and W_L.

Lemma 5.8. *The dimensions of G invariant subspaces of W_V and W_L are as follows:*

$$\dim W_V^G = |G|^{|V|-1},$$
$$\dim W_L^G = |G|^{|L|-1}.$$

Proof. Let us consider the basis of W_V given by $(\otimes_{v\in V}w_{\chi_v})$. The action of g in this basis is diagonal, so the space of invariant vectors is spanned by invariant elements of this basis. As $g(w_\chi)=\chi(g^{-1})w_\chi$ we obtain:

$$g(\otimes_{v\in V}w_{\chi_v}) = \otimes_{v\in V}\chi_v(g^{-1})w_{\chi_v} = \prod_{v\in W}\chi_v(g^{-1})\otimes_{v\in V}w_{\chi_v},$$

so an element $\otimes_{v\in V}w_{\chi_v}$ is invariant if and only if for any $g\in G$ we have $\prod_{v\in V}\chi_v(g)=1$. This is equivalent to the condition that $\sum_{v\in V}\chi_v$ is equal to the trivial character (we use additive notation for the group of characters G^*). From this we see that the dimension $\dim W_V^G$ is equal to the number of sequences, indexed by vertices of the tree, of characters that sum up to a neutral character. This gives us $|G^*|^{|V|-1}$ sequences and proves the first equality, as for abelian groups $|G^*|=|G|$. The proof of the second equality is the same. □

Remark 5.9. The basis $\{\otimes_{v\in V}w_{\chi_v}\}$ of W_V depends on the choice of the bijection between the set S and G. However the basis $\{\otimes_{v\in V}w_{\chi_v} : \sum_{v\in V}\chi_v=\chi_0\}$ of W_V^G is natural. Changing the bijection multiplies w_χ by $\chi(g)$ for a fixed $g\in G$. As $\sum_{v\in V}\chi_v=\chi_0$, then $(\sum_{v\in V}\chi_v)(g)=1$ and the vectors remain unchanged.

One can easily see that the image of $\widehat{W_E}$ in W_V is invariant with respect to the action of G.

Proposition 5.10. *The map $\widehat{\psi}$ is an isomorphism of vector spaces $\widehat{W_E}$ and W_V^G. It takes the base $\{\otimes_{e\in E}|G|l_{\chi_e}\}$ bijectively onto the base $\{\otimes_{v\in V} w_{\chi_v} : \sum_{v\in V}\chi_v = \chi_0\}$, where χ_0 is the trivial character.*

Proof. Using Proposition 5.7 we can see that:

$$(\otimes_{v\in V} a_v)^*(\widehat{\psi}(\otimes_{e\in E}|G|l_{\chi_e})) = \prod_{e=(v_1,v_2)\in E}(-\chi_e)(g_{a_{v_1}})\chi_e(g_{a_{v_2}}).$$

For given characters χ_e let us define characters χ_v for all v vertices of the tree as:

$$\chi_v := \sum_{(v,w)\in E}\chi_{(v,w)} - \sum_{(w,v)\in E}\chi_{(w,v)}.$$

This corresponds to summing all characters on edges adjacent to v with appropriate signs, depending on the orientation of the edge. We consider an element $\otimes_{v\in V} w_{\chi_v}$ that is clearly in the chosen basis of W_V^G as each character χ_e is taken twice with different signs, so the sum of all χ_v is the trivial character. Moreover

$$\otimes_{v\in V} w_{\chi_v} = \otimes_{v\in V}(\sum_{a\in S}\chi_v(g_a)a),$$

so $(\otimes_{v\in V} a_v)^*(\otimes_{v\in V} w_{\chi_v}) = \prod_{v\in V}\chi_v(g_{a_v})$, which proves the theorem. □

Corollary 5.11. *The following morphism:*

$$\psi : \prod_{e\in E}\mathbb{P}(\widehat{W_e}) \to \mathbb{P}(W_V^G),$$

is given by a full Segre system. In the basis from Proposition 5.10 it is given by monomials. □

Our aim will be to obtain a result similar to Proposition 5.10 for the map $\pi \circ \widehat{\psi}$. Let us notice that apart from the action of G on $W \otimes W$ given by $g \otimes g$ that allowed us to define $\widehat{W}$, we have got another action of G on $W \otimes W$ given by $g \otimes id$, where id is the identity map.

Lemma 5.12. *The action $g \otimes id$ restricts to $\widehat{W}$.*

Proof. It is enough to prove that the image of the action of $g \otimes id$ on any element that is invariant with respect to the action $g' \otimes g'$ is also invariant. Let C be any element of $\widehat{W}$.

$(g'\otimes g')((g\otimes id)C = (g'g\otimes g')(C) = (gg'\otimes g')(C) = (g\otimes id)(g'\otimes g')(C) =$

$$(g \otimes id)(C).$$

Here we used the fact that G is abelian. □

Definition 5.13 (The group G_N)**.** *We define $\rho^g_{v,e}$ for each $v \in N$, $g \in G$ and $e \in E$ as an isomorphism of the space $\widehat{W}_e$. The action on $\widehat{W}_e$ depends on e and v. If e is not adjacent to v it is the identity. If e is an outgoing edge from v it is equal to $g \otimes id$ and if e is an incoming edge it is equal to $g^{-1} \otimes id$.*

For each $v \in N$ and $g \in G$ we define an isomorphism of $\widehat{W}_E$ given by $\rho^g_v := \otimes_{e\in E}\rho^g_{v,e}$. We also define a group $G_N \subset \mathrm{End}(\widehat{W}_E)$ as a group generated by all ρ^g_v.

Remark 5.14. It is crucial to realize how $g \otimes id$ acts on elements of $\widehat{W}$ considered as morphisms. One can check that $g \otimes id(A_{g'}) = A_{g'} \circ A_{g^{-1}}$, so the action of $g \otimes id$ composes given morphism with $A_{g^{-1}}$.

To obtain a nice description of the morphism $\pi \circ \widehat{\psi}$ we need a technical lemma.

Lemma 5.15. *The group $G_N \cong G^{|N|}$. There is a base in which G_N acts diagonally on $\widehat{W}_E$.*

Proof. Using 5.14 we obtain:

$$(g \otimes id(l_\chi))(w_{\chi'}) = l_\chi A_{g^{-1}}(w_{\chi'}) =$$
$$= l_\chi A_{g^{-1}}(\sum_{a\in A}\chi'(g_a)a) = l_\chi(\sum_{a\in S}\chi'(g_a)g^{-1}a) =$$
$$= l_\chi(\sum_{a\in S}\chi'(g_a g)a) = \chi'(g)l_\chi(w_{\chi'}) = \chi(g)l_\chi(w_{\chi'}),$$

where the last equality follows from the fact that $l_\chi(w_{\chi'})$ is non zero only if $\chi = \chi'$. This proves that $g \otimes id(l_\chi) = \chi(g)l_\chi$, what proves the theorem. □

Let F be any abelian group. In our examples $F = G$ or $F = G^*$. Let us consider two groups F^E and F^N. The elements of each are associations of group elements respectively to edges and to nodes of the tree.

Definition 5.16 (Adding morphism add, projection p_v)**.** We define a morphism $add : F^E \to F^N$. Let $m \in F^E$ and $p_v : F^N \to F$ be a projection onto the component indexed by a vertex $v \in N$. The element $p_v(add(m))$ is equal to the sum of group elements associated by m to edges incoming into v minus the sum of group elements associated to edges outgoing from v.

Example 5.17. Consider $F = \mathbb{Z}_3$. Let T be a claw tree with three edges. We have

$$add : (\mathbb{Z}_3)^3 \to \mathbb{Z}_3,$$

where add is the usual sum in $\mathbb{Z}_3$.

Definition 5.18 (trivial signed sum)**.** *We say that an element $m \in F^E$ has got trivial signed sum around a vertex v if and only if $p_v(add(m))$ is the neutral element of F.*

Definition 5.19 (map add')**.** *We define a map $add' : F^L \to F$. This map sends an association of group elements to leaves to their sum.*

Remark 5.20. As in Proposition 5.10 elements of the base of $\widehat{W_E}$ are bijective with the sequences of characters indexed by edges of a tree. In other words an element of the basis of $\widehat{W_E}$ can be described as an association of a character of G to each edge of a tree. Moreover the elements of the basis of $\widehat{W_E}$ that are invariant with respect to the action of G_N are exactly such associations that the signed sum of characters around each inner vertex is the trivial character.

Lemma 5.21. *The map $\pi : W_V \to W_L$ can be described as follows:*

$$\pi(\otimes_{v\in V} w_{\chi_v}) = |G|^{|N|} \otimes_{l\in L} w_{\chi_l}$$

if all the characters χ_v for the inner vertices are trivial or zero otherwise.

Proof. First let us look at $\otimes_{v\in V} w_{\chi_v}$ in the old coordinates:

$$\otimes_{v\in V} w_{\chi_v} = \otimes_{v\in V}(\sum_{a\in S}\chi_v(g_a)a) = \sum_{(a_u)_{u\in V}\in S^V}(\prod_{v\in V}\chi_v(g_{a_v}))(\otimes_{v\in V} a_v),$$

where the sum $\sum_{(a_u)_{u\in V}\in S^V}$ is taken over all $|V|$-tuples (indexed by vertices) of basis vectors. In other words this sum parameterizes the basis of W_V made of tensor products of base vectors corresponding to elements of G. This is equal to:

$$\sum_{(a_u)_{u\in N}\in S^N}\ \sum_{(a_l)_{l\in L}\in S^L}\ \prod_{v\in N}\chi_v(g_{a_v})\prod_{f\in L}\chi_f(g_{a_f}) \otimes_{v\in N} a_v \otimes_{f\in L} a_f.$$

We see that $\pi(\otimes_{v\in V} w_{\chi_v})$ is equal to:

$$\sum_{(a_u)_{u\in N}\in S^N}\ \sum_{(a_l)_{l\in L}\in S^L}\ \prod_{v\in N}\chi_v(g_{a_v})\prod_{f\in L}\chi_f(g_{a_f}) \otimes_{f\in L} a_f =$$

$$(\prod_{v\in N}(\sum_{g\in G}\chi_v(g)))\sum_{(g_l)_{l\in L}\in G^N}\ \prod_{f\in L}\chi_f(g_l) \otimes_{f\in L} a_f.$$

The product $\prod_{u\in N}(\sum_{g\in G}\chi_u(g))$ is equal to zero unless all characters χ_u for $u \in N$ are trivial. In the latter case the product is equal to $|G|^{|N|}$. Of course

$$\sum_{(g_l)_{l\in L}\in G^N}(\prod_{f\in L}\chi_f(g_l))(\otimes_{l\in L}g_l) = \otimes_{l\in L}w_{\chi_l},$$

which proves the proposition. □

The following theorem is a direct generalization to arbitrary abelian groups of Theorem 2.12 from [BW07].

Theorem 5.22. *The spaces* (W_L^G) *and* $(\widehat{W}_E)^{G_N}$ *are isomorphic.*

Proof. One can prove it using dimension argument, but it is better to look how the basis are transformed. The base of $(\widehat{W}_E)^{G_N}$ is given by $\otimes_{e\in E}|G|l_{\chi_e}$, where the signed sum of all characters at any vertex is trivial. This, thanks to Proposition 5.10, by the morphism $\widehat{\psi}:\widehat{W}_E\to W_V$ is transformed bijectively into an independent set $\otimes_{v\in V}w_{\chi_v}$, where characters for inner vertices are trivial and the sum of all characters is trivial. Using Lemma 5.21 the image of this set by π gives the set $|G|^{|N|}\otimes_{l\in L}w_{\chi_l}$, where the characters χ_l sum up to the trivial character. The last set forms a base of W_L^G. □

Corollary 5.23. *The morphism* $\pi\circ\widehat{\psi}$ *is a toric morphism.*

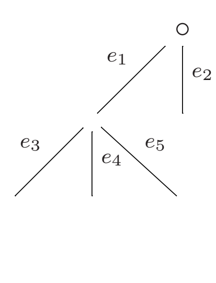
Proof. Follows from the proof of Theorem 5.22. □

Our aim is to describe the monomials that define $\pi\circ\widehat{\psi}$. This motivates the following definitions of groups of sockets and networks.

Definition 5.24 (Groups $\mathfrak{S}$ and $\mathfrak{N}$)**.** *We fix an abelian group* $F = G^*$. *The group of networks* $\mathfrak{N}$ *is the kernel of the morphism add. The group of sockets* $\mathfrak{S}$ *is the kernel of the morphism add'.*

Hence a socket is an association of characters from G^* to each leaf such that the sum of all these characters is the trivial character. A network is an association of characters from G^* to each edge such that the signed sum of characters at each inner vertex gives the trivial character.

Example 5.25. Let us consider the group $G \cong G^* = \mathbb{Z}_3$ and the following tree:



Here e_2, e_3, e_4 and e_5 are leaves. An example of a socket is an association $e_2 \to 1$, $e_3 \to 1$, $e_4 \to 2$, $e_5 \to 2$.

Example 5.26. We consider the same tree as in Example 5.25. We can make a network using the same association and extending it by $e_1 \to 2$.

Remark 5.27. Networks and sockets were introduced in [BW07] – see the discussion below. As the construction presented here directly generalizes the previous one we decided to keep the name. However, networks could also be named **group based flows**. Indeed, the condition that at each vertex the sum of elements associated to incoming edges equals the sum of elements associated to outgoing edges is the well known condition for a flow. The only difference is that we associate elements of an arbitrary group. As we will see in Proposition 5.30 there is a bijection between sockets and networks. This is similar to the theorem that for a flow the sum over all sources equals the sum over all sinks.

In [BW07] for the group $\mathbb{Z}_2$ the socket was defined as an even subsets of leaves. That corresponds to associating 1 to chosen leaves and 0 to the other leaves. The condition that the subset has got even number of elements is just the condition that the elements from the group sum up to the neutral element. We see that this definition is compatible. Networks were defined as subsets of edges such that there was an even number chosen around each inner vertex – this is also the condition of summing up to the neutral element around each inner vertex.

Let us generalize the results on sockets and networks from [BW07].

Lemma 5.28. *There are exact sequences of abelian groups:*

$$0 \to \mathfrak{N} \to (G^*)^E \stackrel{add}{\to} (G^*)^N \to 0,$$

$$0 \to \mathfrak{S} \to (G^*)^L \stackrel{add'}{\to} G^* \to 0.$$

Proof. As add and add' are surjective the lemma follows from Definition 5.24. □

Definition 5.29 (morphism fo and bi)**.** *There is a group morphism $fo : (G^*)^E \to (G^*)^L$ that forgets all the components indexed by edges not adjacent to leaves. From the diagrams in Lemma 5.28 the image of $\mathfrak{N}$ by fo is contained in $\mathfrak{S}$. We define $bi : \mathfrak{N} \to \mathfrak{S}$ to be the restriction of fo.*

There is the following diagram:

$$\begin{array}{ccccccc} 0\to & \mathfrak{N}\to & (G^*)^E & \stackrel{add}{\to} & (G^*)^N \to & 0, \\ & \downarrow^{bi} & \downarrow^{fo} & & \downarrow^{-sum} & \\ 0\to & \mathfrak{S}\to & (G^*)^L & \stackrel{add'}{\to} & G^* \to & 0. \end{array}$$

The map $-sum : (G^*)^N \to G^*$ associates to an $|N|$-tuple of characters minus their sum.

Proposition 5.30. *For any tree and any abelian group G the morphism bi that associates a socket to a network is a group isomorphism.*

Proof. Let n be a network. We know that the signed sum $p_v(add(n))$ around each inner vertex v is the neutral element. Hence $\sum_{v\in N} p_v(add(n)) = e$, where e is the neutral element. Let us consider an edge directed from v_1 to v_2, where $v_1, v_2 \in N$. Let us note that the group elements $n(v_1, v_2)$ and $n(v_1, v_2)^{-1}$ appear in $p_{v_1}(add(n))$ and $p_{v_2}(add(n))$. We see that $\sum_{v\in N} p_v(add(n)) = \sum_{l\in L} n(l)$. This means that a restriction of the network to leaves gives a socket.

Given a socket s we can define a function $n : E \to G$ inductively, starting from leaves, using the condition of summing up to the neutral element around inner edges. The only nontrivial thing is to notice that the sum around the root also gives the neutral element. This follows from the previous equality $\sum_{v\in N} p_v(add(n)) = \sum_{l\in L} n(l)$ and the fact that $p_v(add(n)) = e$ for each node v different from the root. □

Each network determines naturally an element of the basis of $(\widehat{W}_E)^{G_N}$ and each socket an element of the basis of W_L^G. The isomorphism in Theorem 5.22 just uses the natural bijection 5.30. This motivates the following definition.

Definition 5.31 (Spaces $\widetilde{W}_E$, $\widetilde{W}_L$). *We define the subspace $\widetilde{W}_E := (\widehat{W}_E)^{G_N} \subset \widehat{W}_E$. Recall that basis elements of $\widehat{W}_E$ are indexed by elements of $(G^*)^E$ as in Remark 5.20. The basis elements of $\widetilde{W}_E$ correspond to elements of $\mathfrak{N}$.*

We define the subspace $\widetilde{W}_L := W_L^G \subset W_L$. The basis elements of $\widetilde{W}_L$ correspond to associations that form a socket – cf. proof of Lemma 5.8.

Using Theorem 5.22 we know that the variety $X(T, W, \widehat{W})$ is the closure of the image of the rational map induced by $\pi \circ \widehat{\psi}$:

$$\check{\psi} : \prod \widehat{W}_e = \mathbb{C}^{|G||E|} \to \widetilde{W}_L,$$

where the coordinates of the domain are indexed by pairs (e, χ) for $e \in E$ and $\chi \in G^*$. The coordinates of the codomain are indexed by sockets (or equivalently networks). In fact the codomain is a regular representation of the group $\mathfrak{N}$. In forthcoming sections we will use the action of this group on the variety $X(T, W, \widehat{W})$.

Note that for a fixed basis of a vector space, the points with nonzero coordinates form an algebraic torus that acts on the space. Let us describe the affine map $\pi \circ \widehat{\psi}$ in toric terms.

Definition 5.32 (Lattices M_S, M_e, M_E)**.** *To each edge e we associated a vector space $\widehat{W}_e$ with the distinguished basis given by ω_χ. The points with nonzero coordinates in this basis form an algebraic torus with the action given by coordinatewise multiplication. We define M_e as the character lattice of this torus.*

The product vector space $\prod_{e \in E} \widehat{W}_e$ has got a basis induced from each $\widehat{W}_e$. The points with nonzero coordinates form an algebraic torus with the character lattice given by M_E.

The vector spaces $\widetilde{W}_E \cong \widetilde{W}_L$ have got the distinguished basis with elements corresponding to sockets. The points with nonzero coordinates form an algebraic torus with the character lattice given by M_S.

Let us note that the coordinate system on the vector space distinguishes the basis of the lattice. The basis of each lattice M_e is indexed by characters. As $M_E = \bigoplus_{e \in E} M_e$ the basis of M_E is indexed by pairs (e, χ) where e is an edge and χ a character of G. The basis elements of M_S corresponds to sockets or networks. The rational map $\check{\psi} : \prod_{e \in E} W_e \to \widetilde{W}_E \cong \widetilde{W}_L$ is an equivariant parametrization of a toric variety.

Definition 5.33 (Morphism $\widetilde{\psi}$)**.** *The morphism $\widetilde{\psi} : M_S \to M_E$ is the morphism of lattices induced by $\check{\psi}$.*

In this setting the description of $\widetilde{\psi}$ is particularly simple. Let $f_n \in M_S$ be a basis vector corresponding to a network n. The element $\widetilde{\psi}(f_n)$ will be an element of the unit cube in M_E. Let $h_{(e,\chi)} \in M_E$ be the basis vector indexed by a pair $(e, \chi) \in E \times G^*$ and let $h^*_{(e,\chi)}$ be its dual. We have:

$$h^*_{(e,\chi)}(\widetilde{\psi}(f_n)) = \begin{cases} 1 \text{ if } n(e) = \chi \\ 0 \text{ otherwise.} \end{cases}$$

We come to the **most important definition of this section**.

Definition 5.34 (Polytope P)**.** *We define the polytope $P \subset M_E$ to be the convex hull of the image of the basis of M_S by $\widetilde{\psi}$. In other words the vertices of the polytope P correspond to networks. More precisely each vertex has got 1 on coordinates indexed by pairs that form a network* *and 0 on other coordinates. Note that the polytope P is a subpolytope of a unit cube. Hence all its integer points are vertices.*

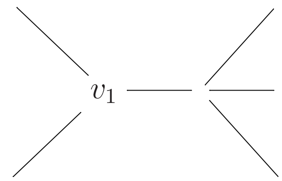
Example 5.35. Let us consider the tree T with one inner vertex and three leaves l_1, l_2 and l_3. Let $G \cong G^* = \mathbb{Z}_2$. The lattice M_S is the 4 dimensional lattice generated freely by vectors $e_{(0,0,0)}$, $e_{(1,1,0)}$, $e_{(1,0,1)}$, $e_{(0,1,1)}$ that correspond to sockets/networks on T. The lattice M_E is a 6 dimensional lattice with basis vectors $f_{(l_i,g)}$ with $1 \leq i \leq 3$ and $g \in \mathbb{Z}_2$. We have $\widehat{\psi}(e_{(a,b,c)}) = f_{(l_1,a)} + f_{(l_2,b)} + f_{(l_3,c)}$. Hence each vertex of P will have three coordinates equal to zero and three to one. Let us consider the base of M_E in the following order $f_{(l_1,0)}, f_{(l_1,1)}, \dots, f_{(l_3,0)}, f_{(l_3,1)}$. The vertex corresponding to $e_{(0,0,0)}$ is $(1,0,1,0,1,0)$. In the same order $e_{(1,1,0)} \to (0,1,0,1,1,0)$, $e_{(1,0,1)} \to (0,1,1,0,0,1)$ and $e_{(0,1,1)} \to (1,0,0,1,0,1)$. These are of course all vertices of P.

Remark 5.36. *Suppose that a tree T has got a vertex v of degree two. Let $e_1 = (u,v)$ and $e_2 = (v,w)$ be respectively an incoming and outgoing edge. Consider any network n. We have $n(e_1) = n(e_2)$. Let T' be a tree obtained from T be removing the vertex v, edges e_1, e_2 and adding an edge (u,w). We see that the polytope associated to T is isomorphic to the polytope associated to T'.*

The polytope P is the polytope associated to the toric variety $X(T,G)$. The algebra of this variety is the algebra associated to the monoid generated by P in M_E. The generating binomials of a toric ideal associated to a polytope P correspond to integral relations between integer points of this polytope, Corollary 2.6. Hence in our situation phylogenetic invariants correspond to relations between networks. Each such relation can be described in the following way. We number all edges of a tree from 1 to e. The networks are specific e-tuples of group elements. For example for the claw tree these are e-tuples of group elements summing up to the neutral element. Each relation of degree d between the networks is encoded as a pair of matrices with d columns and e rows with entries that are group elements. We require that each column represents a network. Moreover the rows of both matrices are the same up to permutation.

Example 5.37. Consider the binary Jukes-Cantor model and the following tree.

(5.1)



The leaves adjacent to v_1 have got numbers 1 and 2. We assign 3 to the inner edge. An example of a relation is given by a pair of matrices:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

The numbers 0 and 1 are treated as elements of $\mathbb{Z}_2$. Due to the definition of the socket the third row has to be the sum of both the first two and last three rows.

Note that P does not have to generate the lattice M_E.

Definition 5.38 (Lattice $\widehat{M_E}$)**.** *We define the lattice $\widehat{M_E}$ as a sublattice of M_E generated by vertices of P.*

The lattices defined so far corresponded to affine objects. A rational map from a vector space to its projectivization is well defined on points with non zero coordinates. Hence it induces a surjective morphism of tori, what corresponds to an injective morphism of character lattices.

Definition 5.39 (Degree functions $\deg_e$)**.** *Note that for a character lattice M with a distinguished basis we can define a function* $\deg : M \to \mathbb{Z}$ *that sums up coordinates. The degree of a lattice element is the degree of the monomial function associated to it. For lattices M_e the corresponding degree functions are denoted by* $\deg_e$.

Definition 5.40 (Lattices $M_{S,0}$, $M_{E,0}$ and $\widehat{M_{E,0}}$)**.** *For a lattice M_S we define $M_{S,0}$ as a sublattice of elements with the sum of coordinates equal to zero. In particular $M_{S,0}$ is the character lattice of the torus whose points are identified with points of $\mathbb{P}(\widetilde{W_E})$ with all coordinates different from zero.*

We define $M_{E,0}$ as a sublattice of M_E defined by equalities $\deg_e = 0$ *for each edge e. This is the character lattice of the torus whose points are identified with points of $\prod \mathbb{P}(W_e)$ with all coordinates different from zero.*

We define $\widehat{M}_{E,0} := M_{E,0} \cap \widehat{M}_E$. This is the character lattice of the torus whose points are identified with points of the projective toric variety $\mathbb{P}(X(T))$ with all coordinates different from zero.

Recall that the basis of the lattice M_E is indexed by pairs (e, χ) where e is an edge and χ is a character of G. Also to each such pair we can associate a one parameter subgroup in the dual of M_E. This is given as a morphism from M_E to $\mathbb{Z}$ that is the dual vector to the vector of the base of M_E that is indexed by the pair (e, χ). In particular for each leaf l and character $\chi \in G^*$ we obtain a one parameter subgroup λ_l^χ. Using the morphism dual to $\tilde{\psi} : M_S \to M_E$, for each pair (e, χ) we obtain a one parameter subgroup in the lattice dual to M_S. For each $t \in \mathbb{C}^*$ we have an action of $\lambda_l^\chi(t)$ on $\mathbb{A}^{(|L|-1)\times|G|} \supset X$. The weight of this action on the coordinate indexed by a socket s is either 0 or 1 depending on whether the socket s associates to the leaf l character χ (in this case 1) or not (in this case 0).

Remark 5.41. In [BW07] the authors considered only one one parameter subgroup for each leaf although their group had two elements. Notice however that in our notation for the group $\mathbb{Z}_2$ the weights of the action of λ_l^0 are completely determined by the weights of the action of λ_l^1 – one weights are negations of the others. In our notation the authors considered only λ_l^1.

The setting presented here, where an abelian group G acts transitively and freely on the set of states is the most well-understood. The models obtained in this way are called general group-based models. Although this definition is quite clear, the question what is a group-based model is much less obvious. This motivates the discussion of the next section 5.2.

5.2. **Notation.** In Section 5.1 we have introduced the general group-based models. The key point of the definition was that the vector space $\widehat{W}$ was given as the subspace of $\operatorname{End} W$ invariant under the action of an abelian group that acts transitively and freely on the basis of W. This setting enabled us to apply the discrete Fourier transform and associate toric varieties with the models. There are a few possibilities to generalize this construction depending on the assumptions on the group, its action on the space W and properties of the obtained associated variety.

The first idea would be to consider any action of any group on W. Even more general construction is presented in [DK09], where the vector space W may vary depending on the vertex of the tree. Such models are called equivariant models. Of course, in this case, in general one cannot apply the discrete Fourier transform, as the group G is not abelian. Moreover if the group G is small the transition matrices may be to general and the associated variety will not be toric. For example if G has got only one element it is abelian. However the model corresponding to it is just the general Markov model. The varieties associated to this model are an object of intensive study, see for example [AR08] and references therein. They are very far from being toric and establishing their properties even for the simplest tree is a great challenge. For example it is an open problem to determine the ideal in case of the tripod [BO10].

As we want to work with toric varieties it is reasonable to make further assumptions. Let us notice that the adjective "general" indicates that other group-based models should be more specific. In other words the subspace $\widehat{W}$ for a group-based model should contain specific transition matrices of a general group-based model. Thus we fix an abelian group H that acts on the space W transitively and freely. A group-based model will be obtained by requiring further conditions on the space of transition matrices.

Before stating definitions that will be used in this thesis let us present the state of art. In the literature one can find many references to group-based models [SS05], [APRS11], [PS05, p. 327]. In this setting one assumes that there is a bijection between elements of an abelian group and elements of S, as in general group-based models. One also requires that the entries of the transition matrices depend only on the difference of group elements labelling the row and the column of the given entry. However we allow the parameters for different differences to be the same – a formal definition is presented in 5.43. This is a very general definition that covers many models, like Jukes-Cantor on any number of states, 2-Kimura or any general group-based model. However for example in [APRS11] [SS05, p.460] one can also find theorems, usually originating to [ES93] that group-based models are toric. We do not believe that this is true in such a general setting. The example is presented in the Appendix 1, where after the Fourier transform we do not get monomials but polynomials. The reason for this is that equality of variables before Fourier transform does not imply equality of parameters after it. We would like to stress that the fact that Jukes-Cantor and 2-Kimura give rise to toric varieties was known before. To give a formal definition of group based-models we use a method of labellings due to Sturmfels and Sullivant [SS05, Section 3].

Definition 5.42 (Labelling function)**.** *Let Lab be any finite set and H an abelian group. A labelling function is any function $f : H \to Lab$.*

Later, we will consider special labellings, induced by group actions, that will turn out to have interesting properties.

Definition 5.43 (Group-based model)**.** *We define group-based models by specifying the space of transition matrices $\widehat{W}$. Suppose that an abelian group H acts on the set of states S transitively and freely. For any two states $s_1, s_2 \in S$ we define a morphism $p_{s_1,s_2} : \operatorname{End} W \to \mathbb{C}$. It is given by the equality $p_{s_1,s_2}(M) = (s_2^*)(M(s_1))$ where $s_1 \in W$ is an element of the basis and s_2 is an element of the dual basis. Let $g_{s_1,s_2} \in H$ be the unique element sending s_1 to s_2.*

We fix any labelling function f on H. We define $\widehat{W}$ as the largest subspace of transition matrices M satisfying the following condition:

For any $s_1, s_2, s_3, s_4 \in S$ such that $f(g_{s_1,s_2}) = f(g_{s_3,s_4})$ we have $p_{s_1,s_2}(M) = p_{s_3,s_4}(M)$.

Less formally, but more intuitively one labels the rows and columns of transition matrices with elements of H. The condition requires that entries labelled by (g_1, g_2) and (g_3, g_4) equal if $(f(g_1), f(g_2)) = (f(g_3), f(g_4))$. Notice that the space $\widehat{W}$ is obtained from the space of transition matrices of a general group-based model by specific hyperplane sections. It is important to understand that in this setting the class of group-based models is much larger than the class of general group-based models. The latter are called "general" because the space $\widehat{W}$ is the most general. They correspond to labellings that are injective. The main drawback of this setting is that varieties associated to group-based models do not have to be toric. Because of the hyperplane sections, the parametrization after the discrete Fourier transform does not have to be given by monomials. Although, as we have already said, in many cases it is. This is a motivation for the next Section 5.3. We will distinguish a class of group based-models, so called G-models. For them, we will require that the labelling is given by a specific group action. In this setting the associated varieties will be toric.

5.3. G-models.

This section contains results from [Mic11b]. Our main aim is to introduce the general framework that would include all models of interest described as group-based, but still would give rise to toric varieties. Moreover we obtain a particulary nice description of the associated polytope.

The setting of this section is sufficiently general to cover many Markov processes, in particular this will be a generalization of the results of Section 5.1. However the inspiration is the 2-Kimura model, that is the phylogenetic model in which the transition matrices are of the following type:

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

In this case, as in the previous section, we also have an abelian group $H = \mathbb{Z}_2 \times \mathbb{Z}_2$ that acts on the basis (A, C, G, T) of a four dimensional vector space W. As we have seen the fixed points of the action of H on $W \otimes W$ define the 3-Kimura model. We may however define a larger group G, namely the dihedral group of order 8, that contains H as a normal subgroup. The action of G on $W \otimes W$ defines the 2-Kimura model. Details of this construction can be found in [BDW09]. This motivates the following setting.

Let S be an n-element set of states. Let G be a subgroup[5] of $S_n = Sym(S)$ acting on S. Suppose moreover that the group G contains a *normal, abelian* subgroup H and the action of H on S is transitive and free. Elements of S once again correspond to states of vertices of a phylogenetic tree T. We define W as in Definition 4.1.

The basic difference with the abelian case is that we define elements of $\widehat{W}$ as matrices fixed not only by the action of H, but by the whole action of G. We assume that $End(W) \cong W \otimes W$, cf. Remark 4.5.

Definition 5.44. *Let*

$$\widehat{W} = \{ \sum_{a_i, a_j \in S} \lambda_{a_i, a_j} a_i \otimes a_j : \lambda_{a_i, a_j} = \lambda_{g(a_i), g(a_j)} \forall g \in G \}.$$

Remark 5.45. The characterization of $\widehat{W}$ from Remark 5.4 is still valid. However due to additional symmetries the dimension is different.

Remark 5.46. The situation of the previous section corresponds to $G = H$.

Remark 5.47. As before by choosing an element $e \in S$ we make a bijection between S and H. An element associated to $a \in S$ will be denoted by $h_a \in H$. The element e corresponds to the neutral element of H and is the index of the first row of transition matrices. Notice that the action of G on S (as permutation) will not generally be the same as the action of G on H (as a group).

We will often use the following easy observation.

Lemma 5.48. *Let $h \in H$ be an element that as a permutation sends a to b, where $a, b \in S$. Then $h = h_b h_a^{-1}$.*

[5]not necessarily abelian

Proof. Both elements send a to b, so because H acts on S freely, they have to be equal. □

Definition 5.49 (G-model)**.** *Let G be a finite group acting on a finite set S. Suppose that G contains a normal, abelian subgroup H that acts on the set S transitively and freely. A G-model is an algebraic variety $X(T, W, \widehat{W})$ for W and $\widehat{W}$ as in Definitions 4.1 and 5.44.*

Our aim is to prove that also in this generalized setting we will obtain toric varieties. We will proceed in four steps.

(i) We introduce a general method for constructing endomorphisms of W from complex functions on H. We prove that under certain conditions (namely a function should be constant on orbits of the conjugation action of G on H), the obtained endomorphism is in $\widehat{W}$. Such functions can regarded as a generalization of class functions to pairs of groups.
(ii) We prove that some sums (over the orbits of the action of G on H^*) of characters of H are functions that can define elements of $\widehat{W}$. We also notice that we obtain a set of independent vectors of $\widehat{W}$.
(iii) Using dimension arguments we prove that the set defined in step 2 is in fact a basis.
(iv) Finally, using theorems from Section 5.1, we prove, using the new coordinates, that our variety is toric.

Definition 5.50. *We define $\widehat{W_H}$ to be the vector space of matrices fixed by the action of H.*

Remark 5.51. From the previous subsection we know that the closure of the image of the map:

$$\psi : \prod_{e \in E} \mathbb{P}(\widehat{(W_H)_e}) \dashrightarrow \mathbb{P}(W_L),$$

is a toric variety. Moreover we also found the base in which the described morphism is given by monomials. As $\widehat{W} \subset \widehat{W_H}$, our aim is to prove that the restriction of the previous map is also given by monomials in certain base. We will use the base on $\widehat{W_H}$ to define the base of $\widehat{W}$.

5.3.1. *Step 1: Correspondence between functions on H and endomorphisms of W.* We are going to define some endomorphisms of W.

Definition 5.52. *Let $f : H \to \mathbb{C}$ be any function. We define:*

$$l_f = \frac{1}{|H|} \sum_{a,b \in S} f(h_a^{-1} h_b) a \otimes b.$$

Remark 5.53. Notice that due to Proposition 5.7 the previous definition is consistent with the definition of l_χ for $\chi \in H^*$. Moreover the vector l_f depends only on the function f and not the bijection between S and H, as $h_a^{-1} h_b$ is the only element from H that sends a to b.

Proposition 5.54. *Let us consider the conjugation action of G on H:*

$$(g, h) \to ghg^{-1}.$$

If f is constant on orbits of this action then $l_f \in \widehat{W}$.

Proof. Consider any element $g \in G$. We focus on two entries of the matrix l_f, namely (a_1, b_1) and (a_2, b_2), where

$$g(a_1) = a_2 \text{ and } g(b_1) = b_2.$$

These entries are from the definition of l_f respectively $f(h_{a_1}^{-1} h_{b_1})$ and $f(h_{a_2}^{-1} h_{b_2})$. Due to Remark 5.4 we want to prove that $f(h_{a_1}^{-1} h_{b_1}) = f(h_{a_2}^{-1} h_{b_2})$. Consider an element $g h_{b_1} h_{a_1}^{-1} g^{-1}$. Clearly it is an element of H (because H was a normal subgroup of G) that sends a_2 to b_2. From Lemma 5.48 we obtain:

$$g h_{b_1} h_{a_1}^{-1} g^{-1} = h_{b_2} h_{a_2}^{-1}.$$

This completes the proof, as f was constant on orbits of the conjugation action. □

5.3.2. *Step 2: Appropriate functions on H.* In the abelian case we considered characters of H. As G was equal to H, these functions were of course constant on (one element) orbits of the action of G on H. In a general case it may happen that we do not have an equality

$$\chi(ghg^{-1}) = \chi(h).$$

Of course this equality holds if a character of H extends to a character of G, but this is not always the case. If we define the vectors l_χ for $\chi \in H^*$ they may not be in $\widehat{W}$. To obtain the vectors in $\widehat{W}$ we will sum up some characters to obtain functions that satisfy the condition of Proposition 5.54. Consider the action of G on H^*:

$$\chi^g(h) = \chi(ghg^{-1}).$$

Let O be the set of orbits of this action. Elements of O give a *partition* of H^*. Let us define for each element $o \in O$ a function $f_o : H \to \mathbb{C}$.

Definition 5.55 (Function f_o). *Let* $f_o = \sum_{\chi \in o} \chi$. *Here we are summing characters as complex valued functions, not as characters, so this is the usual sum, not the product. We obtain* $l_{f_o} = \sum_{\chi \in o} l_\chi$.

Proposition 5.56. *The function* f_o *satisfies the conditions of Proposition 5.54 that is, it is constant on orbits of the conjugation action of* G *on* H.

Proof. As the action of g' is a permutation of the orbit o we have:

$$f_o(g'hg'^{-1}) = \sum_{\chi \in o} \chi(g'hg'^{-1}) = \sum_{\chi \in o} (g', \chi)(h) = \sum_{\chi \in o} \chi(h) = f_o(h).$$

□

Corollary 5.57. *The vectors* l_{f_o} *for* $o \in O$ *are in* $\widehat{W}$. *Moreover, as* l_χ *forms a basis of* $\widehat{W}_H$, *and* l_{f_o} *are sums over a partition of this basis, they are independent.*

Proposition 5.58. *Any complex function constant on orbits of* O *is a linear combination of the functions* f_o.

Proof. Let us fix a function f constant on orbits. As the characters of H span the space of all functions we know that $f = \sum_{\chi \in H^*} a_\chi \chi$. We have to prove that coefficients of χ in the same orbit are the same. Let $\chi_1^g = \chi_2$. We know that for any $h \in H$ we have

$$\sum_{\chi \in H^*} a_\chi \chi(h) = f(h) = f(ghg^{-1}) = \sum_{\chi \in H^*} a_\chi \chi(ghg^{-1}) = \sum_{\chi \in H^*} a_\chi \chi^g(h).$$

From the independence of characters we see that $a_{\chi_1} = a_{\chi_2}$ which completes the proof. □

Corollary 5.59. *The number of orbits in* O *(and so the number of vectors* l_{f_o}*) is equal to the number of orbits of the conjugation action of* G *on* H.

Proof. This follows from comparing dimensions of spaces of complex functions on H that are constant on orbits. □

5.3.3. *Step 3: Dimension of* $\widehat{W}$. We are going to prove that the dimension of $\widehat{W}$ is equal to the number of orbits $|O|$. First let us note that all coefficients of any matrix in $\widehat{W}$ (in the basis S) are determined by coefficients in the first row. This follows from Section 5.1. We see that $\dim \widehat{W}$ is equal to the number of independent parameters in the first row, that is indexed by e. The action of G imposes some conditions, namely the coefficient in the e-th row and a-th column and the coefficient in the e-th row and b-th column for $a, b \in S$ have to be equal if and only if there exists an element $g \in G$ such that:

$$g(e) = e \text{ and } g(a) = b.$$

Lemma 5.60. *The following conditions are equivalent:*

- *(i) there exists $g \in G$ that sends e to e and a to b,*
- *(ii) the elements h_a and h_b are in the same orbit with respect to the action $(g, h) = ghg^{-1}$.*

Proof. Of course h_a and h_b are in the same orbit if and only if h_a^{-1} and h_b^{-1} are in the same orbit. For the proof we concentrate on the second variant.

i)$\Rightarrow$ ii): From Lemma 5.48 we know that $gh_a^{-1}g^{-1} = h_b^{-1}$, because both elements send b to e.

i)$\Leftarrow$ ii): Suppose that $gh_a^{-1}g^{-1} = h_b^{-1}$. Let $g' = h_b^{-1}gh_{g^{-1}(b)}$. The element g' sends e to e, but $g' = gh_a^{-1}h_{g^{-1}(b)}$, hence it also sends a to b. □

Proposition 5.61. *The dimension of $\widehat{W}$ is equal to the number of orbits $|O|$.*

Proof. Classes of equal parameters in the first row of matrices in $\widehat{W}$ correspond bijectively to orbits of the action of G on H from Lemma 5.60 and remarks at the beginning of this subsection. By Corollary 5.59 this finishes the proof. □

Corollary 5.62. *The elements l_{f_o} for $o \in O$ form a basis of $\widehat{W}$.*

Proof. The vectors l_{f_o} are independent due to Corollary 5.57. The number of vectors equals the dimension of the space due to Proposition 5.61. □

5.3.4. *Step 4: G-models are toric.* Let us define a basis on $\widehat{W_e}$ consisting of vectors l_{f_o}. We consider the inclusion map $i : \widehat{W_e} \to \widehat{(W_H)}_e$, in the basis made respectively of l_{f_o} and l_χ. We know that $l_{f_o} = \sum_{\chi \in o} l_\chi$. Let us describe the morphism i in the coordinates corresponding to the basis l_{f_o} on $\widehat{W_e}$ and to the basis l_χ on $\widehat{(W_H)}_e$. Fix $\chi \in o$. We have $l_\chi^*(i(x)) = l_{f_o}^*(x)$.

This shows that the map from $\prod_{e \in E} \mathbb{P}(\widehat{W_e})$ to $\mathbb{P}(W_L)$ that parameterizes the model is also given by monomials – these are exactly monomials from Section 5.1, where we just make some variables equal to each other. Let us describe which variables are identified. We recall that variables in the abelian case correspond to networks. Fix two networks n_1 and n_2. We identify them if and only if for each edge e the characters $n_1(e)$ and $n_2(e)$ are in the same orbit of the adjoint G action.

We have got the following commutative diagram:

$$\begin{array}{ccccc} \prod_{e\in E}\mathbb{P}(\widehat{W}_e) & \to & \mathbb{P}(\widehat{W}_E) & \dashrightarrow & \mathbb{P}(W_L) \\ \downarrow & & \downarrow & & \updownarrow \\ \prod_{e\in E}\mathbb{P}(\widehat{W_{He}}) & \to & \mathbb{P}(\widehat{W_{HE}}) & \dashrightarrow & \mathbb{P}(W_L) \end{array}$$

This proves the main theorem of this section.

Theorem 5.63. *Let G be a finite group that acts faithfully on a finite set S. Let H be a normal, abelian subgroup of G. Suppose that the action of H on S is transitive and free. Let $\widehat{W}$ be the space of matrices invariant with respect to the action of G and let W be the vector space spanned freely by elements of S. Then the G-model $X(T, W, \widehat{W})$ is toric for any tree T.*

We will now describe the lattices of characters of the tori that appear in the construction. As in Section 5.1 there is a lattice M_S with basis elements corresponding to sockets and two lattices $\widehat{M_{E,H}} \subset M_{E,H}$. The letter has got basis elements indexed by pairs (e, χ) where $e \in E$ is an edge of the tree and $\chi \in H^*$ is a character.

Definition 5.64 (Lattice $M_{E,G}$)**.** *Let $M_{E,G}$ be a lattice with basis elements indexed by pairs (e, o), where $e \in E$ and o is an orbit of the adjoint action of G on H^*.*

Let $f_{e,\chi} \in M_{E,H}$ be a basis element indexed by the pair (e, χ). Let $f_{e,o} \in M_{E,G}$ be a basis element indexed by the pair (e, o). There is a natural projection $M_{E,H} \to M_{E,G}$. To an element $f_{e,\chi}$ we associate $f_{e,o}$, where $\chi \in o$. The image of a polytope $P \subset M_{E,H}$ for the general group-based model is a polytope $\tilde{P}$ that is associated to the variety representing the G model. Hence $\tilde{P}$ is a subpolytope of a unit cube. An element $\sum_{e\in E} f_{e,o_e}$ is a vertex of $\tilde{P}$ if and only if there exist characters $\chi_{o_e} \in o_e$ such that $\sum_{e\in E} f_{e,\chi_{o_e}}$ is a vertex of P. The lattice spanned by $\tilde{P}$ will be denoted by $\widehat{M_{E,G}}$. The following diagram commutes.

$$\begin{array}{ccc} M_S & \longrightarrow & M_{E,H} \\ & \searrow & \downarrow \\ & & M_{E,G} \end{array}$$

The morphisms from M_S correspond to embeddings of both models in an affine space. The surjective vertical morphism corresponds to inclusion of models. Indeed, by introducing new conditions on transition matrices for a G-model we restrict the image, hence there is a natural inclusion in a general group-based model.

We finish this section by presenting relations of G-models to labellings 5.42. From Lemma 5.60 it follows that the entries of transition matrix labelled respectively by $(h_1, h_2) \in H^2$ and $(h_3, h_4) \in H^2$ are equal if the elements $h_1^{-1}h_2$ and $h_3^{-1}h_4$ are in the same orbit of the adjoint action of G on H. Let Lab be the set of orbits of the adjoint action of G on H. The labelling function $f : H \to Lab$ associates to an element its orbit.

Definition 5.65 (m-friendly labelling, friendly labelling, [SS05, Definition 8])**.** *Let H be any abelian group and Lab any finite set. Fix a labelling function $f : H \to Lab$. For $m \geq 3$ consider the set*

$$Z = \{(g_1, \dots, g_m) \in H^m : \sum_{i=1}^{m-1} g_i = g_m\}.$$

Consider the induced map $\tilde{f} : Z \subset H^m \to Lab^m$ and denote by π_i the projection $\pi_i : H^m \to H$ onto the i-th coordinate. The function f is called m-friendly *if, for every $l = (l_1, \dots, l_m) \in \tilde{f}(Z) \subset Lab^m$,*

$$\pi_i(\tilde{f}^{-1}(l)) = f^{-1}(l_i) \quad \text{for all } i = 1, \dots, m.$$

A labelling is friendly if it is m-friendly for all $m \geq 3$.

Lemma 5.66. *The labellings for G-models are friendly.*

Proof. Fix an m-uple of orbits $(o_1, \dots, o_m)$ for the adjoint action of G on an abelian normal subgroup H. Suppose that there exist elements $h_i \in o_i$ such that $\prod_{i=1}^{m-1} h_i = h_m$. Fix any element $\tilde{h}_{i_0} \in o_{i_0}$. There is an element $g \in G$ such that $\tilde{h}_{i_0} = g h_{i_0} g^{-1}$. Consider an element $(gh_1g^{-1}, \dots, gh_mg^{-1})$. Let $\tilde{f}$ and π_i be as in Definition 5.65. Of course $\tilde{f}(gh_1g^{-1}, \dots, gh_mg^{-1}) = (o_1, \dots, o_m)$. Moreover $\pi_{i_0}(gh_1g^{-1}, \dots, gh_mg^{-1}) = \tilde{h}_{i_0}$, which proves that the labelling is friendly. □

The main reason to introduce friendly labellings is that they allow to apply a very important inductive procedure. Assuming that we are dealing with a model given by friendly labelling the variety associated to any tree T can be described in terms of the varieties associated to claw trees. The polytope associated to a tree T is a fiber product of polytopes associated to claw trees. More information can be found in Section 5.5 and articles [Sul07], [SS05, Lemma 12].

At this point we should make a remark about the difference between group elements and characters. To define the space of transition matrices for a G-model we used a G action on the space $\mathrm{End}(W)$. We considered the basis of W that corresponded to states, or by choosing a bijection to elements of an abelian group. The adjunction action of G on H allowed us to define the labelling that described a G-model. Note however that this is *not* the labelling that identifies the coordinates of the parametrization of the variety. In the latter case the variables correspond to pairs (e, χ) where $\chi \in H^*$. The labelling identifies the variables corresponding to pairs with characters on the second coordinate that are in the same orbit. Hence the set of labels is the set of orbits of the adjoint action of G on H^*. The labelling associates to a character its orbit in the adjoint action. The same proof as in the Lemma 5.66 shows that this is also a friendly labelling.

5.4. **Example of 2-Kimura model.** In this subsection we will show how the construction from the previous subsection works on Kimura models. We will also present the algorithm for constructing a polytope of a model for a given group G with a normal subgroup H. The method was described in a different language in [SS05]. The main difference (apart from the notation) is that the authors assumed the existence of a friendly labelling function, that described which characters are identified. In case of G-models we exactly know this function: it associates to a given character its orbit of the G action. This is a friendly labelling.

If $G = H$ the construction is particularly easy. The polytope has got $|G|^{|E|-|N|}$ vertices and the algorithm works in time $O(|N|(|G|^{|E|-|N|}))$ assuming that we can perform group operations in unit time.

Algorithm 1. *INPUT: A rooted tree T and an abelian group G*

OUTPUT: Vertices of the polytope associated to the toric variety representing the model for the tree T and the group G

(i) *Orient the edges of the tree from the root.*
(ii) *For each inner vertex choose one outgoing edge.*
(iii) *Make a bijection $b : G \to B \subset \mathbb{Z}^{|G|}$, where B is the standard basis of $\mathbb{Z}^{|G|}$.*
(iv) *Consider all possible associations of elements of G with not-chosen edges (there are $|G|^{|E|-|N|}$ such associations).*
(v) *For each such association, make a full association by assigning an element of G to each chosen edge in such a way that the signed sum of elements around each inner vertex gives a neutral element in G.*
(vi) *For each full association output the vertex of the polytope: $(b(g_e)_{e\in E})$, where g_e is the element of the group associated to edge e.*

Example 5.67. For the 3-Kimura model, corresponding to the group $\mathbb{Z}_2 \times \mathbb{Z}_2$, on a tree with one inner vertex and three leaves the vertices of P correspond to triples of characters of the group that sum up to a neutral character:

1) $(0,0),(0,0),(0,0)$ *2)* $(0,0),(1,0),(1,0)$ *3)* $(1,0),(0,0),(1,0)$
4) $(1,0),(1,0),(0,0)$ *5)* $(0,0),(0,1),(0,1)$ *6)* $(0,1),(0,0),(0,1)$
7) $(0,1),(0,1),(0,0)$ *8)* $(0,0),(1,1),(1,1)$ *9)* $(1,1),(0,0),(1,1)$
10) $(1,1),(1,1),(0,0)$ *11)* $(0,1),(1,0),(1,1)$ *12)* $(0,1),(1,1),(1,0)$
13) $(1,0),(1,1),(0,1)$ *14)* $(1,0),(0,1),(1,1)$ *15)* $(1,1),(0,1),(1,0)$
16) $(1,1),(1,0),(0,1)$

This in the coordinates of the lattice gives us vertices of the polytope:

1) 1,0,0,0,1,0,0,0,1,0,0,0 *2) 1,0,0,0,0,1,0,0,0,1,0,0*
3) 0,1,0,0,1,0,0,0,0,1,0,0 *4) 0,1,0,0,0,1,0,0,1,0,0,0*
5) 1,0,0,0,0,0,1,0,0,0,1,0 *6) 0,0,1,0,1,0,0,0,0,0,1,0*
7) 0,0,1,0,0,0,1,0,1,0,0,0 *8) 1,0,0,0,0,0,0,1,0,0,0,1*
9) 0,0,0,1,1,0,0,0,0,0,0,1 *10) 0,0,0,1,0,0,0,1,1,0,0,0*
11) 0,0,1,0,0,1,0,0,0,0,0,1 *12) 0,0,1,0,0,0,0,1,0,1,0,0*
13) 0,1,0,0,0,0,0,1,0,0,1,0 *14) 0,1,0,0,0,0,1,0,0,0,0,1*
15) 0,0,0,1,0,0,1,0,0,1,0,0 *16) 0,0,0,1,0,1,0,0,0,0,1,0*

The basis for $\widehat{W}$ for 3-Kimura (in previous notation vectors $l_\chi = \sum \chi(h_a^{-1}h_b) a \otimes b)$ is the following:

$$l_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, l_2 = \begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix},$$

$$l_3 = \begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}, l_4 = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}.$$

For the 2-Kimura model the four elements of H, treated as permutations decomposed into cycles, are in order:

$$(1)(2)(3)(4); (1,2)(3,4); (1,3)(2,4); (1,4)(2,3).$$

The group G is spanned by H and the transposition $(3,4)$.

If we consider the action of G on H^* we obtain three following orbits:

(i) The orbit of the trivial character contains only the trivial character. This tells us that the vector

$$f_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix},$$

is in $\widehat{W_G}$ and will be considered as the first basis vector.

(ii) The orbit of the character that associates -1 to $(1,3)(2,4)$ and $(1,4)(2,3)$ and 1 to other elements. It has got also only one element. For example let us notice that

$$\chi((3,4)(1,3)(2,4)(3,4)) = \chi((1,4)(2,3)) = -1 = \chi((1,3)(2,4)).$$

This means that the vector

$$f_2 = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}.$$

will be a basis vector of $\widehat{W_G}$.

(iii) The orbit that contains the two remaining characters. If we take their sum (as functions, not characters) we obtain a function that associates 2 to $(1)(2)(3)(4)$, -2 to $(1,2)(3,4)$ and 0 to other two elements. This gives us an element:

$$f_3 = \begin{bmatrix} 2 & -2 & 0 & 0 \\ -2 & 2 & 0 & 0 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & -2 & 2 \end{bmatrix}$$

This is the sum of two other l_χ.

We obtain $f_1 = l_1$, $f_2 = l_4$, $f_3 = l_2 + l_3$. Let $F = \{f_1, f_2, f_3\}$ and $L = \{l_1, \ldots, l_4\}$. From the previous section we know that F is the basis of $\widehat{W_G}$ and L of $\widehat{W_H}$. This can be checked directly in this example. Let us now look at the map for the tripod tree $\perp$. Elements of $\widehat{W_G}$ are special elements of $\widehat{W_H}$. We have a map:

$$(f_j^{e_i})_{j=1,\ldots,3, i=1,\ldots,3} \to (l_j^{e_i})_{j=1,\ldots,4, i=1,\ldots,3}.$$

Here j parameterizes base vectors and i parameterizes edges. Our model is the composition of this map and a model map for H. The image of the first map is a subspace given by a condition that the coordinates corresponding to $l_2^{e_i}$ and $l_3^{e_i}$ are equal for each $i = 1, \ldots, 3$. Let us see this directly.

The fixed bijection b from the Algorithm 1 is the following:

$$b(e) = (1,0,0,0), \quad b(\chi_3) = (0,1,0,0)$$
$$b(\chi_1) = (0,0,1,0), \quad b(\chi_2) = (0,0,0,1)$$

where χ_1 and χ_3 are in the same orbit. The domain of $\widehat{\psi}$ for the group H is $\{(x_1,\dots,x_{12}) : x_i \in \mathbb{C}\}$ in the order as in Example 5.67 (we fix an isomorphism with $\chi_1 = (1,0)$ and $\chi_3 = (0,1)$). This tells us that the subspace $\prod_{e\in E}(\widehat{W_G})_e$ is given by conditions $x_2 = x_3$ (the coordinates of l_2 and l_3 for $\widehat{W}_H^{e_1}$), $x_6 = x_7$, $x_{10} = x_{11}$.

This procedure works generally. After having fixed the polytope for a subgroup H, that is in the lattice M (whose coordinates are indexed by edges and characters of H) we consider a morphism from M onto the lattice M' (whose coordinates are indexed by edges and orbits of characters of H) that just assigns a character to a given orbit. This morphism sums up coordinates that are in the same orbit of the action of G on H^*. The image of the polytope P is a polytope of our model. For 3-Kimura we sum up coordinates ordered as in Example 5.67 obtaining a polytope for 2-Kimura model:

1) 1,0,0,1,0,0,1,0,0 2) 1,0,0,0,1,0,0,1,0
3) 0,1,0,1,0,0,0,1,0 4) 0,1,0,0,1,0,1,0,0
5) 1,0,0,0,1,0,0,1,0 6) 0,1,0,1,0,0,0,1,0
7) 0,1,0,0,1,0,1,0,0 8) 1,0,0,0,0,1,0,0,1
9) 0,0,1,1,0,0,0,0,1 10) 0,0,1,0,0,1,1,0,0
11) 0,1,0,0,1,0,0,0,1 12) 0,1,0,0,0,1,0,1,0
13) 0,1,0,0,0,1,0,1,0 14) 0,1,0,0,1,0,0,0,1
15) 0,0,1,0,1,0,0,1,0 16) 0,0,1,0,1,0,0,1,0

After removing double entries we get the following vertices:

1) 1,0,0,1,0,0,1,0,0 2) 1,0,0,0,1,0,0,1,0
3) 0,1,0,1,0,0,0,1,0 4) 0,1,0,0,1,0,1,0,0
5) 1,0,0,0,0,1,0,0,1 6) 0,0,1,1,0,0,0,0,1
7) 0,0,1,0,0,1,1,0,0 8) 0,1,0,0,1,0,0,0,1
9) 0,1,0,0,0,1,0,1,0 10) 0,0,1,0,1,0,0,1,0

5.5. **Further notation and applications.** In this section we will introduce notation concerning specific group-based models. We start by introducing the so called "time-reversibility" condition. This condition forces the transition matrices to be symmetric [PS05, Lema 17.2]. It is satisfied for many models considered in applications, for example for the 3-Kimura model. One can notice that a general group-based model gives rise to symmetric transition matrices if and only if all nonneutral group elements are of order two. We have to point out that in the literature often one adds to the definition of group-based models the requirement that matrices are symmetric [BDW09], [PS05, p. 328]. We do not use this convention. This leads to the following definition.

Definition 5.68 (general symmetric group-based model, symmetric group-based model)**.** *Let H be an abelian group acting transitively and freely on the set of states S. We define the general symmetric group-based model, as the model associated to the vector space $\widehat{W}$ given as the maximal space of* ***symmetric*** *matrices invariant with respect to the H action.*

Analogously we define the symmetric group-based model, as a model associated to a subspace of $\widehat{W}$ given by hyperplane sections that make some parameters of the transition matrices equal.

Symmetric group-based models do not have to be toric. For a counter example one can consider the general group-based model for $\mathbb{Z}_6$. The transition matrices are of the following type:

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

Let us consider a symmetric submodel with transition matrices of the following type:

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

After the Fourier transform we do not get a map given by monomials – see the Appendix 1. However the general symmetric group-based models always give rise to toric varieties.

Proposition 5.69. *General symmetric group-based models give rise to toric varieties.*

Proof. This is the corollary of Theorem 12.1. Suppose that H is any abelian group. We take G to be a semi direct product of H by $\mathbb{Z}_2$ where the action of $1 \in \mathbb{Z}_2$ on h gives h^{-1}. In this case the assumptions of the Theorem 12.1 are satisfied and the subspace invariant with respect to the G action gives the general symmetric group-based model. $\square$

There are two abelian groups of order 4. For $\mathbb{Z}_2 \times \mathbb{Z}_2$ the general symmetric group-based model is the same as the general group-based model and is the 3-Kimura model. For $\mathbb{Z}_4$ the general symmetric group-based model is the 2-Kimura model. Notice however that the class of general symmetric group-based models does not include Jukes-Cantor on four states that is a G-model. It can be obtained for example by an embedding of $\mathbb{Z}_2 \times \mathbb{Z}_2$ in S_4 as a normal subgroup. More precisely as $\{id; (12)(34); (13)(24); (14)(23)\}$. In conclusion we believe that the G-models form the largest known class of group-based models that give rise to toric varieties.

We would like to finish this subsection by restating the results of Sturmfels and Sullivant obtained for group-based models, in the case of G-models. We have seen that to each tree T and a G-model we can associate a polytope P. Fix a group G with a normal abelian subgroup H. The polytope P defines a projective toric variety as described in 2 and this is the variety representing the model. For general group-based model the points of P correspond to networks 5.24, that is special associations of characters of a group to edges of the tree. Using the labelling method we identify two networks if for each edge the associated characters are in the same orbit of the adjoint action of G on H^*.

Definition 5.70 (Join of two trees, split of a tree into two subtrees)**.** *Fix a tree T with an inner edge $e = (v_1, v_2)$. We distinguish two subsets S_1 and S_2 of vertices of T. The set S_1 contains all descendants of v_1, including v_1. The set S_2 contains all vertices that are not descendants of v_2, including v_2. Let T_1 and T_2 be induced subtrees of T with vertices given respectively by S_1 and S_2. Note that the edge e is a distinguished leaf both in T_1 and T_2. One can specify the roots of T_1 and T_2 arbitrarily. A canonical choice is to take respectively v_1 and v_2.*

We call the trees T_1 and T_2 the split of T. The tree T is a join of T_1 and T_2 (with a distinguished edge e).

Friendly labellings allow to describe the polytope associated to T as a fiber product of the polytopes associated to T_1 and T_2. In particular we can give a description of the polytope of any tree knowing just the polytopes associated to claw trees.

Recall that the polytope associated to the tree T is contained in the lattice $M_{E,G}$ with the basis given by pairs (k, o), where k is an edge of T and o is an orbit of the adjoint G action on H.

Fact 5.71 ([Sul07, Theorem 12], [SS05, Theorem 23])**.** *Let T be a join of two trees T_1 and T_2 with a distinguished edge e. Let M be the lattice associated to the tree T. Consider a G-model associated to a group G*

with a normal abelian subgroup H. Let M_1 and M_2 be the corresponding lattices for the trees T_1 and T_2. Let M_e be the lattice generated by the basis elements (e, o), where o is any orbit of the adjoint G action on H and e is a fixed edge. There are natural projections $p_1 : M_1 \to M_e$ and $p_2 : M_2 \to M_e$.

The polytope associated to the tree T is a fiber product over the projections p_1 and p_2 of the polytopes associated to trees T_1 and T_2. □

5.6. **Normality of G-models.** We have seen that the models associated to a group containing a normal, abelian subgroup are toric. The monomial parametrization map is sufficient for the applications. However for an algebraic geometer this would not be enough, as one would also need to prove the normality of these varieties. We will now address this problem. By normality we will mean projective normality, that is normality of the affine cone equivalent to normality of polytopes. We will see that in general one cannot expect a G-model to be normal, but in many cases it is. First let us start with a technical lemma. Different versions of it that worked only for polytopes with a unimodular cover were presented in [BW07] and [Zwi]. Recently these results were generalized in the paper [EKS11].

Lemma 5.72. *Let P_1 and P_2 be two normal polytopes contained respectively in lattices L_1 and L_2 spanned by the points of the polytopes. Suppose that we have got morphisms $p_i : L_i \to L$ of lattices for $i = 1, 2$ such that $p_i(P_i) \subset S$, where S is a standard simplex (convex hull of standard basis). Then the fiber product $P_1 \times_L P_2$ is normal in the lattice spanned by its points.*

Proof. Let $q \in n(P_1 \times_L P_2)$ for some positive integer n. Let q_i be the projection of q to L_i. Suppose q is in the lattice spanned by points of $P_1 \times_L P_2$. Hence q is equal to the sum of points that belong to $P_1 \times_L P_2$ with integral coefficients summing up to n. We know that it is in the convex hull of $n(P_1 \times_L P_2)$. Hence each q_i is the sum of points that belong to P_i with coefficients summing up to n and is in the convex hull of nP_i. This means that $q_i \in nP_i \cap L_i$. From the assumptions we obtain:

$$q_i = \sum_{j=1}^{n} v_j^i,$$

with each $v_j^i \in P_i$. We also know that $p_1(q_1) = p_2(q_2)$ and this is an element of nS. Moreover $p_i(v_j^i) \in S$. Let us notice that each element of nS can be *uniquely* written as the sum of n elements of S. This means that the collections $(p_1(v_1^1), \ldots, p_1(v_n^1))$ and $(p_2(v_1^2), \ldots, p_2(v_n^2))$ are the same up to permutation, so we can assume that $p_1(v_j^1) = p_2(v_j^2)$. Thus we can lift each pair (v_j^1, v_j^2) to a point $v_j \in P_1 \times_L P_2$ that projects respectively to v_j^1 and v_j^2. One obtains $q = \sum_{j=1}^n v_j$ which completes the proof. □

Due to Fact 5.71 the polytope associated to a tree with more then one inner vertex is the fiber product of polytopes associated to trees with strictly smaller number of inner vertices. Due to Lemma 5.72 if we want to prove normality of a polytope associated to any trivalent tree we only have to consider normality of a polytope for a tripod. More generally if we want to prove normality of a polytope associated to a tree with vertices of valency less or equal to m we have to check the normality of polytopes associated to claw trees with at most m leaves.

Proposition 5.73. *Let us consider a trivalent tree. The G-models for the abelian groups: $\mathbb{Z}_2$, $\mathbb{Z}_2 \times \mathbb{Z}_2$, $\mathbb{Z}_3$ and $\mathbb{Z}_4$ are normal.*

Proof. One can find the polytopes for the tripod and check their normality using Macaulay computer program [GS]. The proposition then follows from Lemma 5.72. □

Proposition 5.74. *The polytope of the 2-Kimura model for the tripod is not normal. Moreover the projective variety associated to the model is not normal.*

Proof. As the second part of the statement is stronger we prove only that part. The polytope of the 2-Kimura model has for vertices:

1) 1,0,0,1,0,0,1,0,0	2) 1,0,0,0,1,0,0,1,0
3) 0,1,0,1,0,0,0,1,0	4) 0,1,0,0,1,0,1,0,0
5) 1,0,0,0,0,1,0,0,1	6) 0,0,1,1,0,0,0,0,1
7) 0,0,1,0,0,1,1,0,0	8) 0,1,0,0,1,0,0,0,1
9) 0,1,0,0,0,1,0,1,0	10) 0,0,1,0,1,0,0,1,0

Let $Q = (1, 0, 0, 1, 0, 0, 1, 0, 0)$ be a vertex of P. Due to Fact 2.15 it is enough to prove that the monoid C generated by integral points of $P - Q$ is not saturated. Let us consider the cone $\tilde{C}$ that is the saturation of C. The point $L = (-1, 0, 1, -1, 0, 1, -1, 0, 1)$ is in C, as $2L$ is equal to

$$(-1, 0, 1, -1, 0, 1, 0, 0, 0)+(-1, 0, 1, 0, 0, 0, -1, 0, 1)+(0, 0, 0, -1, 0, 1, -1, 0, 1).$$

The point L is also in the lattice spanned by the vertices as

$$L = (0, 1, 0, 0, 1, 0, 0, 0, 1) - (0, 1, 0, 0, 1, 0, 1, 0, 0) + (0, 1, 0, 0, 0, 1, 0, 1, 0)$$

$$-(0, 1, 0, 0, 0, 1, 0, 1, 0) + (0, 0, 1, 0, 1, 0, 0, 1, 0) - (0, 0, 1, 0, 1, 0, 0, 1, 0).$$

However it is not an integral sum with positive coefficients of vertices of $P - Q$. Indeed each vertex of $P - Q$ with 0 on the second, fifth and eighth coordinate has got an even sum of third, sixth and ninth coordinates. However the sum of these coordinates for L is odd. □

In a joint work with Maria Donten-Bury [DBM] we managed to get further results. Using the implementation of the Algorithm 1 one can obtain the set of vertices of the polytope related to the investigated group and the tripod. We applied Polymake [GJ00] to check the normality of this polytope (in the lattice generated by its vertices). We obtained:

Computation 5.75. *The polytope associated with G-model for the tripod and the group $G = H = \mathbb{Z}_6$ is not normal. Hence the affine algebraic variety representing this model is not normal.*

In particular, the class of abelian models contains non-normal models. We believe it can be difficult to characterize the class of groups for which G-models are normal, or even to determine a big (infinite) class of normal, toric G-models. On the other hand one has the following result:

Proposition 5.76. *Let T be a phylogenetic tree and let G_1 be a subgroup of an abelian group G_2. If the variety corresponding to the tree T and group G_1 is not normal then the variety corresponding to the tree T and group G_2 is also not normal.*

Proof. Let M_i be a lattice whose basis is indexed by pairs of an edge of a tree and an element of the group G_i. The inclusion $G_1 \subseteq G_2$ gives us a natural injective morphism $f : M_1 \to M_2$. Let $P_i \subset M_i$ be the polytope associated to the model for the tree T and group G_i. Let $\tilde{M}_i \subset M_i$ be a sublattice spanned by vertices of the polytope P_i.

As P_1 is not normal in the lattice spanned by its vertices, there exists a point $x \in nP_1 \cap \tilde{M}_1$, that is not a sum of n vertices of the polytope P_1. Let us consider $y = f(x)$. The vertices of P_1 are mapped to vertices of P_2. We see that $y \in nP_2 \cap \tilde{M}_2$. If P_2 was normal in $\tilde{M}_2$ we would be able to write $y = \sum_{i=1}^{n} q_i$ with $q_i \in P_2$.

Let us notice that each point in the image $f(M_1)$ has got zero on each entry of the coordinates indexed by any edge and any element of the group $g \in G_2 \setminus G_1$. In particular y has got zero on these entries. As all entries of all vertices of P_2 are nonnegative, this proves that all entries indexed by any edge and any element of the group $g \in G_2 \setminus G_1$ are zero for q_i. However, we see that vertices of P_2 that have got all non zero entries on coordinates indexed by pairs of an edge and an element $g \in G_1$ are in the image of P_1. Hence $q_i = f(p_i)$ for $p_i \in P_1$. We see that $x = \sum p_i$, which is impossible. □

In particular we see that all abelian groups G such that $|G|$ is divisible by 6 give rise to non-normal models.

6. Description of the variety using the group action

Let us describe precisely the characters of the torus that is the dense orbit of the variety associated to the model. Let us fix a tree T and an abelian group H. We have got the following diagram:

$$\begin{array}{ccc} \widehat{\psi} : M_S & \twoheadrightarrow & \widehat{M_E} \subset M_E \\ \uparrow & \nearrow & \\ M_{S,0} & & \end{array}$$

Let us define a sublattice of M_E.

Definition 6.1 (M_{deg})**.**

$$M_{deg} = \{m \in M_E : \deg_{e_1}(m) = \deg_{e_2}(m) \quad e_1, e_2 \in E\}$$

Proposition 6.2. *The lattice $\widehat{M_E}$ is contained in the sublattice M_{deg}.*

Proof. For any basis element $b \in M_S$ corresponding to a socket and for any edge $e \in E$ we have $deg_e(\widehat{\psi}(b)) = 1$. Hence the image of any element of M_S satisfies the relations in the definition of M_{deg}. □

Of course the elements of $\widehat{M_E}$ satisfy more relations. We will describe them now.

Definition 6.3 (Morphism $\mathfrak{add}$)**.** *There is a natural surjective group morphism $\mathfrak{add} : M_E \to (H^*)^N$. For a node $n \in N$ let $p_n : (H^*)^N \to H^*$ be the projection onto the corresponding factor. Let $f_{e,\chi} \in M_E$ be a basis element corresponding to an edge e and a character $\chi \in H^*$. We define*

$$p_n(\mathfrak{add}(f_{e,\chi})) = \begin{cases} \chi_0 \text{ if and only if } n \text{ is not adjacent to } e \\ \chi \text{ if and only if } e \text{ is an edge incoming to } n \\ -\chi \text{ if and only if } e \text{ is an edge outgoing from } n, \end{cases}$$

where χ_0 is the neutral character.

We say that an element $m \in \widehat{M_E}$ has a trivial sum around a node n if and only if $p_n(\mathfrak{add}(m)) = \chi_0$.

Consider the composition $\mathfrak{add} \circ \widehat{\psi}$. Let $s \in M_S$ be a basis element corresponding to a network $\tilde{s} \in \mathfrak{N} \subset (H^*)^E$. We have $\mathfrak{add} \circ \widehat{\psi}(s) = add(\tilde{s})$. However due to Definition 5.24 we have $add(\tilde{s}) = 0$, hence $\mathfrak{add} \circ \widehat{\psi}(s) : M_S \to (H^*)^N$ is equal to zero. This means that $\widehat{M_E}$ is contained in the kernel of the morphism $\mathfrak{add}$. We will prove that there is an exact sequence:

$$0 \to \widehat{M_E} \to M_{deg} \to (H^*)^N \to 0,$$

where the last morphism is the restriction of $\mathfrak{add}$ to M_{deg}. In particular ranks of $\widehat{M_E}$ and M_{deg} are equal.

Corollary 6.4. *The dimension of the affine variety associated to the model, is equal to the dimension of the dense torus orbit that is*

$$\dim \widehat{M_E} = \dim M_{deg} = (|H|-1)|E| + 1.$$

The dimension of the projective variety equals $(|H|-1)|E|$.

We have to prove the following lemma.

Lemma 6.5. *Every element of* M_{deg} *that is in the kernel of* $\mathfrak{add}$ *belongs to* $\widehat{M_E}$.

Proof. We proceed by induction on the number of inner vertices of the tree. First let us assume that the tree T is a claw-tree with l leaves. The elements of M_{deg} can be described by sequences of length l given by elements $(\sum a^1_\chi \chi, \dots, \sum a^l_\chi \chi)$ with the condition $\sum a^1_\chi = \dots = \sum a^l_\chi$.

We prove that elements of the form $(g_1+g_2-g_1g_2-\chi_0, 0, \dots, 0)$, where $g_1, g_2 \in H^*$ are any characters are in $\widehat{M_E}$. Such an element is equal to $(g_1, g_1^{-1}, \chi_0, \dots, \chi_0)+(g_2, \chi_0, g_2^{-1}, \chi_0, \dots, \chi_0)-(g_1g_2, g_1^{-1}, g_2^{-1}, \chi_0, \dots, \chi_0)-(\chi_0, \dots, \chi_0)$. Each element of the sum is given by a socket, hence it is in $\widehat{M_E}$.

We now fix any element $(\sum a^1_\chi \chi, \dots, \sum a^l_\chi \chi) = m \in M_{deg}$ that is in the kernel of $\mathfrak{add}$. We will reduce it modulo the image of M_S to zero. Let us assume that $\sum a^1_\chi = \dots = \sum a^l_\chi = d$.

Using elements as above we can reduce m and assume that for $\chi \neq \chi_0$ the coefficient a^j_χ for each $1 \leq j \leq l$ is zero apart from one character for each j for which the coefficient can be equal to one. Precisely if there are two characters with a positive (resp. negative) coefficients we can replace them with their sum plus (resp. minus) the trivial character. If one entry is equal to $g_1 - g_2$ we add $g_2 + g_1g_2^{-1} - g_1 - \chi_0$. If there is one negative g on an entry we add $g + g^{-1} - 2\chi_0$.

In other words m is equal to $(\chi_1, \dots, \chi_l)+(d-1)(\chi_0, \dots, \chi_0)$ modulo the image of M_S. As $\sum \chi_j = \chi_0$ in H^* this element is in the image of M_S.

Now we will prove the induction step. Let us fix a tree T with at least two inner vertices. We may choose an inner edge e of T, such that cutting along the edge e we obtain two trees T_1 and T_2 (the tree

T is a join of T_1 and T_2) with strictly lower number of inner vertices. In one of the trees, say T_2, we have to choose a root – this will be a vertex belonging to the edge e. In this way all edges of T_2 are oriented as in T apart from e which has an opposite direction. An element $m \in M_{deg}$ gives us two elements $m_i \in M^i_{deg}$ for $i = 1, 2$ that are also in the kernels of $\mathfrak{add}$ for both trees. By induction hypothesis we can find two elements $s_i \in M^i_S$ which images give m_i. Let $s_i = \sum c^i_j b^i_j$ where b^i_j is the basis of M^i_S corresponding to sockets on T_i. Let us consider the multisets Z_i that are the projections of $\sum c_j b^i_j$ onto the edge e – each b_j distinguishes an element on e. The multiset Z_i has c_j elements distinguished by b^i_j with a minus sign if $c_j < 0$. Z_i is a signed multiset of characters. Let Z'_i be a multiset obtained by reductions cancelling χ with $-\chi$ in the multiset Z_i. The multiset Z'_1 is just the signed multiset of characters corresponding to m_e. The multiset Z'_2 gives the same multiset as Z'_1 if we inverse all characters. This means that we can pair together elements from Z'_1 and Z'_2 such that each pair gives rise to a socket on the tree T. The image of the sum of these sockets does not have to give m yet. We have to lift also the sockets that we cancelled by passing from Z_i to Z'_i. This is done as follows. Suppose that two sockets b_1 and b'_1 give χ on the edge e and so, b_1 and $-b'_1$ were cancelling each other in Z_1. We choose any socket s on T_2 that gives χ^{-1} on the edge e. We can glue together b_1 and s obtaining a socket (b_1, s) of the tree T and analogously (b'_1, s). The image of the difference of sockets $(b_1, s) - (b'_1, s)$ on the edges of the tree T_1 is the same as the difference of $b_1 - b'_1$ and zero on the edges belonging to T_2. In this way we obtain the sockets of T which image agrees with $\sum c_j b^i_j$ on T_i, hence is equal to m. □

Corollary 6.6. *For the tree T and the group H the dense torus orbit of the affine variety representing the model has a natural description as a quotient of the dense orbit of the torus of the parameter space by the $H^N \times (\mathbb{C}^*)^{|E|-1}$ action.*

Proof. The characters of the dense orbit of the parameter space are given by the lattice M_E. Its algebra is $\mathbb{C}[M_E] = \mathbb{C}[x^{\pm 1}_{(e,\chi)}]_{e \in E, \chi \in H^*}$. First let us describe the action of $Gr = (\mathbb{C}^*)^{|E|-1}$. We regard this torus as a subtorus of $(\mathbb{C}^*)^{|E|}$ with an additional condition that the product of all coordinates is one. Hence an element of Gr is just an association of a nonzero complex number to each edge of the tree T, such that the product of all these numbers is one. The action of Gr just multiplies $x_{(e,\chi)}$ by the complex number associated to e. In this way the invariant monomials are those whose degree with respect to each edge is the same, hence $M_E^{Gr} = M_{deg}$.

The coordinates of the group H^N are indexed by nodes. There is a natural diagonal action of the group H^N on the algebra $\mathbb{C}[M_E]$. Let us fix a node $v \in N$. The action of the $h \in H$ considered as an element of H^N, equal to h on the coordinate indexed by v and the neutral element on the other coordinates is as follows:

- for an edge e incoming to v we have $h(x_{(e,\chi)}) = \chi(h)x_{(e,\chi)}$
- for an edge e outgoing from v we have $h(x_{(e,\chi)}) = (\chi(h))^{-1}x_{(e,\chi)}$
- for the other edges $h(x_{(e,\chi)}) = x_{(e,\chi)}$.

First let us notice that elements of $\widehat{M_E}$ are invariant by the action of H^N. They are in the kernel of $\mathfrak{add}$, so the signed sum of characters around each inner vertex gives a trivial character. But the action of $h \in H \subset H^N$ just multiplies the monomial by the value on h of the character that is a signed sum of characters associated to edges adjacent to v, hence by 1. Conversely if the signed sum of characters on any $h \in H$ is 1, then the sum has to be a trivial character. So an element of M_{deg} is invariant with respect to the H^N action if and only if it is in the kernel of $\mathfrak{add}$, so by 6.5 if and only if it belongs to $\widehat{M_E}$. □

The group $H^N \times (\mathbb{C}^*)^{|E|-1}$ acts also on the algebra of the parameter space $\mathbb{C}[x_{(e,\chi)}]_{e\in E,\chi\in H^*}$. However the quotient is not equal to the variety representing the model, contrary to what is stated in [CFS08, Theorem 3.6]. Indeed the algebra of the variety is generated by the polytope (contained in the positive quadrant of M_{deg}) and is invariant by the action of $H^N \times (\mathbb{C}^*)^{|E|-1}$. However the invariant monomials of $\mathbb{C}[x_{(e,\chi)}]_{e\in E,\chi\in H^*}$ correspond to all the monomials of $\widehat{M_E}$ that are in the positive quadrant of M_E. Not all such monomials are generated by the polytope. For example for the 3-Kimura model the monomial $x_{e_0,\chi}^2 \prod_{e_i\in E} x_{e_i,e}^2$, where e is the trivial character is invariant for any χ and any distinguished edge e_0 (because $\chi+\chi = e$). This is not however the sum of any two vertices of the polytope associated to the variety.

Let us present some applications.

Corollary 6.7. *There is an exact sequence of groups:*

$$M_{S,0} \to M_{E,0} \to (H^*)^{|N|} \to 0.$$

The first map is given by $\widehat{\psi}$. The second one is the restriction of $\mathfrak{add}$ to $M_{E,0}$. □

This corollary can be applied in the identifiability problem to determine the parameters of transition matrices. We will do this in Section 11.4.1.

Let us fix an abelian group H and a tree T. We will prove that the group of networks $\mathfrak{N}$ acts on the variety $X(T,G)$. Recall that the ambient space $\widetilde{W}_L$ is a regular representation of $\mathfrak{N}$.

Proposition 6.8. *The action of the group of networks $\mathfrak{N}$ on $\widetilde{W}_L$ restricts to the variety $X(T,G)$.*

Proof. Consider the parametrization morphism $\pi_L \circ \widehat{\psi} : \mathbb{C}^{|E||H^*|} \to \widetilde{W}_L$. The basis vectors of the affine space $\mathbb{C}^{|E||H^*|}$ are indexed by pairs $(e,\chi) \in E \times H^*$. We denote the corresponding basis elements by $b_{(e,\chi)}$. For $t \in \mathbb{C}^{|E||H^*|}$ we define $t_{(e,\chi)} := b^*_{(e,\chi)}(t)$. The basis elements of $\widetilde{W}_L$ are indexed by networks $n \in (H^*)^E$. We identify a network with a sequence of characters $n = (n_e := \chi_e)_{e\in E}$ indexed by edges. Note that the group of networks acts also on the domain $\mathbb{C}^{|E||H^*|}$ by:

$$(n(t))_{(e,\chi)} := t_{(e,n_e^{-1}\chi)}.$$

It is easy to check that the morphism $\pi_L \circ \widehat{\psi}$ is equivariant. □

7. Phylogenetic invariants

The section contains results of joint work with Maria Donten-Bury [DBM]. We investigate the most important objects of phylogenetic algebraic geometry – ideals of phylogenetic invariants. The main problem in this area is to give an effective description of the whole ideal of the variety associated to a given model on a tree. Our task is to find an efficient way to compute generators of these ideals.

We suggest a way of obtaining all phylogenetic invariants of a claw tree of a G-model – more precisely we conjecture that our invariants generate the whole ideal of the variety. These, together with Fact 5.71, could provide an algorithm listing all generators of the ideal of phylogenetic invariants for any tree and for any G-model (so in particular for a general group-based model).

7.1. **Inspirations.** The inspirations for our method were the conjectures made by Sturmfels and Sullivant in [SS05]. They are still open but, as we will see, they strongly support our ideas. In particular, we will prove later that our algorithm listing the generators of the ideal works for the 3-Kimura model if we assume that the weaker conjecture made in [SS05] holds.

First we introduce some notation. As before let $K_{n,1}$ be a claw tree with n leaves. Let $\phi(G,n) = d$ be the least natural number such that the ideal associated to $K_{n,1}$ for the group based model G is generated in degree d. The phylogenetic complexity of the group G is defined as

$\phi(G) = sup_n \phi(G,n)$. Note that due to [SS05, Theorem 23] (see also [Sul07, Theorem 12]) the number $\phi(G,n)$ bounds the degree in which the ideal associated to any tree of valency at most n is generated. Based on numerical results Sturmfels and Sullivant suggested the following conjecture:

Conjecture 7.1. *For any abelian group G we have $\phi(G) \leq |G|$.*

This conjecture was separately stated for the 3-Kimura model, that is for $G = \mathbb{Z}_2 \times \mathbb{Z}_2$.

Still very little is known about the function ϕ apart from the case of the binary Jukes-Cantor model (see also [CP07]):

Proposition 7.2 (Sturmfels, Sullivant [SS05])**.** *In case of the binary Jukes-Cantor model $\phi(\mathbb{Z}_2) = 2$.* □

There are also some computational results – to the table in [SS05] presenting the computations made by Sturmfels and Sullivant a few cases can be added.

Computation 7.3. *Using 4ti2 software* [tt] *we obtained the following:*

- $\phi(\mathbb{Z}_3, 6) = 3$,
- $\phi(\mathbb{Z}_5, 4) = 4$,
- $\phi(\mathbb{Z}_8, 3) = 8$,
- $\phi(\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2, 3) = 8$.

For the 3-Kimura model we do not even know whether the function ϕ is bounded. As we will see later, this conjecture is strongly related to the one stated in the next section.

7.2. **A method for obtaining phylogenetic invariants.** We propose a method that is inspired by the geometry of the varieties we consider. First we have to introduce some notation.

Definition 7.4. *Let V_i be the set of vertices of a tree T_i for $i = 1, 2$. Let e be an inner edge of T_2 joining $v_1, v_2 \in V_2$. We say that the tree T_1 is obtained from the tree T_2 by* contraction of an edge *e if:*

- $V_1 = \{v\} \cup (V_2 \setminus \{v_1, v_2\})$,
- *for $w \in V_1 \setminus \{v\}$ a pair (v, w) is an edge of T_1 if and only if (v_1, w) or (v_2, w) is an edge of T_2,*
- *for $w \in V_1 \setminus \{v\}$ a pair (w, v) is an edge of T_1 if and only if (w, v_1) or (w, v_2) is an edge of T_2,*
- *for $w, u \in V_1 \setminus \{v\}$ a pair (w, u) is an edge of T_1 if and only if (w, u) is an edge of T_2.*

In such a situation we say that T_2 is a prolongation of T_1.

Remark 7.5. Note that these definitions are not the same as the definitions of flattenings introduced in [AR08] and further studied in [DK09].

Assume that we are in an abelian case, that is we are dealing with a general group–based model. Using Algorithm 1 one can see that vertices of the polytope correspond to sockets. As explained in Section 2 vertices of the polytope correspond to coordinates of the ambient space of the variety. In this setting the variety $X(T_1)$ associated to the tree T_1 is in a natural way a subvariety of $X(T_2)$. Notice that we can identify sockets of both varieties, as we may identify their leaves, so both varieties are contained in $\mathbb{P}^s$, where s is the number of sockets. The natural inclusion corresponds to the projection of character lattices: we forget all the coordinates corresponding to the edge joining the vertices v_1 and v_2. Details are presented in Proposition 8.1. In this setting the following conjecture is natural:

Conjecture 7.6. *The variety $X(K_{n,1})$ is equal to the (scheme theoretic) intersection of all the varieties $X(T_i)$, where T_i is a prolongation of $K_{n,1}$ that has only two inner vertices, both of them of valency at least three.*

As $X(K_{n,1})$ is a subvariety of $X(T_i)$ for any prolongation T_i one inclusion is obvious. Note also that the valency condition is made, because otherwise the conjecture would be obvious – one of the varieties that we intersect would be equal to $X(K_{n,1})$ by Remark 5.36. All T_i have got a strictly smaller maximal valency than $K_{n,1}$, so if the conjecture holds then we can inductively use Theorem 23 of Sturmfels and Sullivant [SS05] (see also Theorem 12 [Sul07]) to obtain all phylogenetic invariants for a given model for any tree of any valency, knowing just the ideal of the tripod. In such a case the ideal of $X(K_{n,1})$ is just the sum of ideals of trees with smaller valency. More precisely, if 7.6 holds then the degree in which the ideals of claw trees are generated cannot grow when the number of leaves gets bigger. This means that $\phi(G) = \phi(G, 3)$ which can be computed in many cases. In particular, the conjecture 7.6 implies all cases of the conjecture 7.1 in which we can compute $\phi(G, 3)$ – this includes the most interesting 3-Kimura model.

Remark 7.7. Let us note that varieties $X(T_1)$ and $X(T_2)$ are naturally contained in the same ambient space for any model, even if it does not give rise to toric varieties. Indeed using the construction of the variety presented in Section 4 one can see that the ambient space depends only on leaves of the tree. Hence if we can identify the leaves of trees we can identify ambient spaces of associated varieties. Thus conjecture 7.6 can help to compute the ideals of claw trees for a large class of phylogenetic models.

Of course one may argue that the conjecture 7.6 above is too strong to be true. Later we will prove it for the binary Jukes-Cantor model. We will also consider two modifications of this conjecture to weaker conjectures that can still have a lot of applications. The first modification just states that the conjecture 7.6 holds for n large enough.

Proposition 7.8. *For any G-model the conjecture 7.6 holds for n large enough if and only if the function ϕ is bounded.*

Proof. One implication is obvious. Suppose that 7.6 holds for $n > n_0$. We choose d such that the ideals associated to $K_{l,1}$ are generated in degree m for $l \leq n_0$. Using 7.6 and the results of [SS05] we can describe the ideal associated to $K_{n,1}$ as the sum of ideals generated in degree m. It follows that this ideal is also generated in degree m, so the function ϕ is bounded by m.

For the other implication let us assume that $\phi(n) \leq m$. Let us consider any binomial B that is in the ideal of the claw tree and is of degree less or equal to m. We prove that B belongs to the ideal of some prolongation of a tree T, which is in fact more than the statement of Conjecture 7.6.

Such a binomial can be described as a linear relation between (at most m) vertices of the polytope of this variety. Each vertex is given by an association of orbits of characters to edges such that there exist representatives of orbits that sum up to a trivial character. Let us fix such representatives, so that each vertex is given by n characters summing up to a trivial character.

Now the binomial B can be presented as a pair of matrices A_1 and A_2 with characters as entries. Each column of the matrices is a vertex of the polytope. The matrices have got at most m columns and exactly n rows. Let us consider the matrix $A = A_1 - A_2$, that is entries of the matrix A are characters that are differences of entries of A_1 and A_2. We can subdivide the first column of A into groups of at most $|H|$ elements summing up to a trivial character. Then inductively we can subdivide the rows into groups of at most $|H|^i$ elements summing up to a trivial character in each column up to the i-th one.

For $n > |H|^m + 1$ we can find a set S of rows of A such that the characters sum up to a trivial character in each column restricted to S, such that both the cardinality of S and of its complement are greater then 1. Note that the sums of the entries lying in a chosen column and in the rows in S are the same in A_1 and A_2. Therefore, adding to both matrices an extra row whose entries are equal to the sum of the entries in the subset S gives a representation of a binomial B on a prolongation of T. □

In particular, this proof shows that if the conjecture 7.1 of Sturmfels and Sullivant holds for the 3-Kimura model, then conjecture 7.6 also holds for this model for $n > 257$. Later we will significantly improve this estimation.

For the second modification of the conjecture 7.6 let us recall a few facts on toric varieties. Let T_1 and T_2 be two tori with lattices of characters given respectively by M_1 and M_2. Assume that both of them are contained in a third torus T with the character lattice M. The inclusions give natural isomorphisms $M_1 \simeq M/K_1$ and $M_2 \simeq M/K_2$, where K_1 and K_2 are torsion free lattices corresponding to characters that are trivial when restricted respectively to T_1 and T_2. The ideal of each torus (inside the algebra of the big torus) is generated by binomials corresponding to such trivial characters. The points of T are given by monoid morphisms $M \to \mathbb{C}^*$. The points of T_i are those morphisms that associate 1 to each character from K_i. We see that the points of the intersection $T_1 \cap T_2$ are those morphisms $M \to \mathbb{C}^*$ that associate 1 to each character from the lattice $K_1 + K_2$. Of course the (possibly reducible) intersection Y is generated by the ideal corresponding to $K_1 + K_2$. This lattice may be not saturated, but Y contains a distinguished torus T', that is one of its connected components. If K' is the saturation of the lattice $K_1 + K_2$ then the characters of T' are given by the lattice M/K'. Suppose that X is a toric variety that contains the dense torus orbit equal to T. Let X_i be the toric variety that is the closure of T_i and X' be the closure of T' in X. We call the toric variety X' the *toric intersection* of X_1 and X_2. The definition extends to a greater number of toric varieties embedded equivariantly in one toric variety. The most important case that we will use is when X is the affine space and X_i are affine toric varieties.

In the setting of 7.6 we conjecture the following:

Conjecture 7.9. *The toric variety $X(T)$ is the toric intersection of all the toric varieties $X(T_i)$.*

This conjecture differs from the previous one by the fact that we allow the intersection to be reducible, with one distinguished irreducible component equal to $X(T)$. We state this conjecture, because it can be checked using only the tori. As the points important from the biological point of view are contained in the torus (see [CFS08, Definition 2.13]), this conjecture is a weaker version of Conjecture 7.6 which is still suitable for applications. Moreover, it is quite easy to check it for trees with small enough number of leaves using computer programs. To explain it properly, let us consider the following general setting.

Assume that the tori T_i are associated to polytopes P_i and that T is just the torus of the projective space $\mathbb{P}^n \supseteq T_i$ consisting of the points with all coordinates different from zero. Let A_i be a matrix whose columns represent vertices of the polytope P_i. The characters trivial on T_i or respectively binomials generating the ideal of T_i are exactly represented by integer vectors in the kernel of A_i. The characters trivial on the intersection are given by integer vectors in the sum of lattices $\ker A_1 + \ker A_2$.

Note that the ideal of the toric intersection T' of the tori T_i in T is generated by binomials corresponding to characters trivial on T', that is by the saturation of $\ker A_1 + \ker A_2$. These binomials define a toric variety in $\mathbb{P}^n$. This variety is contained in the intersection (in fact it is a toric component) of the toric varieties that are the closures of T_i. The equality may not hold however, as the intersection might be reducible.

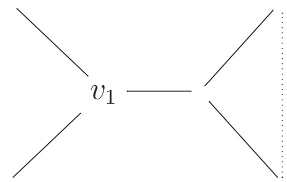
In conjecture 7.9 we have to compare two tori, one contained in the other. To do this, it is enough to compare their dimensions, that is the ranks of the character lattices. Let us note that the dimension of the intersection $T_1 \cap T_2$ is given by n minus the dimension (as a vector space) of $\ker A_1 + \ker A_2$, as it is equal to the rank of the lattice $\mathbb{Z}^n \cap (\ker A_1 + \ker A_2)$. To compute this dimension it is enough to compute the ranks of matrices A_1, A_2 and B, where B is a matrix obtained by putting A_1 under A_2 (that is, $\ker B = \ker A_1 \cap \ker A_2$). This can be done very easily using GAP ([GAP]). The results obtained for small trees will be used in the following section.

7.3. **Main Results.** To support Conjecture 7.6 let us consider the case of binary Jukes-Cantor model. This model is well understood [BW07], [CP07], [SS05].

Proposition 7.10. *Conjecture 7.6 holds for the binary Jukes-Cantor model.*

Proof. We use the same notation as in the proof of Proposition 7.8.

Let us fix a number of leaves l. We claim that we can find two special trees T_1 and T_2 for which the scheme-theoretic intersection $X(T1, \mathbb{Z}_2) \cap X(T_2, \mathbb{Z}_2)$ equals $X(K_{l,1}, \mathbb{Z}_2)$. We number the leaves from 1 to l. The trees T_1 and T_2 are isomorphic as graphs but have different leaf labelling. The topology of the trees is as follows:



For the tree T_1 the leaves adjacent to v_1 have got numbers 1 and 2. For the tree T_2 they are numbered 1 and 3. The ideal of the variety associated to a tree for the group $\mathbb{Z}_2$ is always generated in degree 2 by Proposition 7.2. Hence the generators of the ideals are of the form $n_1n_2 = n_3n_4$ where n_i for $1 \leq i \leq 4$ are coordinates corresponding to networks. Each binomial equality corresponds to a pair of matrices (M_0, M_1), with entries that are group elements, whose columns represent networks and rows are the same up to permutation. Hence each generator of the ideal of $X(K_{l,1}, \mathbb{Z}_2)$ is represented by a pair of $2 \times l$ matrices with entries from $\mathbb{Z}_2$. Moreover the sum in each column is the neutral element and rows of both matrices are the same up to permutation. As we can permute columns of each matrix we may assume that the first rows of both matrices coincide. Let us consider any such generator (M_0, M_1) in the ideal of $X(K_{l,1}, \mathbb{Z}_2)$. First suppose that the entries in the first row are the same, that is either 00 or 11. Then the relation holds both for $X(T_1)$ and $X(T_2)$. Hence we may suppose that the first row is 01 or 10. If the second row would be equal to 00 or 11 then the relation would hold for $X(T_1)$. The same reasoning holds for the third row and $X(T_2)$. Hence all three rows in both matrices are either 01 or 10. If the second (resp. third) rows are the same in both matrices then the relation holds for $X(T_1)$ (resp. $X(T_2)$). So the only possibility left is that the second and third row of M_1 are respectively the negation of the second and third row of M_0. In this case the relation does not hold in any $X(T_i)$ but we can generate it. We consider a matrix M that is equal to M_0 with the first two rows permuted. The pair (M_0, M) represents a relation in $X(T_1)$. Moreover the pair (M, M_1) represents a relation in $X(T_2)$.

□

From the proof above it follows that in fact to obtain the variety of the claw tree for the binary Jukes-Cantor model it is enough to intersect two varieties corresponding just to three subdivisions. This subdivisions correspond to S containing exactly the first and second rows or the first and third rows. Note that it is not enough to intersect two varieties corresponding to *any* prolongations – see Section 8.

Now we prove the following conditional result for the 3-Kimura model:

Proposition 7.11. *If the conjecture 7.1 of Sturmfels and Sullivant holds then the conjecture 7.6 holds for $n > 8$.*

Proof. We use the same notation as in Proposition 7.8. Consider any binomial of degree k represented by a pair of matrices (M_1, M_2) with entries given by group elements. Let $A = M_1 - M_2$, where minus is the group substraction. Matrix A has got k columns with entries from $\mathbb{Z}_2 \times \mathbb{Z}_2$. Consider A' with $2k$ columns and entries from $\mathbb{Z}_2$. The matrix A' is obtained from A by applying two projections $\mathbb{Z}_2 \times \mathbb{Z}_2 \to \mathbb{Z}_2$ to each entry. Recall that matrices M_1 and M_2 had the same rows up to permutation. This means that also after each projection the rows were the same up to permutation. Note that a difference of two vectors with entries from $\mathbb{Z}_2$ that are the same up to permutation has got always an even number of 1. Thus if we consider any row of matrix A' and either odd or even entries of this row, the number of 1 is always even.

Once again we may assume that the entries in the first row of A' are neutral elements, that is they are equal to zero. Let A'' be the matrix obtained by deleting the first row of A'. For each subset of rows of A'' we may consider a vector of length equal to the number of columns of A'', whose entries are given by sums of group elements from the subset. Note that this vector always has an even number of 1 both in even and odd columns. Because we assume conjecture 7.1, the matrix A'' has got at most 8 columns. By pigeonhole principle, if $n > 8$ then we can find two subsets of rows of A'' that are not complements of each other, such that their sum vector is the same. If we take a symmetric difference of these subsets, we obtain a strict, nonempty set S of rows of A'', summing up in each column to the neutral element. We add the first row of A' to S or its complement, so that both sets have more than one element. Thus we obtain a subdivision of the set of rows of A such that the given binomial is in the ideal of the tree corresponding to this division. □

For $n \leq 8$ we checked, using the computer programs Polymake, 4ti2, Macaulay2 and GAP, that the toric intersection of the tori of subdivisions gives the torus of the claw tree. We used the linear algebra described in the previous section. This proves that if the conjecture 7.1 holds for 3-Kimura model, then the conjecture 7.9 holds. Moreover, in all the checked cases it was enough to consider just two subdivisions. This is not a coincidence as we will prove in Section 11.

To summarize, we know that for 3-Kimura model conjecture 7.6 implies both conjectures 7.9 and 7.1 and moreover conjecture 7.1 implies 7.9 and for $n > 8$ also conjecture 7.6.

8. Interactions between trees and varieties

The ideas from the preceding sections are general. We can define an order on trees with l leaves as follows. We say that $T_1 \leq T_2$ if T_1 can be obtained from T_2 by a series of contractions of inner edges. Here by an edge contraction we mean identifying two vertices of a given edge as in Definition 7.4. The smallest tree with l leaves is the claw tree $K_{l,1}$ with one inner vertex. This is a part of a construction of the tree space [BHV01]. We fix an abelian group G.

Proposition 8.1. *If $T_1 \leq T_2$ then $X(T_1, G) \subset X(T_2, G)$.*

Proof. Although the statement is very easy we believe that the following discussion may be helpful to better understand the forthcoming sections. Both trees have got the same number of leaves, so we can make a natural bijection between their sockets. This gives an isomorphism of the ambient spaces $\widetilde{W_E}$. As $T_1 \leq T_2$ we can make an injection from the edges of T_1 to the edges of T_2. Note that a network on T_2, restricted to the edges of T_1 is a network on T_1. This gives us a projection $\pi : M_E^{T_1} \twoheadrightarrow M_E^{T_2}$. The map π simply forgets the coordinates indexed by (e, g), where e is an edge of T_2 not corresponding to an edge of T_1. Moreover the projection of P^{T_2} is equal to P^{T_1}. The following diagram commutes:

$$\begin{array}{ccc} & & M_E^{T_2} \\ & \nearrow & \downarrow \\ M_S & & \\ & \searrow & \\ & & M_E^{T_1}. \end{array}$$

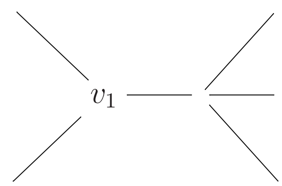
Any relation between the vertices of P^{T_2} is also a relation between the vertices of P^{T_1}. Hence any polynomial in the ideal of $X(T_2, G)$ is also in the ideal of $X(T_1, G)$. □

The surjective morphism of algebras corresponding to the inclusion of varieties is given by the restriction of the surjective morphism between $M_E^{T_2}$ and $M_E^{T_1}$ to the cones spanned by polytopes P^{T_2} and P^{T_1}.

It is natural to ask what is the relation between $X(T_0, G)$ and the scheme theoretic intersection of all $X(T, G)$ for $T_0 < T$. Conjecture 7.6 states that if there exists at least one $T > T_0$, then they are equal. So far we only know that the answer is positive for $G = \mathbb{Z}_2$ [CP07], [SS05], [DBM].

Conjecture 7.6 can be stated for *any phylogenetic model*, not necessarily given by a group[6]. In particular for a general Markov model. One would be also interested to know exactly what is an intersection of a few varieties associated to different trees. In particular how many ideals do we have to sum to obtain the ideal associated to the claw tree. One could also hope that the intersection of $X(T_1, G)$ and $X(T_2, G)$ is equal to $X(T, G)$ where T is the largest tree smaller than T_1 and T_2. Here we present a counterexample. We will prove that a scheme theoretic intersection $X(T_1, \mathbb{Z}_2) \cap X(T_2, \mathbb{Z}_2)$ does not have to be equal to $X(K_{l,1}, \mathbb{Z}_2)$ even if $K_{l,1}$ is the only tree smaller then T_1 and T_2. We consider the case of five leaves $l = 5$. The trees T_1 and T_2 are isomorphic as graphs but have different leaf labelling. Their topology is as follows:

(8.1)



For the tree T_1 the leaves adjacent to v_1 have got numbers 1 and 2. The tree T_2 is isomorphic, with two distinguished leaves labelled with 4 and 5. We consider the relation given by a pair of matrices:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

This corresponds to a generator of the ideal of $X(K_{5,1}, \mathbb{Z}_2)$. Consider any relation involving the first matrix and some other matrix M for $X(T_1)$ or $X(T_2)$. One can see that the first two rows of M must be negations of each other and the third one is 00. Hence it is impossible to generate the relation above.

9. Computational results

This section contains results of the joint work with Maria Donten-Bury [DBM]. We used the implementation of Algorithm 1.

[6] I would like to particularly thank Elizabeth Allman for discussions on this topic.

9.1. **Hilbert-Ehrhart polynomials.** The binary Jukes-Cantor model (for trivalent trees) has an interesting property, stated and proved in [BW07]: an elementary mutation of a tree gives a deformation of the associated varieties (see Construction 3.23). This implies that binary Jukes-Cantor models of trivalent trees with the same number of leaves are deformation equivalent (Theorem 3.26 in [BW07]). As it was not obvious what to expect for other G-models, we computed Hilbert-Ehrhart polynomials, which are invariants of deformation, in some simple cases.

Let us recall basic facts about Hilbert polynomials for projective toric varieties. Suppose that our variety corresponds to a polytope $P \times \{1\}$ contained in the lattice M spanned by its integral points. There are two functions that one can associate to the polytope P.

(i) Let $h : \mathbb{N} \to \mathbb{N}$ be a function. Let $h(n)$ equal the number of points in the monoid generated by $P \times \{1\}$ with the last coordinate equal to n. We call h the Hilbert function.

(ii) Let $e : \mathbb{N} \to \mathbb{N}$ be a function. Let $e(n)$ equal the number of integral points in nP, or equivalently in $n(P \times \{1\})$. We call e the Ehrhart function.

The function e is a polynomial function, thus we call it the Ehrhart polynomial. The function h is a polynomial function for large enough values. The polynomial $\tilde{h}$ such that for n large enough $\tilde{h}(n) = h(n)$ is called the Hilbert polynomial. From the definition of normal polytope 2.13 we see that the Hilbert function equals the Ehrhart polynomial if and only if P is normal, that is if and only if the associated variety is projectively normal. The associated variety is normal if and only if the Hilbert polynomial equals the Ehrhart polynomial [Stu96, Theorem 13.11]. In this case we call it the Hilbert-Ehrhart polynomial.

9.1.1. *Numerical results.* We checked models for two different trees with six leaves (this is the least number of leaves for which there are non-isomorphic trees, exactly two), the *snowflake* and the *3-caterpillar*. The most interesting ones were the cases of the biologically meaningful 2-Kimura and 3-Kimura models.

To determine the Hilbert-Ehrhart polynomial of a G-model we compute the number of lattice points in multiples of its polytope. Even if it is not possible to get enough data to determine the polynomials (eg. because numbers are too big), sometimes we can say that polynomials for two models are not equal, because their values for some n are different.

Before we completed our computations, Kubjas computed numbers of lattice points in the third dilations of the polytopes for 3-Kimura model on the *snowflake* and the *3-caterpillar* with 6 leaves and got 69248000 and 69324800 points respectively [Kub10]. Thus she proved that varieties associated with these models are not deformation equivalent.

Our computations confirm her results as for the 3-Kimura model and also give the following

Computation 9.1. *The varieties associated with 2-Kimura models for the snowflake and the 3-caterpillar trees have different Ehrhart polynomials. In the second dilations of the polytopes there are 56992 lattice points for the snowflake and 57024 for the 3-caterpillar.*

Also the pairs of varieties associated with G-models for the snowflake and the 3-caterpillar trees and

(i) $G = H = \mathbb{Z}_3$,
(ii) $G = H = \mathbb{Z}_4$,
(iii) $G = H = \mathbb{Z}_5$,
(iv) $G = H = \mathbb{Z}_7$

have different Hilbert-Ehrhart polynomials and therefore are not deformation equivalent. (For these pairs G-models are normal, which can be checked using Polymake.) The precise results of the computations are presented in the Appendix 2.

In the cases of

(i) $G = H = \mathbb{Z}_8$,
(ii) $G = H = \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$,
(iii) $G = H = \mathbb{Z}_9$

the varieties have got different Hilbert functions. We were not able to check if they are normal, however if they are then the Hilbert-Ehrhart polynomials are different.

9.2. **Some technical details.** The first attempt to compute numbers of lattice points in dilations of a polytope was the direct method: constructing the list of lattice points in nP by adding vertices of P to lattice points in $(n-1)P$ and reducing repeated entries. This algorithm is not very efficient, but (after adding a few technical upgrades to the implementation) we were able to confirm Kubjas' results [Kub10]. However, this method does not work for non-normal polytopes. As we planned to investigate 2-Kimura model, we had to implement another algorithm.

The second idea is to compute inductively the relative Hilbert polynomials, i.e. number of points in the n-th dilation of the polytope intersected with the fiber of the projection onto the group of coordinates that correspond to a given leaf. Our approach is quite similar to the methods used in [Kub10] and [Sul07].

First we compute two functions for the tripod. Let $P \subset \mathbb{Z}^{3m} \cong \mathbb{Z}^m \times \mathbb{Z}^m \times \mathbb{Z}^m$ be the polytope associated to a tripod. Let $pr_i : \mathbb{Z}^{3m} \cong \mathbb{Z}^m \times \mathbb{Z}^m \times \mathbb{Z}^m \to \mathbb{Z}^m$ be the projection onto the i-th group of coordinates. We distinguish one edge of the tripod corresponding to the third group of coordinates in the lattice. Let f be a function such that $f(a)$ for $a = (a_1, \ldots, a_m) \in \mathbb{Z}^m$ is the number of lattice points in $(a_1+\cdots+a_m)P$ that project to a by pr_3. We compute $f(a)$ for sufficiently many values of a to proceed with the algorithm.

Example 9.2. The polytope P for the binary Jukes-Cantor model has the following vertices:

$$v_1 = (0, 1, 0, 1, 0, 1),$$

$$v_2 = (0, 1, 1, 0, 1, 0),$$

$$v_3 = (1, 0, 0, 1, 1, 0),$$

$$v_4 = (1, 0, 1, 0, 0, 1).$$

These are the only integral points in P. In this case $f(1, 0) = 2$ because there are exactly two points, $(1, 0, 0, 1, 1, 0)$ and $(0, 1, 1, 0, 1, 0)$, that are in $1P = P$ and project to $(1, 0)$ via the third projection.

The function f will be our base for induction. Next, we need to compute the number of points in the fiber of a projection onto two distinguished leaves. Let g be a function such that $g(a, b)$ for $(a, b) = (a_1, \ldots, a_m, b_1, \ldots, b_m) \in \mathbb{Z}^m \times \mathbb{Z}^m$ is the number of lattice points in $(a_1 + \cdots + a_m)P$ that project to a by pr_3 and to b by pr_2. We compute $g(a, b)$ for sufficiently many pairs (a, b) to proceed with the algorithm.

Let T be a tree with a corresponding polytope P and a distinguished leaf l. Let h be a function such that $h(a)$ for $a = (a_1, \ldots, a_m) \in \mathbb{Z}^m$ is equal to the number of points in the fiber of the projection corresponding to leaf l of $(a_1 + \cdots + a_m)P$ onto a. We construct a new tree T' by attaching a tripod to the chosen leaf l of T. We call T' a join of T and the tripod. The chosen leaf of T' will be one of the leaves of the attached tripod. As proved in [BW07], [SS05], [Mic11b], [Sul07] (depending on the model), the polytope associated to a join of two trees is a fiber product of the polytopes associated to these trees. Thus we can calculate the function h' for T' by the following rule: $h'(a) = \sum_b g(a, b)h(b)$, where the sum is taken over all $b \in \mathbb{Z}^m$ such that $g(a, b) \neq 0$.

This allows us to compute inductively the relative Hilbert polynomial. The last tripod could be attached in the same way. Then one obtains the Hilbert function from relative Hilbert functions simply by summing up over all possible projections. However, it is better to do the last step in a different way.

Suppose that as before we are given a tree T with a distinguished leaf l and a corresponding relative Hilbert function h. We compute the Hilbert function of the tree T' that is a join of the tree T and a tripod using the equality $h'(n) = \sum_a f(a)h(a)$, where $a = (a_1, \dots, a_m)$ and $\sum a_i = n$. The function f is the basis for induction introduced above.

Thus, decomposing the *snowflake* and the *3-caterpillar* trees to joins of tripods, we can inductively compute (a few small values of) the corresponding Hilbert functions. This method works also for non-normal models, if only the Hilbert function for the tripod can be computed. In particular, for 2-Kimura model the computations turned out to be possible, because its polytope for the tripod is quite well understood at least to describe fully its second dilation. More precisely the points of the polytope and the point constructed in the proof of Proposition 5.74 generate the cone over the polytope. This way we obtained the results of 9.1.

10. Categorical setting

The aim of this section is to present a category GM of G-models and its connections with other categories. As an application of the theory we will present a proof of Conjecture 7.9 for the 3-Kimura model.

10.1. **Category of G-models.** A G-model is the following set of data:

- a tree T
- a group G
- a normal, abelian subgroup $H \lhd G$.

Let us remind that the group G acts on the characters H^* by adjunction $\chi^g(h) = \chi(ghg^{-1})$. This motivates the following definition.

Definition 10.1 (Compatible morphism of subgroups)**.** *Let us fix two pairs (H_i, G_i) where H_i is an abelian, normal subgroup of G_i for $i = 1, 2$. We say that a morphism $f : H_1 \to H_2$ is* compatible *if the dual morphism $f^* : H_2^* \to H_1^*$ preserves the orbits of groups G_i. That is for any pair of characters χ, $\chi' \in H_2^*$ in the same orbit of the G_2 action the images $f^*(\chi)$ and $f^*(\chi')$ are in the same orbit of the G_1 action.*

Remark 10.2. Let us note that in the abelian case, that is $G_i = H_i$ all morphisms are compatible. Note also that compatible does not mean that the orbits of the adjoint action of G_i on H_i are preserved by f.

Now we are ready to state the definition of the category GM.

Definition 10.3 (Category GM of G-models)**.** *Let GM be a category where the objects are triples (T, G, H), as described above. A morphism in GM between (T_1, G_1, H_1) and (T_2, G_2, H_2) will be a pair of maps $f : T_1 \to T_2$ and $g : H_1 \to H_2$. Here g is a compatible group morphism and f is a morphism of graphs, that is an isomorphism onto the image.*

We define the category of polytopes $Poly$.

Definition 10.4 (Category $Poly$ of polytopes)**.** *Let $Poly$ be a category where objects are pairs $(P, \widehat{M})$, where $\widehat{M}$ is a lattice and P a lattice polytope, that spans the whole lattice. A morphism from $(P_1, \widehat{M_1})$ to $(P_2, \widehat{M_2})$ is a lattice morphism from $\widehat{M_1}$ to $\widehat{M_2}$ that takes points of P_1 to points of P_2.*

10.1.1. *Construction of the functor F.* Our aim is to define a contravariant functor F from the category GM to the category $Poly$. We have already done this on objects; to a tree T and a group $G \triangleright H$ we associate a pair $(\tilde{P}, \widehat{M_{E,G}})$ as in the discussion after Definition 5.64. Let us define the functor F on morphisms. Suppose that we have a morphism in GM, that is a pair of morphisms $f : T_1 \to T_2$ and $g : H_1 \to H_2$. Let $P_i \subset \widehat{M_i}$ be the polytope and the lattice corresponding to the tree T_i with the group $G_i \triangleright H_i$. Let also M_i be the lattice with the basis elements indexed by (e, o) – cf. Definition 5.64 – where e is an edge of T_i and o an orbit in H_i^*. The lattice M_i contains the lattice $\widehat{M_i}$. Morphism g gives us a morphism of characters $g^* : H_2^* \to H_1^*$. We proceed in two steps.

Step 1. The group morphism.

We consider a polytope $\widetilde{P}$ associated to the tree T_2 with the group $G_1 \triangleright H_1$. Let M' be the lattice associated to this tree. The basis of M' is indexed by pairs (e, o), where e is an edge of T_2 and o is an orbit in H_1^*. Using the morphism g^* we can define a morphism $m : M_2 \to M'$ by sending a character over an appropriate edge to its image by g^*. Of course the points of P_2 are mapped to the points of $\widetilde{P}$, because the condition of summing up to a trivial character is preserved by the action of the morphism and so are the orbits. This means that we can restrict m to the morphism $m' : \widehat{M_2} \to \widehat{M'}$, where $\widehat{M'}$ is a sublattice of M' spanned by points of $\widetilde{P}$. This gives us a morphism in $Poly$ from $(P_2, \widehat{M_2})$ to $(\widetilde{P}, \widehat{M'})$.

Step 2. The tree morphism.

Here we forget the coordinates corresponding to edges that are not in the image. Of course the condition of summing up to a trivial character around vertices that are in the image is preserved.

Remark 10.5. In the "big" lattice M_i our morphism has got always a form of:

-first summing up coordinates (that correspond to the orbits of characters in the inverse image of a given orbit)

-second forgetting coordinates indexed by pairs (e, o), where e is an edge not in the image of the morphism of trees.

However, each time we have to remember about smaller lattices and the fact that the image of our polytope may not span the whole "small" lattice $\widehat{M_i}$ (if the morphism g^* is not surjective).

Next we consider a covariant functor from *Poly* to the category of algebras. We associate to a polytope $P \subset M$ an algebra, that is defined as a monoid algebra for the submonoid of $\mathbb{Z} \times M$, spanned by $\{1\} \times P$. The contravariant functor from the category of algebras to the category of varieties is well known. In the toric case it was described in Section 2. Composing all we obtain a covariant functor from the category *GM* to the category of toric varieties.

Remark 10.6. Note that first we associate to a polytope $P \subset M$ an algebra, that is defined as an algebra associated to the submonoid of $\mathbb{Z} \times M$, spanned by $\{1\} \times P$. This is not necessarily a cone, as P does not have to be normal. Then we associate to this algebra a variety. This does *not* have to be a toric variety associated to a polytope in the sense of [Ful93], [CLS] – that construction always gives a normal variety.

10.2. **Morphisms of groups and rational maps of varieties.** The motivation for this subsection is the following observation: if we look at graded algebras (or respectively projective varieties), then the map of graded algebras obtained from the map of polytopes in general gives us only a rational map of varieties. However we obtain a morphism for example if the map of graded algebras is surjective.

This observation allows us to define a functor G from *GM* to *Proj*, where *Proj* is the category of embedded projective varieties with rational morphisms. The functor G is a composition of the functor F from the previous section, a natural functor that associates to a polytope a graded algebra generated in degree one (cf. Remark 10.6) and a well-known functor that associates to a graded algebra a projective variety [Har77, p. 76].

In particular let us consider the abelian case, that is a full subcategory $GM^{ab} \subset GM$ containing all objects for which $G = H$. Then to each morphism of groups $G_1 \to G_2$ we can associate a rational morphism of projective varieties. Note that this is a well defined morphism of affine cones over the projective varieties. More information on the abelian case can be found in Section 10.3.

Let us consider a G-model (T_1, G_1, H_1). The affine variety associated to this model can be realized as a subvariety of $\mathbb{A}^s$, where s is the number of vertices of the associated polytope. Notice that the morphism between two G-models that is an identity on trees induces an equivariant morphism of ambient spaces.

The following description of the morphism between the varieties will be useful in the following sections. Consider two G-models (T, G_1, H_1) and (T, G_2, H_2). Let $f : H_1 \to H_2$ be a compatible morphism that, together with an identity on T, induces a morphism of G-models. Let P_1 and P_2 be the polytopes associated to corresponding models. As in Definition 5.64 the polytope P_i is contained in the lattice M_{E,G_i} with basis elements indexed by pairs (e, o) for e an edge of T and o an orbit of G_i action on H_i^*. The vertices of P_i correspond also to coordinates of the affine space embedding the affine variety associated to a model. Note that f^* induces a morphism $m : M_{E,G_2} \to M_{E,G_1}$. Each vertex of P_2 can be *represented* by an association of characters from H_2^* to edges. The morphism m is simply an application of f^* to the representants.

Proposition 10.7. *Consider the setting described above. Let s_i be the number of vertices of P_i and let $\mathbb{A}^{s_i}$ be the affine space embedding the affine variety associated to (T, G_i, H_i). The morphism of G-models induces the morphism of affine spaces $\tilde{m} : \mathbb{A}^{s_1} \to \mathbb{A}^{s_2}$. This is an equivariant morphism induced by a restriction of m to positive quadrants. Precisely, let e_v^* be the coordinate corresponding to a vertex $v \in P_2$. We have $e_v^*(\tilde{m}(x)) = e_{m(v)}^*(x)$.* □

Let us now fix morphisms from (T, G_i, H_i) to (T, G_0, H_0) that are identities on trees and are given by compatible group morphisms $f_i : H_i \to H_0$. Let P_i be the polytope associated to the model (T, G_i, H_i). Let M_{S_i} be the lattice with basis elements indexed by vertices of P_i. We obtain a morphism of lattices $m : M_{S_0} \to \prod M_{S_i}$. Let s_i be the dimension of M_{S_i}. Let $p_j : \prod M_{S_i} \to M_{S_j}$ be the projection to the j-th factor.

Remark 10.8. The morphism of lattices described above corresponds to the morphism of ambient spaces $\prod \mathbb{A}^{s_i} \to \mathbb{A}^{s_0}$. It can be described in coordinates as follows:

A coordinate corresponding to a vertex $v_0 \in P_0$ is a product of all coordinates corresponding to vertices $p_j(m(v_0)) \in P_j$.

10.3. **Abelian case.** In this section we will establish connections between morphisms of abelian groups and morphisms of corresponding varieties. Once again our main aim is application in geometry. We are building the set up of the next section. That is why we restrict to special cases. This reduces the complexity of the language but still gives a geometric insight. Let us fix a tree T.

Let $f : G_1 \to G_2$ be a morphism of abelian groups. It induces morphisms of groups of sockets $\mathfrak{S}^{G_2} \to \mathfrak{S}^{G_1}$. This gives the following commutative diagram :

$$\begin{array}{ccc} M_{S,G_1} & \longrightarrow & \widehat{M}_{E,G_1} \\ \uparrow & & \uparrow \\ M_{S,G_2} & \longrightarrow & \widehat{M}_{E,G_2} \end{array}$$

Hence the morphism $\widehat{M}_{E,G_1} \to \widehat{M}_{E,G_2}$ of character lattices restricts to cones over polytopes. This gives a morphism of algebras of associated varieties. The morphism $M_{S,G_2} \to M_{S,G_1}$ restricts to positive quadrants of both lattices. Hence we get a morphism of ambient spaces $\widehat{f} : \widehat{W}_{L,G_1} \to \widehat{W}_{L,G_1}$ compatible with morphism of varieties $\widehat{f'} : X(T, G_1) \to X(T, G_2)$. This gives a covariant functor from the category of abelian groups to the category of embedded affine toric varieties. Moreover if f^* is injective (resp. surjective) then $\widehat{f'}$ is dominant (resp. injective). The second assertion is an easy exercise. We also need the following setting. Suppose that we have morphisms $\phi_i : G_i \to G$ for $i = 1, \dots, m$. Just as above this gives us a morphism of embedded varieties $f_i : X(T, G_i) \to X(T, G)$. Let P be the polytope associated to $X(T, G)$ and let P_i be the polytope associated to $X(T, G_i)$. Consider the induced morphism $\tilde{f} : \widehat{M}_{E,G} \to \prod \widehat{M}_{E,G_i}$. If the product $f_1^* \times \cdots \times f_m^* : G^* \to \prod G_i^*$ is surjective, then $\tilde{f}$ restricted to the monoid spanned by P is surjective onto the monoid spanned by $\prod P_i$. However, in general, if the product $f_1^* \times \cdots \times f_m^*$ is injective then the restriction of $\tilde{f}$ to the monoid generated by P does not have to be injective. If $\tilde{f}$ is injective, than it induces a dominant map from the product $\prod X(T, G_i)$ to $X(T, G)$.

11. Applications to the 3-Kimura model, part 1

Our aim is to prove the Conjecture 7.9 for $G = \mathbb{Z}_2 \times \mathbb{Z}_2$.

Conjecture 11.1. *The dense torus orbit of the toric variety* $X(K_{l,1}, \mathbb{Z}_2 \times \mathbb{Z}_2)$ *is the intersection of the dense torus orbits of the varieties* $X(T, \mathbb{Z}_2 \times \mathbb{Z}_2)$*, where* T *is any tree with* l *leaves different from the claw tree.*

Note that all dense torus orbits are contained in the dense torus orbit O of the projective (or affine) ambient space. In the algebraic set O all the considered orbits are closed subschemes. Hence Conjecture 11.1 can be regarded in a set-theoretic or in a scheme-theoretic version. Both of them are equivalent. This follows for example from a more general statement [ES96, Corollary 2.2] and is particularly simple in toric case. However because the proofs of both versions are basically the same for $G = \mathbb{Z}_2 \times \mathbb{Z}_2$ we have decided to include both. Moreover this also gives an idea how the elements of the ideal of $X(K_{l,1}, \mathbb{Z}_2 \times \mathbb{Z}_2)$ can be generated by elements of ideals of $X(T, \mathbb{Z}_2 \times \mathbb{Z}_2)$.

The main idea of the proof is to extend the results known for binary models to the 3-Kimura model. The binary model is very well understood and has a lot of special properties [BW07]. In particular from 7.10 we know that Conjecture 7.6 holds for $G = \mathbb{Z}_2$. As G is abelian we will be identifying G with G^*. In particular, in this subsection we assume that networks and sockets associate to edges group elements, not characters. This convention does not change anything, but simplifies the language.

We have got three natural projections $f_i : \mathbb{Z}_2 \times \mathbb{Z}_2 \to \mathbb{Z}_2$ for $i = 1, \dots, 3$. The map $f_1 \times f_2 \times f_3 : \mathbb{Z}_2 \times \mathbb{Z}_2 \to \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ is injective. Moreover it induces a dominant map from the product of three binary models onto the 3-Kimura model. This map is the key tool that will allow us to transfer some of the properties from the binary model to the 3-Kimura model. Unfortunately the map is not surjective, but just dominant. We can projectivise the varieties, but then we get a rational map. It turns out that a combine use of both of the maps allows to derive the main theorem.

Let $f_i^* : M_{S,\mathbb{Z}_2 \times \mathbb{Z}_2} \to M_{S,\mathbb{Z}_2}$ be a morphism of lattices induced by f_i. More precisely a socket that associates to an edge e a group element $g \in \mathbb{Z}_2 \times \mathbb{Z}_2$ is send to a socket that associates to e and element $f_i(g) \in \mathbb{Z}_2$. Let $i : M_{E,\mathbb{Z}_2 \times \mathbb{Z}_2} \to M_{E,\mathbb{Z}_2} \times M_{E,\mathbb{Z}_2} \times M_{E,\mathbb{Z}_2}$ be the morphism of lattices induced by $f_1 \times f_2 \times f_3$. A basis vector indexed by a pair (e, g) is send to the product of three basis vectors indexed respectively by pairs $(e, f_1(g))$, $(e, f_2(g))$ and $(e, f_3(g))$. For sublattices spanned by basis vectors indexed by a fixed edge the morphism i can be described in coordinates as:

$$(a, b, c, d) \to (a + c, b + d, a + b, c + d, a + d, b + c).$$

In particular we see that i is indeed injective. Let $g : M_{S,\mathbb{Z}_2} \to M_{E,\mathbb{Z}_2}$ be the morphism of lattices that corresponds to the parametrization map of the binary model – cf. Definition 5.33. Let $g_0 : M_{S,\mathbb{Z}_2\times\mathbb{Z}_2} \to M_{E,\mathbb{Z}_2\times\mathbb{Z}_2}$ be the morphism of lattices that corresponds to the parametrization map of the 3-Kimura model.

We have got the following commutative diagram:

$$\begin{array}{ccc} M_{S,\mathbb{Z}_2} \times M_{S,\mathbb{Z}_2} \times M_{S,\mathbb{Z}_2} & \xrightarrow{g\times g\times g} & M_{E,\mathbb{Z}_2} \times M_{E,\mathbb{Z}_2} \times M_{E,\mathbb{Z}_2} \\ {\scriptstyle f_1^*\times f_2^*\times f_3^*}\uparrow & & \uparrow{\scriptstyle i} \\ M_{S,\mathbb{Z}_2\times\mathbb{Z}_2} & \xrightarrow{g_0} & M_{E,\mathbb{Z}_2\times\mathbb{Z}_2} \end{array}$$

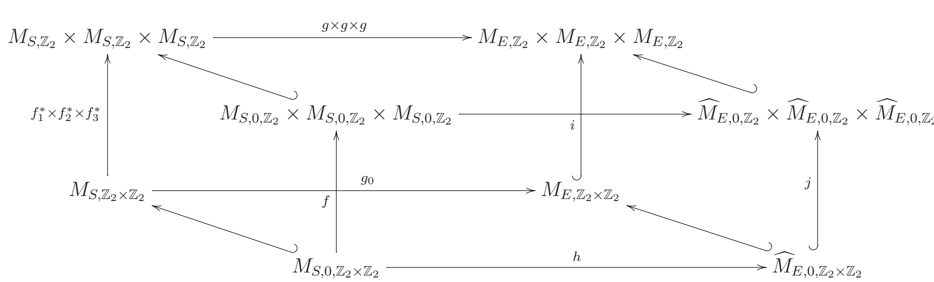
The following Fact follows from Corollary 6.4.

Fact 11.2. The dimension of the affine 3-Kimura model is equal to $3|E|+1$. The dimension of the product of three affine binary models is equal to $3(|E|+1)$. The dimension of the projective 3-Kimura model is equal to $3|E|$. The dimension of the product of three projective binary models is equal to $3|E|$. □

It follows that if we consider projective varieties representing the models, the dominant morphism from the product of three binary models to the 3-Kimura model described above becomes a rational, generically finite map. As the map between projective varieties is not a morphism we will restrict our attention only to dense orbits of the tori. On these tori orbits all maps are well defined and are represented by morphism of lattices.

11.1. **Maps of dense torus orbits.** Let us consider the following diagram:

(11.1)



The rectangle on the back is just the previous diagram. The rectangle in the front is induced from it by taking sublattices – cf. Definition 5.40. On the level of varieties the back is the affine picture, while the front is the projective one. The left square with lattices of type M_S corresponds to morphisms of ambient spaces. The square on the right describes the maps between varieties, or parameterizing spaces. The upper square corresponds to the product of three binary models, while the bottom square to the 3-Kimura model.

Let us explain the morphism j. It is injective, as it is a restriction of i. The lattice $\widehat{M}_{E,0}$ is the character lattice of the torus acting on the projective toric variety representing the model. The morphism j is induced by the rational finite map from the product of three $\mathbb{P}(X(T,\mathbb{Z}_2))$ to $\mathbb{P}(X(T,\mathbb{Z}_2\times\mathbb{Z}_2))$. Due to the coordinate system we can identify dense torus orbits with the tori.

Definition 11.3 (The torus $\mathbb{T}_X$)**.** *Let X be any toric variety in an affine or projective space with a distinguished coordinate system. Suppose that X is embedded equivariantly, as in Section 2. The dense torus orbit of X will be denoted by $\mathbb{T}_X \subset X$. Recall that $\mathbb{T}_X$ consists precisely of those points of X that have got all coordinates different from* 0.

The morphism j of character lattices is induced by the finite morphism from $\mathbb{T}_{(\mathbb{P}(X(T,\mathbb{Z}_2)))^3} = (\mathbb{T}_{\mathbb{P}(X(T,\mathbb{Z}_2))})^3$ to $\mathbb{T}_{\mathbb{P}(X(T,\mathbb{Z}_2\times\mathbb{Z}_2))}$. Due to the discussion in the proof of Proposition 8.1 we also know that the morphism of ambient spaces does not depend on the tree, but only on the number of leaves l. Hence the vertical morphisms of lattices on the left hand side of Diagram 11.1 are the same for all trees with l leaves.

11.2. **Idea of the proof.** The main reason for passing to tori is that we want to have a well defined dominant finite map. This allows us to take advantage of toric geometry. For example we know that the number of points in the fiber of the morphism of tori $(\mathbb{T}_{\mathbb{P}(X(T,\mathbb{Z}_2))})^3 \to \mathbb{T}_{\mathbb{P}(X(T,\mathbb{Z}_2\times\mathbb{Z}_2))}$ is equal to the index I_1 of the image of j in $(\widehat{M}_{E,0,\mathbb{Z}_2})^3$.

For the projective ambient spaces the situation is a little bit different. The morphism $f : M_{S,0,\mathbb{Z}_2\times\mathbb{Z}_2} \to (M_{S,0,\mathbb{Z}_2})^3$ is not injective, so the corresponding morphism of tori is not surjective. We will show that the image of f in $(M_{S,0,\mathbb{Z}_2})^3$ is of finite index, say I_2. It means that the corresponding morphism of tori is finite with each fiber having I_2 elements. Moreover we will show that $I_2 = I_1$. Hence we get the diagram:

$$\begin{array}{ccc} \mathbb{T}_{(\mathbb{P}(\widetilde{W}_{E,\mathbb{Z}_2})^3)} & \longrightarrow & \mathbb{T}_{\mathbb{P}(\widetilde{W}_{E,\mathbb{Z}_2\times\mathbb{Z}_2})} \\ \uparrow & & \uparrow \\ \mathbb{T}_{(\mathbb{P}(X(T,\mathbb{Z}_2)))^3} & \twoheadrightarrow & \mathbb{T}_{\mathbb{P}(X(T,\mathbb{Z}_2\times\mathbb{Z}_2))} \end{array}$$

where the horizontal maps are finite, étale of the same degree.

This means that if we consider the morphism of projective ambient spaces, then the preimage of $\mathbb{T}_{\mathbb{P}(X(T,\mathbb{Z}_2\times\mathbb{Z}_2))}$ is precisely $\mathbb{T}_{(\mathbb{P}(X(T,\mathbb{Z}_2)))^3}$.

Hence any intersection results that hold for the binary model must also hold for the 3-Kimura model. In particular as Conjecture 7.6 holds for the binary model we obtain a set-theoretic version of Conjecture 11.1 for the 3-Kimura model. By easy algebraic arguments we will also prove Conjecture 11.1 scheme-theoretically for 3-Kimura model.

11.3. **Proof.** Our first step will be to understand the morphism of projective ambient spaces $(\mathbb{P}(\widetilde{W}_{E,\mathbb{Z}_2}))^3 \dashrightarrow \mathbb{P}(\widetilde{W}_{E,\mathbb{Z}_2\times\mathbb{Z}_2})$. This is a well defined map on dense tori orbits. The map of tori corresponds to morphism of lattices $f: M_{S,0,\mathbb{Z}_2\times\mathbb{Z}_2} \to (M_{S,0,\mathbb{Z}_2})^3$. This morphisms depend only on the number of leaves, not on the tree.

By the definition we can embed the group of sockets $\mathfrak{S}$ in G^l. We can also view the group $\mathfrak{S}$ as a $\mathbb{Z}$-module. This gives us group morphisms $M_S \to \mathfrak{S} \to G^l$. The element of the basis of M_S indexed by a socket s is mapped to the socket s.

Example 11.4 (The case of the binary model and trivalent claw tree)**.** Let us consider the tree $K_{3,1}$ and the group $\mathbb{Z}_2$. We have got 4 sockets: $(0,0,0),(1,1,0),(1,0,1),(0,1,1)$. By coordinate-wise action they form a subgroup of $(\mathbb{Z}_2)^3$. The lattice M_S is freely generated by four basis vectors $e_{(0,0,0)}$, $e_{(1,1,0)}$, $e_{(1,0,1)}$, $e_{(0,1,1)}$. The morphism $M_S \to \mathfrak{S}$ maps $e_{(a,b,c)}$ to (a,b,c). Of course $ke_{(a,b,c)}$ is mapped to $k(a,b,c)$. For example $3e_{(1,1,0)}$ is mapped to $(1,1,0)+(1,1,0)+(1,1,0)=(1,1,0)$.

Lemma 11.5. *We have an exact sequence of groups:*

$$M_{S,0,\mathbb{Z}_2\times\mathbb{Z}_2} \to (M_{S,0,\mathbb{Z}_2})^3 \to (\mathbb{Z}_2)^l.$$

The first morphism is given by f. The second is the sum of three morphisms $M_{S,0,\mathbb{Z}_2} \to (\mathbb{Z}_2)^l$ described above[7].

Proof. It is clear that this is a complex. Let $(b'_i)_{i\geq 0}$ be the basis of $M_S^{\mathbb{Z}_2}$ corresponding to sockets. Let s_i be the socket corresponding to b'_i. Moreover suppose that b'_0 corresponds to the trivial socket, that is the neutral element of $\mathfrak{S}$. Let b_i be the basis of $M_{S,0,\mathbb{Z}_2}$ defined as $b_i = b'_i - b'_0$ for $i>0$. Note that an element (b'_i, b'_j, b'_k) is in the image of $f_1^* \times f_2^* \times f_3^*$ if and only if the corresponding three sockets s_i, s_j, s_k sum up to the neutral element of $\mathfrak{S}$. Hence the elements of the form $(b_i,b_i,0)=(b'_i,b'_i,b'_0)-(b'_0,b'_0,b'_0)$ are in the image of f. We see that $(2b_i,0,0)=(b_i,b_i,0)+(b_i,0,b_i)-(0,b_i,b_i)$ is also in the image. Furthermore for any two sockets s_i and s_j there exists a socket $s_k := s_i + s_j$ such that (b_i,b_j,b_k) is in the image of f. This reduces any element from $(M_{S,0,\mathbb{Z}_2})^3$ to an element $(b_i,0,0)$ modulo the image

[7] In this case the second operation is often called XOR.

of f or to 0. Hence any element is in the image if the XOR of all its coordinates is zero. □

Definition 11.6 (The kernel K)**.** *For any tree T let $K^T = K_1^T \times K_2^T \times K_3^T \subset M_{S,0,\mathbb{Z}_2} \times M_{S,0,\mathbb{Z}_2} \times M_{S,0,\mathbb{Z}_2}$ be the restriction of the kernel of the morphism $g \times g \times g$ to $M_{S,0,\mathbb{Z}_2} \times M_{S,0,\mathbb{Z}_2} \times M_{S,0,\mathbb{Z}_2}$.*

Each character in K^T is a character of $(\mathbb{T}_{\mathbb{P}(\widetilde{W}_E^{\mathbb{Z}_2})})^3$, that is the trivial character when restricted to the product $(\mathbb{T}_{\mathbb{P}(X(\mathbb{Z}_2))})^3$. Each such character is a triple of characters of $\mathbb{T}_{\mathbb{P}(\widetilde{W}_E^{\mathbb{Z}_2})}$. Each character of the triple is a quotient of monomials $\frac{m_1}{m_2}$ of the same degree on the projective space $\mathbb{P}(\widetilde{W}_E^{\mathbb{Z}_2})$. The polynomials $m_1 - m_2$ span[8] the ideal of the toric variety $\mathbb{P}(X(\mathbb{Z}_2))$. We want to view characters as functions. Hence we restrict our attention to $(\mathbb{T}_{\mathbb{P}(\widetilde{W}_E^{\mathbb{Z}_2})})^3$. In the algebra of this torus the ideal of $(\mathbb{T}_{\mathbb{P}(X(\mathbb{Z}_2))})^3$ is generated by elements $k - 1$, where $k \in K^T$.

Definition 11.7 (The kernel D)**.** *For any tree T let D^T be the kernel of the map h defined on Diagram 11.1.*

The elements of D represent characters trivial on the projective 3-Kimura variety. In the setting described at the end of Subsection 7.2 we want to prove that sublattices D^T for different trees T with l leaves generate the sublattice $D^{K_{l,1}}$. The idea is to push the lattices D to $(M_{S,0,\mathbb{Z}_2})^3$ using the morphism f. Next we use the results on binary models to obtain the generation for $f(D)$. Using properties of the image of f we are able to conclude the generation in $M_{S,0,\mathbb{Z}_2\times\mathbb{Z}_2}$. The following lemma enables us to restrict to the image of f instead of regarding whole lattice $(M_{S,0,\mathbb{Z}_2})^3$.

Lemma 11.8. *For any tree T the kernel K^T is a sublattice of the image of f.*

Proof. It is enough to show that $K_1^T \times \{0\} \times \{0\} \subset \operatorname{Im} f$. Suppose that $m = \sum_i a_i b_i \in K_1^T$, where each b_i is as in the proof of Lemma 11.5. Hence $b_i = (g_1^i - e, \dots, g_l^i - e)$, where e is the neutral element of $\mathbb{Z}_2$ and $g_j^i \in \mathbb{Z}_2$ are elements forming a socket. We know that $g(m) = 0$. In particular the coordinates of M_E indexed by leaves are equal to zero. Let us fix k that is a number of a leaf $1 \le k \le l$. Let us look at all coordinates indexed by pairs (k, q) where $q \in \mathbb{Z}_2$. The restriction of M_E to these coordinates is a free abelian group spanned by elements of $\mathbb{Z}_2$. Hence $\sum_i a_i(g_k - e) = 0$ in the free abelian group generated formally by elements of $\mathbb{Z}_2$. Hence, a fortiori, $\sum_i a_i(g_k - e) = e$ where

[8] They do not only generate the ideal, but even span it as the vector space.

now the sum is taken in $\mathbb{Z}_2$. As the action in $\mathfrak{S}$ is coordinate-wise we see that the image of m in $\mathfrak{S}$, and hence in $\mathbb{Z}_2^l$, is the neutral element. Using Lemma 11.5 we see that $m \in \operatorname{Im} f$. □

Proposition 11.9. *The index of the image of f in $(M_{S,0,\mathbb{Z}_2})^3$ is equal to the index of the image of j in $(\widehat{M}_{E,0,\mathbb{Z}_2})^3$.*

Proof. This is a consequence of Lemma 11.8. □

Corollary 11.10. *Conjecture 11.1 holds set-theoretically.*

Proof. The index of the image of f equals the degree of the finite map of tori. In particular we are in the situation of Diagram 11.2. The corollary follows from the discussion at the beginning of Section 11.2. □

Now we will prove Conjecture 11.1 scheme-theoretically. Let $T_0 = K_{l,1}$. We consider trees T_i such that the ideal of $\mathbb{T}_{\mathbb{P}(X(T,\mathbb{Z}_2))}$ is the sum of the ideals $\mathbb{T}_{\mathbb{P}(X(T_i,\mathbb{Z}_2))}$. Let K^{T_i} be the kernel of $g \times g \times g$ for the tree T_i. Let D^{T_i} be the kernel of h for the tree T_i. We know from Proposition 7.10 that the lattices K^{T_i} for $i > 0$ span K^{T_0}.

Theorem 11.11. *The lattices D^{T_i} for $i > 0$ span D^{T_0}. Conjecture 11.1 holds scheme theoretically.*

Proof. Let $a \in D^{T_0}$. We know that $f(a) \in K^{T_0}_{\mathbb{Z}_2}$, so $f(a) = \sum k_i$, where $k_i \in K^{T_i}_{\mathbb{Z}_2}$. Using Lemma 11.8 we can find $k_i' \in D^{T_i}$ such that $f(k_i') = k_i$. This means that $a - \sum k_i'$ is in the kernel of f. In particular, as j is injective, $a - \sum k_i'$ belongs to every D^{T_i}, hence we obtain the desired decomposition. □

Remark 11.12. From Proposition 7.10 it is enough to take two (particular) different $i > 0$ to span D^{T_0}, as it was in the case of binary model.

11.4. **Applications to phylogenetics.** In this section we present a few applications. The basic result that we use is due to Marta Casanellas and Jesús Fernández-Sánchez [CFS08]. It states that all points important for biologists are contained in the dense torus orbit of $X(T, \mathbb{Z}_2 \times \mathbb{Z}_2)$. Thus, following [CFS08], we call points of the dense torus orbit biologically meaningful. In Section 11 we gave a precise description of this orbit for any tree. This is sufficient for biologists.

People dealing with applications are usually interested in trivalent trees. Let us motivate the use of other trees. The first, obvious reason is that they can appear (at least hypothetically) as right models of evolution. This however is a degenerate situation that is often neglected. The next subsection presents a different reason.

11.4.1. *Identifiability.* Dealing with applications we are given a point P in the space of all possible probabilities $\widetilde{W_L}$. The first question is for which trees this point can be realized. More precisely for which trees T we have an inclusion $P \in X(T, \mathbb{Z}_2 \times \mathbb{Z}_2)$. We are interested in knowing if this is only one tree T or there are several possibilities. This is a first part of the *identifiability* problem. Hence Conjecture 7.6 is a question about the locus of points for which the identifiability problem cannot be resolved at all. Of course a generic point that belongs to any of the varieties belongs to exactly one $X(T, \mathbb{Z}_2 \times \mathbb{Z}_2)$ with T trivalent. Much more is known about the identifiability of different models. For the precise results the reader is advised to look in [AR06] or [APRS11] and the references therein.

In particular we see that points that belong to some $X(T, \mathbb{Z}_2 \times \mathbb{Z}_2)$ where T is not trivalent cannot identify the tree topology. Hence the question about the locus of these points, or equivalently about the polynomials defining such varieties may give some results for trivalent trees. However, as situation in Section 8 shows, the phylogenetic invariants of two varieties $X(T, \mathbb{Z}_2)$ for two different trees, do not generate the ideal of the variety associated to their degeneration.

The second, but equally important question about the identifiability is to give the description of the fiber of the parametrization map of the model $\check{\psi}^{-1}(P)$. The biologist aim at distinguishing one point in the fiber. This would enable to identify not only the tree topology, but also corresponding probabilities of mutation. The algebraic setting allows us to give a description of this fiber. We assume that P is biologically meaningful, that is is contained in the dense torus orbit. Equivalently all coordinates of P after the Fourier transform are different from zero. We prefer to work up to multiplicity, that is regard the projectivization of $\check{\psi}$ denoted by $\check{\psi}_{\mathbb{P}}$. The fiber $\check{\psi}_{\mathbb{P}}^{-1}(P)$ is contained in the dense torus orbit of $\prod \mathbb{P}(W_e)$. As this parameter space is of the same dimension as the image, we know that $\check{\psi}_{\mathbb{P}}$ is a generically finite map. Moreover when restricted to dense torus orbits it is étale and finite. Hence each fiber is finite and contains the same number of points, independent from P. This number is the index of lattice $\widehat{M_E}$ in a saturated sublattice of M_E. Of course we do not claim that all the points in the fiber have got a probabilistic meaning. We just prove that from the algebraic point of view there is always a fixed, finite number of possible candidates for transition matrices.

We will now give a precise description of a general fiber for a general group-based model corresponding to an abelian group H. Due to Corollary 6.4 we know that the map of projective tori parameterizing the model is a finite map. By dualizing the exact sequence in Corollary 6.7 we see that the kernel has got a group structure isomorphic to $H^{|N|}$. Due to [CFS08] the only biologically meaningful points are contained in the dense torus orbit.

Corollary 11.13. *Let T be any tree and H any abelian group. Let $\mathbb{P}(X)$ be the projective variety associated to the model. Let $x \in \mathbb{P}(X)$ be a biologically meaningful point. Up to multiplication by a constant there are exactly $|H|^{|N|}$ parameters in the fiber of x. In other words there are exactly $|H|^{|N|}$ possible transition matrices.* □

Note that we do not use further restrictions on the parameters of transition matrices. For example we do not assume that the parameters are real. This condition for sure further decreases the number of possible transition matrices. However we see that when we consider complex parameters the number of possible parameters is already finite and moreover independent from the considered point.

11.4.2. *Phylogenetic invariants.* The main theorem gives an inductive way of obtaining phylogenetic invariants of any tree. It is an open problem if these invariants generate the whole ideal. It is proved however that they give a description of all biologically meaningful points in case of the 3-Kimura model. The method is very simple. Suppose that we know the phylogenetic invariants for all trees with vertices of degree less or equal to d. Due to the results of [SS05] it is enough to describe the phylogenetic invariants for the claw tree $K_{d+1,1}$. For 3-Kimura, to obtain the description of the dense torus orbit we just take the sum of two ideals – cf. Remark 11.12. They are both associated to trees with the same topology. The tree has got two inner vertices v_1 and v_2 of degrees 3 and d respectively. The difference between the ideals is a consequence of different labelling of leaves. For one tree the leaves adjacent to v_1 are labeled by 1 and 2. For the second tree they are labeled 1 and 3. Notice that in fact we have to compute just one ideal. The second one can be obtained by permuting the variables.

12. Applications to the 3-Kimura model, part 2

The aim of this subsection is to further investigate Conjecture 7.1 for the 3-Kimura model. Let I_n be the ideal of the variety $X(T, \mathbb{Z}_2 \times \mathbb{Z}_2)$ where T is a claw tree with n leaves. Let I'_n be the subideal of I_n generated in degree 4. The conjecture of Sturmfels and Sullivant states that $I_n = I'_n$ for any n. In this subsection we will prove that I_n and I'_n define the same projective scheme. This is equivalent to the fact that their saturations are equal [Har77, Exercise 5.10 b)]. In particular it follows that they define the same affine set. One concludes that in order to check if any point belongs to the variety it is enough to consider phylogenetic invariants of degree four. Due to [SS05, Theorem 23] the result will follow for any tree. Let us state the main theorem of this subsection.

Theorem 12.1. *Consider any tree T and the 3-Kimura model. The ideal of the variety associated to it and the subideal generated by polynomials of degree at most four define the same projective scheme.*

We hope that the method presented in this section can be applied to other problems of the type *"prove that a toric projective scheme can be defined by an ideal generated in degree d"*. In general let I be an ideal of a projective toric variety. Let I' be the subideal generated in degree d. The aim is to prove that the saturation of I' with respect to the irrelevant ideal equals I.

Suppose that the variety is given by a polytope P, with points corresponding to coordinates of the ambient projective space – as in Section 2. Proving that the saturation of I' equals I is equivalent to proving that I' and I are equal in each localization with respect to any coordinate, represented by a point $Q \in P$. Thus we have to prove that any generator of I multiplied by a sufficiently high power of the variable corresponding to Q belongs to I'.

Let us translate this condition to combinatorial language. The generators of I correspond to relations between points of $P \times \{1\}$. Let us fix a relation $\sum A_i = \sum B_j$, where $A_i, B_j \in P \times \{1\}$. Multiplying the corresponding element of the ideal by the variable corresponding to Q is equivalent to adding Q to both sides of the relation. Thus we have to prove that the binomial corresponding to the relation $\sum A_i + mQ = \sum B_j + mQ$ is generated by binomials from I of degree at most d for m sufficiently large.

A binomial corresponding to a relation $\sum R_i = \sum S_i$ between points of a polytope is generated in a degree d if and only if one can transform $\sum R_i$ to $\sum S_i$ using a sequence of simple steps. In each single transformation one can replace points $R_1, \ldots, R_k$ for $k \leq d$ by $R'_1, \ldots, R'_k$ if they satisfy the relation $\sum_{i=1}^k R_i = \sum_{i=1}^k R'_i$. In such a case we say that the relation is generated in degree d.

The proof scheme is very simple:

(i) Using degree d relations reduce A_i, B_i to some simple, special points of $P \times \{1\}$ contained in a subset $L_Q \subset P$. (*)

(ii) Show that any relation between the points of L_Q is generated in degree d.

In general any of this two points can be very difficult.

Remark 12.2. It is well known that the projective toric variety defined by a polytope P is covered by affine subsets given by localizations by coordinates corresponding to vertices. Thus one can be tempted to prove that $I = I'$ only in the localizations by vertices. Note however that in general, we do not know if the scheme defined by I' is also covered by localizations by coordinates corresponding to vertices. Indeed, I' and I may be different on the set-theoretical level. For example if $\operatorname{Proj} I'$ contains a point that is zero on the coordinates corresponding to vertices and nonzero on some other coordinates, then such a point will not belong to any localization with respect to vertices. However if $\operatorname{rad} I' = I$, then of course it is enough to consider localizations with respect to vertices.

As our polytopes have only vertices, the problem described in Remark 12.2 does not concern us.

Remark 12.3. We have got the following equivalences for a toric ideal I given by a polytope $P \times \{1\}$.

- All relations between vertices of $P\times\{1\}$ are generated in degree $d \Leftrightarrow$ the ideal I is generated in degree d.
- For any point $Q \in P \times \{1\}$ and any relation there is an integer m such that after adding mQ to both sides of the relation, it is generated in degree $d \Leftrightarrow$ the projective scheme defined by I can be defined by an ideal generated in degree d.
- For any relation there are[9] points $Q_i \in P \times \{1\}$ such that after adding $\sum Q_i$ to both sides, it is generated in degree $d \Leftrightarrow$ the dense torus orbit of the variety is defined by the ideal generated in degree d in the algebra of the ambient torus.

The whole subsection is devoted to the proof of Theorem 12.1. the proof is involved but completely elementary. The first observation is that by Proposition 6.8 the group of networks acts on the variety, hence on the ideals I_n and I'_n. The action is transitive on the points of the polytope, as they correspond to elements of the group. Using this action we can reduce to the case where the point Q of the polytope represents the coordinate corresponding to the trivial network, that is a network assigning neutral elements to all edges. Due to Fact 5.71 we can consider only claw trees. Let us index the edges of a claw tree $K_{n,1}$ with numbers $1, \dots, n$. We will identify a network with an n-tuple of group elements summing up to zero. The sum of such n-tuples will be

[9] not necessarily different

a coordinatewise sum, where each entry is treated as an element of the free abelian group generated by elements of $\mathbb{Z}_2 \times \mathbb{Z}_2$. Each network represents a vertex of a polytope $P \subset M_E$. The addition described above is the addition in this lattice.

Example 12.4. For $n = 4$ we can add in the lattice M_E:

$$(0, (1, 0) + (1, 1), 2 \cdot (0, 1), -3 \cdot (0, 0)) + ((0, 1), (1, 0) - (1, 1), (1, 0), (0, 0)) \\ = ((0, 1), 2 \cdot (1, 0), 2 \cdot (0, 1) + (1, 0), -2 \cdot (0, 0)).$$

The trivial network is $((0, 0), (0, 0), (0, 0), (0, 0))$.

Definition 12.5 (Support of a network)**.** *Let n be any network. The set of indices of edges to which n associates a nonneutral element is called the* support *of n.*

Definition 12.6 (Pair, triple)**.** *We say that a network is a* pair *if and only if the cardinality of the support is equal to two. We say that a network is a* triple *if and only if the cardinality of the support is equal to three.*

By nt we denote the neutral element in the group of networks.

Lemma 12.7. *For any network s, for m sufficiently large, $s + m \cdot nt$ can be reduced using degree two relations to a sum of networks that either:*
1) assign the same nontrivial element to two edges – pairs
2) assign three different nontrivial elements to three edges – triples
and the neutral element to all other edges.

Proof. The proof is inductive on the size of the support. Suppose the support of s is of cardinality at least four. We can choose a strict subset S of the support such that the sum of group elements $\sum_{e \in S} s(e)$ is the neutral element. Consider the networks s' and s'' that agree with s respectively on the set S and its complement and assign to all the other edges the neutral element. We have $s + nt = s' + s''$, which finishes the proof. □

Example 12.8. Consider the tree $K_{4,1}$.

$$((1, 0), (0, 1), (0, 1), (1, 0)) + ((0, 0), (0, 0), (0, 0), (0, 0)) \\ = ((1, 0), (0, 0), (0, 0), (1, 0)) + ((0, 0), (0, 1), (0, 1), (0, 0)).$$

We see that we can assume that f represents a relation only between pairs and triples. This completes the first step of the method (*) presented at the beginning of the section. The set L_Q consists of pairs and triples.

Let us fix any relation $\sum n_i = \sum n'_i$, where n_i and n'_i are networks that are either pairs or triples. Our aim is to transform $\sum n_i$ to $\sum n'_i$ in a series of steps, each time replacing at most four n_i by networks with the same sum[10]. We assume that among n_i there are more or the same number of triples as among n'_i. We first try to reduce the relation, so that consequently:

(i) Among n_i there are as few triples as possible,
(ii) Among n'_i there are as few triples as possible,
(iii) The degree of the relation is as small as possible.

More precisely let t and t' be the number of triples among respectively n_i and n'_i. Let d be the degree of the relation. Our proof will be inductive on (t, t', d) with lexicographic order.

To prove Theorem 12.1 we consider separately three cases depending on the number of triples among n_i. The cases are:

a) there are no triples,
b) there is exactly one triple,
c) there are at least two triples.

We say that a family of networks agrees on an index j if they all associate the same element to j and j belongs to their *support*. We will denote by g_1, g_2 and g_3 the three nontrivial elements of $\mathbb{Z}_2 \times \mathbb{Z}_2$. A triple that associates g_1 to index a, g_2 to index b and g_3 to index c is denoted by (a, b, c). A pair that associates an element g_i to indices d and e will be denoted by $(d, e)_{g_i}$. We say that g_i is contained in a network if there exists an index j, such that the network associates g_i to j. We believe that the following proofs are impossible to follow without a piece of paper. We strongly encourage the reader to note what networks appear in both sides of the relation at each step of the proof.

12.1. **The case with no triples.** First note that there are no triples among n'_i. Without loss of generality we may assume that n_1 is a pair equal to $(a, b)_{g_1}$. Hence there exists n'_i, say n'_1, that is $(b, c)_{g_1}$ for some index c. If $c = a$ we can reduce this pair, hence we assume $c \neq a$. There exists a network, say n_2 that is $(c, d)_{g_1}$. If $d = b$ we can reduce this pair. We consider two other cases:

1) $d \neq a$. Then we use the degree two relation $(a, b)_{g_1} + (c, d)_{g_1} = (a, d)_{g_1} + (b, c)_{g_1}$ and we can reduce $(b, c)_{g_1}$.

2) $d = a$. Then there is a network, say n'_2 given by $(a, e)_{g_1}$. If $e = b$ or $e = c$ we can reduce this pair. In the other cases we use the relation $(a, e)_{g_1} + (b, c)_{g_1} = (a, b)_{g_1} + (e, c)_{g_1}$ and we reduce $(a, b)_{g_1}$.

[10]We are also allowed to add the trivial network to both sides.

Notice that in this case we have only used degree two relations.

12.2. **The case with one triple.** Let n_1 be the only triple among n_i.

Lemma 12.9. *There is exactly one triple among n_i'.*

Proof. Due to the assumption that there are less triples among n_i' than among n_i we know that if there is a triple among n_i' then it is unique. Suppose that there are no triples among n_i'. Then the element g_1 appears among n_i' an even number of times. Indeed each pair contains g_1 twice or does not contain it. As the sum of n_i' equals the sum of n_i also the number of times g_1 appears among n_i must be even. This is impossible as n_1 contains g_1 just once and all pairs contain g_1 twice or do not contain it at all. $\square$

Due to the previous lemma we may assume that n_1' is the only triple among n_i'. Let $n_1 = (1, 2, 3)$.

12.2.1. *Case: The triples do not agree on any element of the support.* We want to reduce to the case where n_1' agrees with n_1 on an index that belongs to the support of both. *Suppose that this is not the case.*

The reduction in this case will have two steps. First, if n_1 and n_1' have the same support we will use the relations to move the supports, reduce the triples or decrease the degree. Next we will show how to deal with the case when the supports are not the same.

1) First step – suppose that $\{1, 2, 3\}$ is also the support of n_1'.

Remember that due to the assumption 12.2.1 the triples n_1 and n_1' do not agree on any element from their support. As n_1' has support $\{1, 2, 3\}$ without loss of generality we may assume that $n_1' = (2, 3, 1)$. Hence there must be a pair $(2, a)_{g_1}$ among n_i and $(1, b)_{g_1}$ among n_i'. If $a = 1$ and $b = 2$ then both pairs are the same and can be reduced. As both cases are symmetric we can assume that $a \neq 1$.

If $a \neq 3$ we can use the relation $(2, a)_{g_1} + (1, 2, 3) = (a, 2, 3) + (2, 1)_{g_1}$. This reduces to the case with different supports. We are left with the case $a = 3$. There must be a pair $(3, z)_{g_1}$ among n_i'. If $z \neq 1$ we can use the relation $(3, z)_{g_1} + n_1' = (z, 3, 1) + (2, 3)_{g_1}$. This would enable to reduce the $(2, 3)_{g_1}$ pair and decrease the degree. So we can assume that $z = 1$. So far we have shown that there must be pairs $(2, 3)_{g_1}$ among n_i and $(3, 1)_{g_1}$ among n_i'[11]. By the same reasoning for g_2 and g_3 we see that we can use the following relation:

$$(1, 2, 3)+(2, 3)_{g_1}+(1, 3)_{g_2}+(1, 2)_{g_3} = (2, 3, 1)+(2, 3)_{g_3}+(1, 3)_{g_1}+(1, 2)_{g_2}.$$

[11]Notice that we have made a symmetry assumption $a \neq 1$. The symmetric assumption would be $b \neq 2$. However as the result we got was symmetric, also for $b \neq 2$ we prove the existence of the same pairs.

Notice that this is a degree four relation. It enables us to reduce triples.

2) Second step – the triples n_1 and n'_1 have different supports.

Once again let $(1,2,3) = n_1$ and let $(a,b,c) = n'_1$. We may assume that a is not in the support of n_1. We see that there must be a pair $(a,f)_{g_1}$ among n_i. If $f \neq 1$ we can use a relation $(a,f)_{g_1} + n_1 = (a,2,3) + (f,1)_{g_1}$. This reduces to the case when the triples agree on a, hence we assume that $f = 1$. Hence there must be a pair $(g,1)_{g_1}$ among n'_i. If $g = a$ we can reduce this pair, so we assume $g \neq a$. Notice that there must be a pair $(g,h)_{g_1}$ among n_i. If $h \neq a$, then we can use relation $(1,a)_{g_1} + (g,h)_{g_1} = (g,1)_{g_1} + (h,a)_{g_1}$ and reduce the pair $(g,1)_{g_1}$. So we can assume $h = a$. Then there must be a pair $(a,i)_{g_1}$ among n'_i. If $i = 1$ then we can reduce it. Otherwise we can use the relation $(g,1)_{g_1} + (a,i)_{g_1} = (g,a)_{g_1} + (1,i)_{g_1}$ and reduce the pair $(g,a)_{g_1}$.

12.2.2. *Case: the triples agree on exactly one element in their support.* So far we reduced to the case where triples agree on at least one element, say 1, in their common support. Now we want to make a further reduction, so that the triples agree on two elements that are in their supports. *Assume this is not the case.*

As before let $n_1 = (1,2,3)$ and $n'_1 = (1,b,c)$. We consider three cases.

1) $b \neq 3$.

There must be a pair $(b,d)_{g_2}$ among n_i. If $d \neq 2$ then we can apply the relation $(b,d)_{g_2} + n_1 = (1,b,3) + (d,2)_{g_2}$. This reduces to the case where triples agree on two elements. So we assume $d = 2$. There must be a pair $(2,e)_{g_2}$ among n'_i. Hence there must also be a pair $(e,f)_{g_2}$ among n_i. If $f \neq b$ we can use a relation $(e,f)_{g_2} + (2,b)_{g_2} = (e,2)_{g_2} + (f,b)_{g_2}$ and reduce $(e,2)_{g_2}$. For $f = b$ we must have a pair $(b,g)_{g_2}$ among n'_i. If $g = 2$ or $g = e$ then this pair can be reduced. In the other case we use the relation $(e,2)_{g_2} + (b,g)_{g_2} = (e,g)_{g_2} + (b,2)_{g_2}$ and reduce $(b,2)_{g_2}$.

2) $c \neq 2$.

This case is analogous to 1).

3) $b = 3$ and $c = 2$.

Lemma 12.10. *If there is a pair $(p,q)_{g_2}$ among n_i, such that $p,q \neq 2$ then we may assume that it is equal to $(1,3)$.*

Proof. Suppose that $p \neq 1,3$ and $q \neq 2$. We apply a relation $(p,q)_{g_2} + n_1 = (1,p,3) + (q,2)_{g_2}$ and reduce to case 1) $b \neq 2$. □

Analogously if there is a pair $(p,q)_{g_2}$ among n'_i, such that $p,q \neq 3$ then this pair equals $(1,2)_{g_2}$.

Notice that there must be a pair $(3,d)_{g_2}$ among n_i and a pair $(2,e)_{g_2}$ among n_i'. From Lemma 12.10 d equals 2 or 1 and e equals 3 or 1. We will consider subcases.

3.1) Suppose that $d = 2$.

If $e = 3$ then we can make a reduction of pairs. If $e = 1$ we must have a pair $(1,f)_{g_2}$ among n_i. If $f = 2$ we make a reduction, hence we assume $f = 3$. This means that there must be a pair $(3,g)_{g_2}$ among n_i'. If $g = 2$ or $g = 1$ we can make a reduction. Otherwise we apply the relation $(1,2)_{g_2} + (3,g)_{g_2} = (1,3)_{g_2} + (2,g)_{g_2}$ and reduce the pair $(1,3)_{g_2}$.

3.2) Suppose that $e = 3$.

This case is similar to 3.1).

3.3) Suppose that $d = 1$ and $e = 1$.

As this is the only case left we may repeat the same reasoning for g_3. In particular, we must have a pair $(1,2)_{g_3}$ among n_i. We see that we can reduce the triples by applying the following relation:

$$(1,2,3) + (1,3)_{g_2} + (1,2)_{g_3} = (1,3,2) + (1,2)_{g_2} + (1,3)_{g_3}.$$

This is a degree three relation.

12.2.3. *Case: the triples agree on at least two elements in their support.* So far we reduced to the case where triples agree on two elements, say 1 and 2, that are in their support. Suppose that $n_1 = (1,2,3)$ and $n_1' = (1,2,c)$. Of course if $c = 3$ we can make a reduction. In other case we must have a pair $(c,d)_{g_3}$ among n_i. If $d \neq 3$ then we use the relation $(c,d)_{g_3} + (1,2,3) = (1,2,c) + (3,d)_{g_3}$ and reduce the triples. Hence $d = 3$. Analogously there must be a pair $(3,c)_{g_3}$ among n_i', hence we can reduce this pair.

12.3. **The case with at least two triples.** We suppose that there are at least two triples among n_i.

Lemma 12.11. *If there are two triples n_1, n_2 among n_i that do not agree on any element of their supports then we can make a reduction. Thus we can assume that any two triples among n_i agree on at least one index.*

Proof. The assumptions are equivalent to $n_1 = (a,b,c)$, $n_2 = (d,e,f)$ with $a \neq d$, $b \neq e$, $c \neq f$. We apply the relation $n_1 + n_2 + nt = (a,d)_{g_1} + (b,e)_{g_2} + (c,f)_{g_3}$ that reduces the number of triples. □

Lemma 12.12. *If there is no index on which all triples from n_i agree then we can make a reduction.*

Proof. Suppose there is no index on which all n_i agree. We may consider only two cases due to Lemma 12.11.

1) Suppose that any two triples from n_i agree on at least two elements.

Consider any triple $n_1 = (1, 2, 3)$. Due to the fact that not all triples from n_i associate g_1 to 1 there is a triple $(a, 2, 3)$ with $a \neq 1$ among n_i. There also must be a triple that does not associate g_2 to 2. It cannot agree both with $(1, 2, 3)$ and $(a, 2, 3)$ on two indices.

2) There exist two triples that agree only on one index.

Let $n_1 = (1, 2, 3)$ and $n_2 = (1, b, c)$ with $b \neq 2$ and $c \neq 3$. Due to the case assumption there is a triple $n_3 = (d, e, f)$ with $d \neq 1$. Remember that any two triples have to agree on at least one element due to Lemma 12.11. Hence without loss of generality we can assume $e = b$ and $f = 3$. We can apply the relation:

$$n_1 + n_2 + n_3 + nt = (d, 1)_{g_1} + (2, b)_{g_2} + (3, c)_{g_3} + (1, b, 3),$$

that reduces the number of triples. □

Due to the previous lemma we may assume that there exists an index, say 1*, such that all triples among* n_i *associate to it the same nonneutral element, say* g_1*.*

Definition 12.13 (k)**.** *Let* k *be the number of indices on which all triples among* n_i *agree. We know that* $1 \leq k \leq 3$*.*

We proceed inductively on k, as for $k = 0$ we already know how to reduce the relation. Hence from now on decreasing k is also a reduction.

Lemma 12.14. *Suppose that all triples from* n_i *associate* g_j *to an index* l*. If there is a pair* $(x, y)_{g_j}$ *among* n_i *with* $l \neq x, y$ *then either* $\{l, x, y\}$ *is the support of all triples among* n_i *or we can make a reduction.*

Proof. To simplify the language assume $g_j = g_1$ and $l = 1$. Suppose that there is a triple $n_1 = (1, b, c)$ with the support different from $\{1, x, y\}$. We can assume $x \neq b, c$. We apply the relation $n_1 + (x, y)_{g_1} = (x, b, c) + (1, y)_{g_1}$ what reduces k. □

Lemma 12.15. *Suppose that all triples from* n_i *associate* g_j *to an index* l*. If all pairs* $(x, y)_{g_j}$ *among* n_i *have* l *in the support then we can reduce all such pairs.*

Proof. Let t be the number of triples among n_i. Let p be the number of g_j pairs among n_i. Let t'_1 and t'_2 be the number of triples in n'_i that respectively assign or do not assign g_j to l. Let p'_1 and p'_2 be the number of g_j pairs among n'_i that respectively have or do not have l

in the support. We know that $t \geq t_1' + t_2'$. Comparing the number of times g_j appears in n_i and n_i' we get:

$$t + 2p = t_1' + t_2' + 2(p_1' + p_2').$$

Comparing the number of times g_j appears on index l we get:

$$t + p = t_1' + p_1'.$$

This forces $t_2' = p_2' = 0$, $t = t_1'$ and $p = p_1'$. Hence all g_j pairs and triples among n_i and n_i' must assign g_j to l. Hence the multisets of pairs must be the same for n_i and n_i'. □

Lemma 12.16. *If there are g_1 pairs among n_i, then we can make a reduction.*

Proof. We will prove that there are no pairs $(a,b)_{g_1}$ among n_i that do not have 1 in the support. Due to Lemma 12.15 this will finish the proof. Suppose that there is a pair $(a,b)_{g_1}$ among n_i with $a, b \neq 1$. Due to Lemma 12.14 all the triples among n_i must have the support $\{1, a, b\}$. So either $k = 1$ or $k = 3$. If $k = 1$ we can apply the relation

$$(1, a, b) + (1, b, a) + (a, b)_{g_1} + nt = (1, a)_{g_1} + (1, b)_{g_1} + (a, b)_{g_2} + (a, b)_{g_3}.$$

This reduces the number of triples. Thus we can assume that all triples among n_i are equal to $(1, a, b)$.

Claim: *Consider any pair $(c, d)_{g_2}$ among n_i. We can assume that its support is contained in $\{1, a, b\}$.*

Proof of the Claim. Suppose this is not the case, that is $c \notin \{1, a, b\}$. Due to Lemma 12.14 we can assume $d = a$.

1) Suppose that there is a g_2 pair among n_i that does not contain a in the support.

It must be equal to $(1, b)_{g_2}$ due to Lemma 12.14. We can apply the relation $(1, b)_{g_2} + (a, c)_{g_2} = (c, 1)_{g_2} + (a, b)_{g_2}$. Applying once again Lemma 12.14 to the pair $(c, 1)_{g_2}$ we can make a reduction.

2) All g_2 pairs among n_i contain a in the support.

Due to Lemma 12.15 we can make a reduction. □

Thus the support of all g_2 pairs among n_i is contained in $\{1, a, b\}$. The same holds for g_1 and g_3 pairs. Thus all networks among n_i have support contained in $\{1, a, b\}$. Hence the same must hold for n_i'. So our relation is a relation only on three indices. It is well known that the ideal for a tree with three edges is generated in degree 4, so in particular the considered relation is generated in degree 4. □

Corollary 12.17. *If all triples among n_i associate g_j to an index l, then there are no g_j pairs among n_i. Consequently there are no g_j*

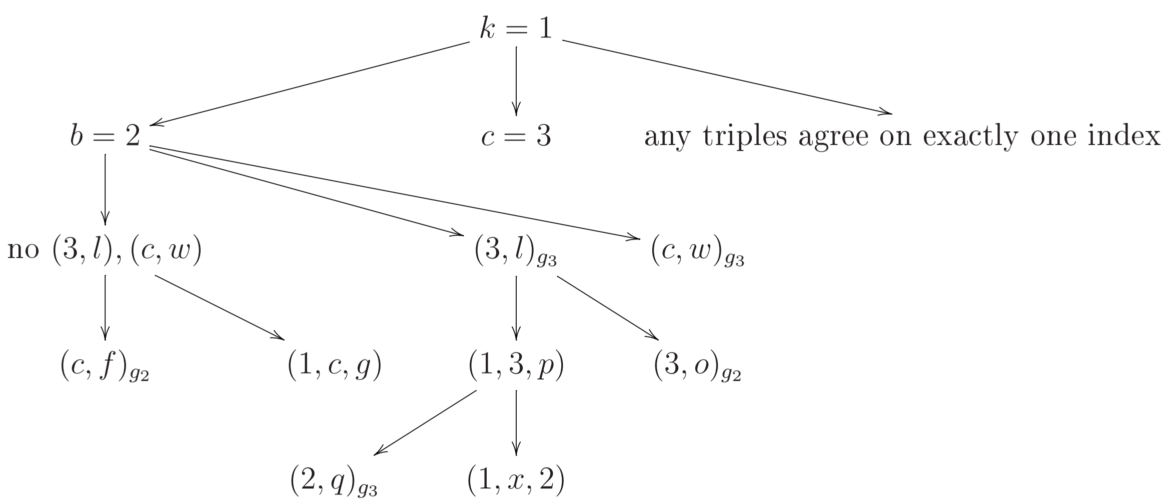
pairs among n_i' and all triples among n_i' associate g_j to l. Moreover the number of triples among n_i equals the number of triples among n_i'. □

In conclusion we reduced to the case were there are no g_1 pairs neither among n_i nor n_i'. Moreover, there is the same number of triples among n_i and n_i' and they all associate g_1 to 1.

Lemma 12.18. *If all the triples among n_i and n_i' have support in $\{1,2,3\}$ then either $k=3$ or we can reduce a triple.*

Proof. In this case $k=1$ or $k=3$. If $k=1$ then among n_i there is a triple $(1,2,3)$ and $(1,3,2)$. One of this triples can be reduced. □

12.3.1. *Case:* $k=1$. We first consider the most difficult case $k=1$. As always let $n_1=(1,2,3)$ and $n_1'=(1,b,c)$. As the proof is quite complicated we decided to include the tree that describes most important cases. While reading the proof we encourage the reader to follow at which node we are. The proof is "depth-first, left-first".



We start with the left node in the second row – assume $b=2$.

We move to the most left node in the third row – suppose that there is no g_3 pair among n_i that has got c in the support and, symmetrically, there is no g_3 pair among n_i' that has got 3 in the support. There must be a triple $(1,e,c)$ among n_i. If $e\neq 3$ then we apply the relation $(1,2,3)+(1,e,c)=(1,2,c)+(1,e,3)$ and reduce the triple $(1,2,c)$. We have $e=3$. Analogously among n_i' there must be a triple $(1,c,3)$. Hence there must be either a pair $(c,f)_{g_2}$ or a triple $(1,c,g)$ among n_i.

We continue to the most left node in the fourth row – suppose that there is a pair $(c,f)_{g_2}$. If $f\neq 2$ we apply the relation $(1,2,3)+(c,f)_{g_2}=(1,c,3)+(f,2)_{g_2}$ and reduce the triple $(1,c,3)$. If $f=2$ we apply the relation $(1,3,c)+(c,2)_{g_2} = (1,2,c)+(3,c)_{g_2}$ and reduce the triple $(1,2,c)$.

Hence we can assume that there is a triple $(1,c,g)$ among n_i – second node in the fourth row. If $g \neq 2$ then we apply the relation $(1,c,g)+(1,2,3)=(1,2,g)+(1,c,3)$ and reduce the triple $(1,c,3)$. For $g=2$ we apply the relation $(1,2,3)+(1,3,c)+(1,c,2)=(1,2,c)+(1,3,2)+(1,c,3)$ and reduce the triple $(1,2,c)$.

We continue to the second node in the third row. We assume that there is a pair $(3,l)_{g_3}$ among n_i'. If $l \neq c$ we apply the relation $(1,2,c)+(3,l)_{g_3}=(1,2,3)+(c,l)_{g_3}$ and reduce the triple $(1,2,3)$. If there was a pair $(c,m)_{g_3}$ among n_i then analogously we would have $m=3$ and we would be able to reduce this pair. So there must be a triple $(1,n,c)$ among n_i. If $n \neq 3$ then we apply the relation $(1,2,3)+(1,n,c)=(1,n,3)+(1,2,c)$ and reduce the triple $(1,2,c)$. So we assume $n_2=(1,3,c)$. Hence there is either a pair $(3,o)_{g_2}$ or a triple $(1,3,p)$ among n_i'.

We move to the third node in the fourth row – suppose that there is a triple $(1,3,p)$ among n_i'. If $p \neq 2$ we apply the relation $(1,2,c)+(1,3,p)=(1,2,p)+(1,3,c)$ and we reduce $(1,3,c)$. So we have $p=2$. There is either a pair $(2,q)_{g_3}$ or a triple $(1,x,2)$ among n_i.

Consider the first node in the fifth row – suppose that there is a pair $(2,q)_{g_3}$ among n_i. If $q \neq c$ then we apply the relation $(1,3,c)+(2,q)_{g_3}=(1,3,2)+(c,q)_{g_3}$ and reduce $(1,3,2)$. If $q=c$ we apply the relation $(1,2,3)+(2,c)_{g_3}=(1,2,c)+(2,3)_{g_3}$ and reduce the triple $(1,2,c)$.

So we can move to the second node in the fifth row – assume that there is a triple $(1,x,2)$ among n_i. If $x \neq c$ we apply the relation $(1,3,c)+(1,x,2)=(1,x,c)+(1,3,2)$ and reduce the triple $(1,3,2)$. If $x=c$ we apply the relation $(1,2,3)+(1,3,c)+(1,c,2)=(1,2,c)+(1,3,2)+(1,c,3)$ and reduce the triple $(1,2,c)$.

We pass to the fourth node in the fourth row – we assume that there is a pair $(3,o)_{g_2}$ and there is no triple $(1,3,p)$ among n_i'. If $o \neq 2$ then we apply the relation $(1,2,c)+(3,o)_{g_2}=(1,3,c)+(2,o)_{g_2}$ and reduce $(1,3,c)$. So there is a pair $(2,3)_{g_2}$ among n_i'. Suppose that this pair appears $r>0$ times among n_i'. Note that there are no pairs $(2,s)_{g_2}$ among n_i. Indeed suppose that there is such a pair. If $s \neq 3$ then we apply the relation $(1,3,c)+(2,s)_{g_2}=(1,2,c)+(3,s)_{g_2}$ and reduce the triple $(1,2,c)$. If $s=3$ we reduce the pair $(2,3)_{g_2}$. Hence there must be at least $r+1$ triples of the type $(1,2,t)$ among n_i. If there is a triple with $t \neq 3$ then we apply the relation $(1,3,c)+(1,2,t)=(1,3,t)+(1,2,c)$ and reduce the triple $(1,2,c)$. Hence we have got at least $r+1$ triples $(1,2,3)$ among n_i. Notice that there are no triples of the type $(1,y,3)$ among n_i'. Indeed, in such a case we could apply the relation $(1, y, 3) + (2, 3)_{g_2} = (1, 2, 3) + (y, 3)_{g_2}$ and reduce $(1, 2, 3)$. Hence there must be at least $r+1$ pairs of the type $(3, u)_{g_3}$ among n_i'. If $u \neq c$ then we apply the relation $(1, 2, c) + (3, u)_{g_3} = (1, 2, 3) + (c, u)_{g_3}$ and reduce the triple $(1, 2, 3)$. Hence we have at least $r+1$ pairs $(3, c)_{g_3}$ among n_i'. Note that there are no pairs of the type $(c, v)_{g_3}$ among n_i. Indeed if $v = 3$ we could reduce this pair. If $v \neq 3$ then we apply the relation $(1, 2, 3) + (c, v)_{g_3} = (1, 2, c) + (3, v)_{g_3}$ and reduce the triple $(1, 2, c)$. Hence we must have at least $r + 1$ triples of the type $(1, z, c)$ among n_i. If $z \neq 3$ then we apply the relation $(1, 2, 3) + (1, z, c) = (1, 2, c) + (1, z, 3)$ and reduce the triple $(1, 2, c)$. So there are at least $r + 1$ triples $(1, 3, c)$ among n_i. Note that the elements g_2 on 3 cannot be reduced – among n_i' there are only r pairs containing them and no triples. The contradiction finishes this case.

Consider the third node in the third row – there is a pair $(c, w)_{g_3}$ among n_i. This is completely analogous to the second node in this row, already considered.

Also the second node in the second row – $c = 3$ – is analogous to the first node in the second row.

We are left with the last, third node in the second column – any two triples n_i and n_j' agree on exactly one index, that is on 1. Due to Lemma 12.18 we can assume $b \neq 2$ and $b \neq 3$. Due to the case assumption there must be a pair $(b, d)_{g_2}$ among n_i. If $d \neq 2$ then we apply the relation $(1, 2, 3) + (b, d)_{g_2} = (1, b, 3) + (d, 2)_{g_2}$ and reduce to the case $b = 2$[12]. Analogously we must have the same pair among n_i' and it can be reduced.

12.3.2. *Case: $k = 2$ or $k = 3$.* Suppose now that $k = 2$. Let $n_1 = (1, 2, 3)$ and $n_1' = (1, 2, c)$. If we cannot reduce n_1' then there must be a pair $(c, d)_{g_3}$ among n_i and a pair $(3, e)_{g_3}$ among n_i'. If $d = 3$ and $e = c$ we can reduce the pairs. Thus we can assume that $d \neq 3$. We apply the relation $(1, 2, 3) + (c, d)_{g_3} = (1, 2, c) + (3, d)_{g_3}$ and reduce the triple $(1, 2, 3)$.

The last, easiest case is $k = 3$. Then all triples are equal to $(1, 2, 3)$ and there are no pairs due to Corollary 12.17. Hence we can reduce the triples. This finishes the proof of the theorem.

13. Open problems

We have already presented a few conjectures in this part of the thesis. Here we would like to give a list of problems that should be much easier, however still we find them interesting.

[12]Notice that we do not reduce to the case $k = 2$ as if this was true we would have already been in the first node in the second column $b = 2$.

We start with the questions concerning normality. We already know that many general group-based models give rise to projectively normal varieties for trivalent trees. However not much is known about trees of higher valency. Of course, due to Proposition 5.72, it is enough to consider claw trees. The questions on normality are important as many toric methods work only for normal polytopes. We have already applied some of them to compute Hilbert functions. Further applications to the conjecture of Sturmfels and Sullivant could be possible due to the methods of "finite generation in rings with infinitely many variables" – for more details see [HS11], [DK11]. The question for the binary model should not be difficult.

Conjecture 13.1. *Let T be any tree. The polytope representing the binary Jukes-Cantor model on T is normal.*

The same question for the 3-Kimura model, in our opinion, is much more ambitious.

Conjecture 13.2. *Let T be any tree. The polytope representing the binary 3-Kimura model on T is normal.*

Recall that in Proposition 9.1 we showed that the *projective* variety representing the model is not normal. We also know that the affine variety representing the general group-based model for $\mathbb{Z}_6$ is not normal.

Conjecture 13.3. *The projective toric variety representing the general group-based model for $\mathbb{Z}_6$ on $K_{1,3}$ is not normal.*

Another question is to what extent the methods of Section 12 can be applied to other abelian groups.

Conjecture 13.4. *The projective scheme associated to the group-based model for $\mathbb{Z}_3$ and any tree can be represented by an ideal generated in degree* 3.

We finish by restating, in our opinion, the most interesting, important and difficult Conjecture 7.6.

Conjecture 13.5. *The variety $X(K_{n,1})$ is equal to the (scheme theoretic) intersection of all the varieties $X(T_i)$, where T_i is a prolongation of $K_{n,1}$ that has only two inner vertices, both of them of valency at least three.*

Appendix 1

Here we show an explicit example when the equality of the parameters before the Fourier transform does not imply the equality after it.

Let $G = \mathbb{Z}_6$. The transition matrices are of the form:

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

The matrix of the type above corresponds to a function $g : G \to \mathbb{C}$, such that $g(0) = a$, $g(1) = b$, $g(2) = c$, $g(3) = d$, $g(4) = e$ and $g(5) = f$. The Fourier transform of g gives us: $\widehat{g}(\chi_0) = a + b + c + d + e + f$, $\widehat{f}(\chi_1) = a+jb+j^2c+j^3d+j^4e+j^5e$, $\widehat{f}(\chi_2) = a+j^2b+j^4c+d+j^2e+j^4e$ etc. where j is a primitive sixth root of unity. We consider a submodel defined by $g(0) = g(1) = g(5)$ and $g(2) = g(4)$. This corresponds to $a = b = f$ and $c = e$. The Fourier transform gives us respectively $(x_0, x_1, x_2, x_3, x_4, x_5) = (3a+2c+d, 2a-c-d, -c+d, -a+2c-d, -c+d, 2a-c-d)$. This defines a linear subspace given by $x_4 = x_2$, $x_5 = x_1$ and $x_1 + 3x_2 + 2x_3 = 0$. This is not an equality of distinct variables.

Appendix 2

Here we present the precise results of the computations of Hilbert-Ehrhart polynomials for a few G-models. The results are from a joint work with Maria Donten-Bury [DBM].

For the groups $\mathbb{Z}_8$, $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ and $\mathbb{Z}_9$ we computed only the Hilbert function and, as we could not check the normality, we do not know if it is equal to Hilbert-Ehrhart polynomial.

Models for $G = H = \mathbb{Z}_3$.

dilation	*snowflake*	*3-caterpillar*
1	243	243
2	21627	21627
3	903187	904069
4	21451311	21496023
5	330935625	331976637
6	3647265274	3662146270
7	30770591364	30920349834
8	209116329075	210269891871
9	1189466778457	1196661601837
10	5831112858273	5868930577941
11	25205348411361	25377886917819

Models for $G = H = \mathbb{Z}_2 \times \mathbb{Z}_2$ **(3-Kimura).**

dilation	*snowflake*	*3-caterpillar*
1	1024	1024
2	396928	396928
3	69248000	69324800
4	5977866515	5990170739
5	291069470720	291864710144
6	8967198289920	8995715702784

Models for $G = H = \mathbb{Z}_4$**.**

dilation	*snowflake*	*3-caterpillar*
1	1024	1024
2	396928	396928
3	69248000	69324800
4	6122557220	6138552524
5	310273545216	311525688320
6	10009786400352	10062179606880

Models for $G = H = \mathbb{Z}_5$**.**

dilation	*snowflake*	*3-caterpillar*
1	3125	3125
2	3834375	3834375
3	2229584375	2230596875
4	640338121875	642089603125

Models for $G = H = \mathbb{Z}_7$**.** In this case the first three dilations of the polytopes have the same number of points. The numbers of points in fourth dilations were too big to obtain precise results. Hence we computed only the numbers of points mod 64, which is sufficient to prove that the Hilbert-Ehrhart polynomials are different.

dilation	*snowflake*	*3-caterpillar*
1	16807	16807
2	117195211	117195211
3	423913952448	423913952448
4	$\equiv 54 \mod 64$	$\equiv 14 \mod 64$

Models for $G = H = \mathbb{Z}_8$**.**

dilation	*snowflake*	*3-caterpillar*
1	32768	32768
2	454397952	454397952
3	3375180251136	3375013036032

Models for $G = H = \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$**.**

dilation	*snowflake*	*3-caterpillar*
1	32768	32768
2	454397952	454397952
3	3375180251136	3375013036032

Models for $G = H = \mathbb{Z}_9$**.**

dilation	*snowflake*	*3-caterpillar*
1	59049	59049
2	1499667453	1499667453
3	20938605820263	20937202945056

Part 2. Semigroups associated to trivalent graphs

La deuxième partie concerne les variétés algébriques associées aux graphes trivalents pour le modèle de Jukes-Cantor binaire. Il s'agit d'un travail en commun avec Weronika Buczyńska, Jarosław Buczyński et Kaie Kubjas. La variété associée á un graphe peut être représentée par un semi-groupe gradué. Nous étudions les liens entre les propriétés du graphe et le semigroupe. Le théorème principal 14.1 borne le degré en lequel le semi-groupe est engendré par le premier nombre de Betti du graphe, plus un.

This part contains results of a joint work with W. Buczyńska, J. Buczyński and K. Kubjas. We use a generalization of the construction that associated a variety to a tree. We will be working with arbitrary trivalent graphs with possible loops and multiple edges between two vertices. However our study concerns only an equivalent of the binary Jukes-Cantor model.

14. Introduction

Let $\mathcal{G}$ be a trivalent graph. For a positive integer d, our main object of study will be a subset $\tau(\mathcal{G})_d$ of all labellings of edges of $\mathcal{G}$ by integers. A labelling is in $\tau(\mathcal{G})_d$, if the following conditions are satisfied:

- [♡♡] *(parity condition)* the sum of the three labels around each inner vertex is even;
- [+] *(non-negativity condition)* each label is non-negative;
- [△] *(triangle inequalities)* the three labels around each inner vertex satisfy the triangle inequalities;
- [°] *(degree inequalities)* the sum of the three labels around each inner vertex is at most $2d$.

We give more details and formal definitions in Section 15. We will be interested in $\tau(\mathcal{G}) = \bigsqcup_{d\in\mathbb{N}} \tau(\mathcal{G})_d$, which has a natural structure of a monoid by edgewise addition, and we call it the **phylogenetic monoid of $\mathcal{G}$**.

14.1. **Motivation.** The combinatorics of the monoid $\tau(\mathcal{G})$ associated to a trivalent graph $\mathcal{G}$ has several incarnations. Buczyńska studied it in [Buc12] as a generalization of the polytope defining the Cavender-Farris-Neyman [Ney71] model of a trivalent phylogenetic tree.

In more recent work Sturmfels and Xu, [SX10] found a universal object for the Cavender-Farris-Neyman model of trivalent trees with the same discrete invariants. More precisely, they proved that given the number of leaves n, the Cavender-Farris-Neyman model of a trivalent tree is a sagbi degeneration of the projective spectrum of the Cox ring of the blow-up of $\mathbb{P}^n$ in $n-3$ points. This variety is closely related to the moduli of quasiparabolic vector bundles on $\mathbb{P}^1$ with $n-2$ marked points.

Further work in this direction was done by Manon in [Man09] and [Man11]. He uses a sheaf of algebras over moduli spaces of genus g curves with n marked points coming from the conformal field theory. The case $g=0$ is the construction of [SX10], thus Manon's work generalises the Sturmfels-Xu construction. The monoid algebras $\mathbb{C}[\tau(\mathcal{G})]$ are the toric deformations of the algebras over the most special points in the moduli of curves in the Manon's construction. Here $\mathcal{G}$ is the dual graph of the reducible curve represented by a special point.

Jeffrey and Weitsmann in [JW92] study the moduli space of flat $SU(2)$-connections on a genus g Riemann surface. In their context a trivalent graph $\mathcal{G}$ describes the geometry of the compact surface of genus g with n marked points. They consider a subset of $\mathbb{Z}$-labellings of the graph, which is exactly $\tau(\mathcal{G})_d$. They prove that the number of elements in this set is equal to the number of Bohr-Sommerfeld fibres associated to $\mathcal{L}^{\otimes d}$, where $\mathcal{L}$ is a natural polarizing line bundle on the moduli space in question. The Bohr-Sommerfeld fibres are the central object of study in [JW92]. By the Verlinde formula, the number of those fibres equals the dimension of the space of holomorphic sections of $\mathcal{L}^{\otimes d}$. This number is the value of the Hilbert function of the toric model of a connected graph with the first Betti number$g$ and n leaves.

Thanks to the Verlinde formula, which arises from mathematical physics, the Hilbert function of the monoid algebras $\mathbb{C}[\tau(\mathcal{G})]$ has significant meaning. In the case of trivalent trees it was also used in [SX10] and then studied by Sturmfels and Velasco in [SV10]. One of the features of this model is that the Hilbert function depends only on the combinatorial data [BW07], [Buc12]. This phenomenon fails to be true for other, even general group-based models, see [Kub10] or Appendix 2 from Part 1.

14.2. **Main results.** If $\omega \in \tau(\mathcal{G})$, then there exists d, such that $\omega \in \tau(\mathcal{G})_d$. Such d is called the **degree** of ω. We are interested in the problem of determining the degrees of elements in the minimal set of generators of the monoid $\tau(\mathcal{G})$. We prove an upper bound for the degree of generators:

Theorem 14.1. *Let $\mathcal{G}$ be any trivalent graph with first Betti number g. Then the degree of each element in the minimal set of generators of $\tau(\mathcal{G})$ is at most $g+1$.*

For $g = 0$, that is $\mathcal{G}$ is a trivalent tree, this result is equivalent to statement that $\tau(\mathcal{G})_1$ is a normal lattice polytope and it has been obtained in [BW07]. For $g = 1$, the result has been obtained in [Buc12]. For $g \geq 2$ it has been previously unknown. We prove the theorem in Section 16.

The lower bounds were presented in [BBKM10]. Let us just state these results.

Theorem 14.2. *Suppose g is even. Then there exists a trivalent graph $\mathcal{G}$ with the first Betti number g and an element $\omega \in \tau(\mathcal{G})$ of degree $g+1$, which cannot be written in a non-trivial way as a sum of two elements $\omega = \omega' + \omega''$ with $\omega', \omega'' \in \tau(\mathcal{G})$. Specifically, $\mathcal{G}$ can be taken as the g-caterpillar graph (see Figure 2), and ω as the labelling on Figure 3.*

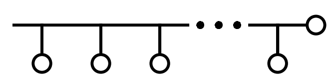


Figure 2: The g-caterpillar graph.

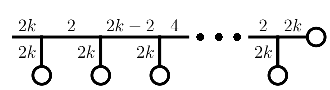


Figure 3: The indecomposable element ω of degree $g+1$ on the g-caterpillar graph for even g.

As for odd g, for all graphs with the first Betti number $g = 1$, the bound is also attained, as proved in [Buc12]. Also there exist graphs with $g = 3$, such that the bound is attained. The simplest of these is the 3-caterpillar graph and we illustrate an indecomposable degree 4 element in Section 17.

We also know the maximal degree of generators of the monoid for the g-caterpillar graph.

Corollary 14.3 ([BBKM10])**.** *For the g-caterpillar graph $\mathcal{G}$, the monoid $\tau(\mathcal{G})$ is generated in degree $g+1$ if either g is even or $g \leq 3$ and it is generated in degree g if $g \geq 5$ and g is odd.*

We present an indecomposable element of degree 6 on a graph with the first Betti number 6 in Section 17. We do not know, if there exist graphs with odd first Betti number $g \geq 5$ such that $\tau(\mathcal{G})$ has a degree $g+1$ minimal generator.

We also present the results of some computational experiments for g-caterpillar graphs with $g \leq 5$. Specifically, we list all the generators of $\tau(\mathcal{G})$ for $g \leq 4$ and enumerate these generators for $g \leq 5$.

15. Semigroup associated with a trivalent graph

In this section we recall the construction of the monoid $\tau(\mathcal{G})$ introduced in [Buc12].

A **graph** $\mathcal{G}$ is a set $\mathcal{V} = \mathcal{V}(\mathcal{G})$ of vertices and a set $\mathcal{E} = \mathcal{E}(\mathcal{G})$ of edges, which we identify with pairs of vertices. A graph is **trivalent** if every vertex has valency one or three. A vertex with valency one is called a **leaf** and an edge incident to a leaf is called a **leaf edge**. A vertex that has valency three is called an **inner vertex**. The set of inner vertices is denoted $\mathcal{N} = \mathcal{N}(\mathcal{G})$.

Notation 15.1. *From now on we shall assume that* ***all*** *graphs and trees are* ***trivalent****.*

The first Betti number of a graph is the minimal number of cuts that would make the graph into a tree. Given the origins of the problem explained in Section 14.1 it is tempting to call this number the *genus* of the graph, but this is inconsistent with the graph theory notation, where the genus of graph is the smallest genus of a Riemann surface such that the graph can be embedded into that surface.

A **path** is a sequence of pairwise distinct edges $e_0, \ldots, e_m$ with $e_i \cap e_{i+1} \neq \emptyset$ for all $i \in \{0, \ldots, m-1\}$, such that either both e_0 and e_m contain a leaf, or $e_0 \cap e_m \neq \emptyset$. In the latter case, if in addition the sequence has no repeated edges, the path is called a **cycle**. A cycle of length one is a **loop**. A trivalent graph with no cycles is a **trivalent**

tree. Two paths are **disjoint** if they have no common vertex. A **network** is a union of pairwise disjoint paths – cf. Remark 5.27 and the discussion afterwards. For consistency we say that the empty set is also a network. An edge which is contained in a cycle is called **cycle edge**.

Definition 15.2. *Given a graph $\mathcal{G}$ let $\mathbb{Z}\mathcal{E} = \bigoplus_{e\in\mathcal{E}} \mathbb{Z}\cdot e$ be the lattice spanned by $\mathcal{E}$, and $\mathbb{Z}\mathcal{E}^\vee = \mathrm{Hom}(\mathbb{Z}\mathcal{E}, \mathbb{Z})$ be its dual. Elements of the lattice $\mathbb{Z}\mathcal{E}$ are formal linear combinations of the edges, thus $\mathcal{E}$ forms the standard basis of $\mathbb{Z}\mathcal{E}$. The dual lattice $\mathbb{Z}\mathcal{E}^\vee$ comes with the dual basis $\{e^*\}_{e\in\mathcal{E}}$. We define*

$$M = \{u \in \mathbb{Z}\mathcal{E} : \forall v \in \mathcal{N} \sum_{e\ni v} e^*(u) \in 2\mathbb{Z}\}.$$

Then the graded lattice of the graph is

$$M^{gr} = \mathbb{Z} \oplus M$$

with the degree map

$$\deg : M^{gr} = \mathbb{Z} \oplus M \to \mathbb{Z},$$

given by the projection onto the first summand.

Remark 15.3. *An element of the lattice $\mathbb{Z}\mathcal{E}$ represents also a labelling of the edges of $\mathcal{G}$ with integers. For $\omega \in \mathbb{Z}\mathcal{E}$ the label of $e \in \mathcal{E}$ equals $e^*(\omega)$.*

Definition 15.4 (a_v, b_v, c_v)**.** *Let $v \in \mathcal{N}$ be an inner vertex and let e_1, e_2 and e_3 be the three edges[13] adjacent to v. For $\omega \in M^{gr}$ we define $a_v(\omega) = e_1^*(\omega)$, $b_v(\omega) = e_2^*(\omega)$, $c_v(\omega) = e_3^*(\omega)$.*

Definition 15.5 (degree)**.** *We define the **degree** of $\omega \in M^{gr}$ at an inner vertex $v \in \mathcal{N}$ as $\frac{1}{2}(a_v(\omega) + b_v(\omega) + c_v(\omega))$.*

We rewrite the definition of $\tau(\mathcal{G})$ given in Section 14 so that $\tau(\mathcal{G})$ is a submonoid of M^{gr}.

Definition 15.6. *For a graph $\mathcal{G}$ we define the **phylogenetic monoid** $\tau(\mathcal{G})$ **on** $\mathcal{G}$ to be the set of elements ω satisfying the following conditions:*

[♡♡] (parity condition) $\omega \in M^{gr}$;
[+] (non-negativity condition) $e^*(\omega) \geq 0$ *for any* $e \in \mathcal{E}$;
[△] (triangle inequalities) *For any inner vertex* $v \in \mathcal{N}$

$$|a_v(\omega) - b_v(\omega)| \leq c_v(\omega) \leq a_v(\omega) + b_v(\omega);$$

[13]If there is a loop at the vertex then $e_1 = e_2$.

[°] (degree inequalities) $\deg(\omega) \geq \deg_v(\omega)$ *for any* $v \in \mathcal{N}$.

The triangle inequalities [△] are symmetric and do not depend on the embedding i_v.

Remark 15.7. *If every edge of* $\mathcal{G}$ *contains at least one inner vertex, then the inequalities above imply* $\deg(\omega) \geq e^*(\omega)$ *for all edges. On the other hand, in the degenerate cases where one of the connected components of* $\mathcal{G}$ *consists of one edge only, for consistency the inequality* $\deg(\omega) \geq e^*(\omega)$ *should be included in Definition 15.6. However, we will not consider these degenerate cases here.*

To define a network in the graded lattice M^{gr}, we first have to do so in the lattice M: we identify paths and networks in $\mathcal{G}$ with elements of the lattice M by replacing union with sum in the group $\mathbb{Z}\mathcal{E}$.

Definition 15.8. *A* ***network in the graded lattice*** M^{gr} *is a pair* $\omega = (1, a) \in M^{gr}$ *where* $a \in M$ *is a network.*

Definition 15.9. *Following 5.24 we define the* ***group of networks*** *to be a subset of*

$$\mathbb{Z}_2\mathcal{E} := \bigoplus \mathbb{Z}_2 \cdot e$$

such that a formal sum in $e_1 + e_2 + \cdots + e_k \in \mathbb{Z}_2\mathcal{E}$ *is in the group of networks if and only if* $\{e_1, e_2, \ldots, e_k\}$ *is a network. Note that this subset forms a subgroup of* $\mathbb{Z}_2\mathcal{E}$.

16. The upper bound

The goal of this section is to prove Theorem 14.1. To do this, we proceed in three steps. First we recall the result of [BW07] that gives Theorem 14.1 in the case $g = 0$ (that is, if $\mathcal{G}$ is a tree). In the second step, we represent a graph $\mathcal{G}$ with first Betti number g as a tree $\mathcal{T}$ together with g distinguished pairs of leaf edges, that are "glued" together. Elements of $\tau(\mathcal{G})$ are in one-to-one correspondence with the elements of $\tau(\mathcal{T})$ that have identical labels on each of the distinguished pairs of leaf edges. Thus for an element $\omega \in \tau(\mathcal{G})$ we consider the decomposition of the corresponding element in $\tau(\mathcal{T})$ into a sum of degree 1 elements of $\tau(\mathcal{T})$. To each such decomposition we assign a matrix with entries in $\{-1, 0, 1\}$. Since the decomposition is not unique, we study how simple modifications of the decomposition affect the matrix. Finally, we apply these modifications to the matrix and prove that any sufficiently high degree element $\tau(\mathcal{G})$ decomposes. The details follow.

16.1. **The case of trees.** The set of degree 1 elements $\tau(\mathcal{G})_1 \subset \{1\} \times M \subset M^{gr}$ consists of networks — see [Buc12, Lem. 2.30].

The monoid $\tau(\mathcal{G})$ is the intersection of the convex polyhedral cone given by inequalities [+], [△], [°] with the lattice M^{gr} — see [♡♡]. If $\mathcal{T}$ is a trivalent tree, then the inequalities defining $\tau(\mathcal{T})_1$ define an integral lattice polytope P in $\{1\} \times M \subset M^{gr}$ — see [BW07, Lem. 2.8]. Furthermore, by [BW07, Prop. A.5] this polytope is *normal*, which means, that any lattice point in the rescaling nP can be obtained as sum of n lattice points in P. This implies that the monoid $\tau(\mathcal{T})$ is generated by $\tau(\mathcal{T})_1$. We summarize by quoting [Buc12, Prop. 2.32]:

Corollary 16.1. *Let $\mathcal{T}$ be a trivalent tree. Then every $\omega \in \tau(\mathcal{T})_d$ can be expressed as $\omega = \omega_1 + \cdots + \omega_d$, where each $\omega_i \in \tau(\mathcal{T})_1$ is a network.*

Note that usually the decomposition in the corollary is not unique.

16.2. **Matrix associated to a decomposition of a lifted element.** To a given connected graph $\mathcal{G}$ with first Betti number g we associate a tree $\mathcal{T}$ with g distinguished pairs of leaf edges. This procedure can be described inductively on g. If $g = 0$, then the graph is a tree with no distinguished pairs of leaf edges. For $g > 0$ we choose a cycle edge e. We divide e into two edges e' and e'' adding two vertices l' and l'' of valency 1. The edges e' and e'' form a distinguished pair of leaf edges. This procedure decreases the first Betti number by one and increases the number of distinguished pairs by one. Note that usually the resulting tree with distinguished pairs of leaf edges is not unique, however a tree with distinguished pairs of leaf edges encodes precisely one graph.

Let $\mathcal{G}$ be a graph and let $\mathcal{T}$ be an associated tree. There is a one-to-one correspondence between elements of $\tau(\mathcal{G})$ and the elements of $\tau(\mathcal{T})$ that assign the same value to the leaf edges in each distinguished pair. Thus we have the natural inclusion $\tau(\mathcal{G}) \subset \tau(\mathcal{T})$. See [Buc12, §2.2–2.3] for a more geometric interpretation of this inclusion.

Let ω be an element of $\tau(\mathcal{G})$. By Corollary 16.1, in the monoid $\tau(\mathcal{T})$ there exists a decomposition $\omega = \omega_1 + \cdots + \omega_{\deg(\omega)}$, where $\omega_i \in \tau(\mathcal{T})_1$. For each such decomposition we consider the matrix $B_{\omega_1,\dots,\omega_{deg(\omega)}}$ with $\deg(\omega)$ rows and g columns indexed by pairs of distinguished leaf edges. The entry in the i-th row and column indexed by a pair of distinguished leaf edges (e', e'') is $e'^*(\omega_i) - e''^*(\omega_i)$. Thus, since ω_i is a network $\omega_i(e) \in \{0, 1\}$ for any edge, entries of $B_{\omega_1,\dots,\omega_{deg(\omega)}}$ are only -1, 0 or 1.

The matrix $B_{\omega_1,\dots,\omega_{deg(\omega)}}$ depends on the tree $\mathcal{T}$ and on the decomposition of ω into the sum of degree one elements. An entry of $B_{\omega_1,\dots,\omega_{deg(\omega)}}$

is zero when the corresponding network is compatible on the corresponding distinguished pair of leaf edges. Our aim is to decompose any element ω with $\deg(\omega) > g+1$ in $\tau(\mathcal{G})$. This means that we are looking for decompositions in $\tau(\mathcal{T})$ that are compatible on the distinguished pairs of leaf edges. Hence, it is natural to consider matrices with as many entries equal to zero as possible.

Let ω be an element of $\tau(\mathcal{T})$. Let $\omega = \omega_1 + \cdots + \omega_{\deg(\omega)}$ be a decomposition of ω into networks. Let $B_{\omega_1,\dots,\omega_{deg(\omega)}}$ be the matrix with $\deg(\omega)$ rows corresponding to the decomposition. Notice that for any subset of indices $\{j_1,\dots,j_p\} \subset \{1,\dots,\deg(\omega)\}$ the following conditions are equivalent:

(i) the element $\omega_{j_1} + \cdots + \omega_{j_p}$ is in $\tau(\mathcal{G})$;
(ii) in each column of $B_{\omega_1,\dots,\omega_{deg(\omega)}}$ the sum of entries in rows $j_1,\dots,j_p$ is equal to zero.

Even if we start from a decomposable ω the associated matrix might not have this property; it depends upon the choice of decomposition of ω in $\tau(\mathcal{T})$. The following lemma shows how to change this decomposition in order to obtain a matrix with the required property.

Lemma 16.2. *Let ω be an element of $\tau(\mathcal{T})$. Consider all decompositions of ω and associated matrices. Let us choose a decomposition of $\omega = \omega_1 + \cdots + \omega_{deg(\omega)}$ that gives a matrix $B_{\omega_1,\dots,\omega_{deg(\omega)}}$ with as many zero entries as possible. Let us choose two entries in the matrix $B_{\omega_1,\dots,\omega_{deg(\omega)}}$ that are in the same column indexed by (e_1', e_1''). Suppose that they are equal, respectively, to 1 and -1. There exists a decomposition of ω that yields a matrix the same as $B_{\omega_1,\dots,\omega_{deg(\omega)}}$, except for those two entries, which are interchanged.*

Proof. Let $\omega = \omega_1 + \cdots + \omega_{\deg(\omega)}$ be the given decomposition. Without loss of generality we may assume that the entries are in the first and second row. Hence ω_1 associates to the edges e_1' and e_1'' values 0 and 1 respectively, and similarly ω_2 associates 1 and 0.

Let us consider all edges of the tree $\mathcal{T}$ on which the networks ω_1 and ω_2 disagree. These edges form the network S on the tree $\mathcal{T}$. In fact, $S = \omega_1 + \omega_2$, where the sum is taken in the group of networks. Define p_1 to be the unique path from S starting at e_1''. Suppose that we have constructed a sequence of paths $p_1,\dots,p_{m-1}$ for $m > 1$, where the first edge of p_i is e_i'' and the last is e_{i+1}' and (e_i', e_i'') is a distinguished pair for $i \in \{1,\dots,m-1\}$. We consider the following cases:

(i) If the edge e_m' is not paired, then we stop the construction. Otherwise we go to Case (ii).

(ii) If there is a distinguished pair (e'_m, e''_m) and ${e'_m}^*(\omega_1) \neq {e''_m}^*(\omega_1)$ or ${e'_m}^*(\omega_2) \neq {e''_m}^*(\omega_2)$ (i.e. at least one of the two entries in the column (e'_m, e''_m) is non-zero), then we stop the construction. Otherwise we go to Case (iii).

(iii) If there is a distinguished pair (e'_m, e''_m) and ${e'_m}^*(\omega_1) = {e''_m}^*(\omega_1)$, ${e'_m}^*(\omega_2) = {e''_m}^*(\omega_2)$, then ω_1 and ω_2 disagree on e''_m, and e''_m is in S. We define p_m to be the unique path from S starting from e''_m. Let e'_{m+1} be the other end of the path p_m. We increase m by 1 and start over from Case (i).

Let us notice that the constructed paths are distinct. In particular, the construction terminates. Indeed, each path p_{i+1} uniquely determines the path p_i. Hence the first path that would have been repeated is p_1. This is possible only if the previous path ends with e'_1. From the assumption, we would have been in Case (ii), hence the construction would terminate.

We define a network $b \subset S$ to be the network, which is the union of paths $(p_1, \dots, p_{m-1})$. We use it to define two new networks ω'_1 and ω'_2. Namely, $\omega'_i = \omega_i + b$, where the sum is taken in the group of networks. In other words, ω'_1 (resp. ω'_2) coincides with ω_1 (resp. ω_2) on all edges apart from those belonging to the network b. On the latter ones ω'_1 (resp. ω'_2) is a negation of ω_1 (resp. ω_2), hence coincides with ω_2 (resp. ω_1). In particular, $\omega_1 + \omega_2 = \omega'_1 + \omega'_2$, where this time the sum is taken in $\tau(\mathcal{T})$.

We get a decomposition $\omega = \omega'_1 + \omega'_2 + \omega_3 + \dots + \omega_{\deg(\omega)}$ with the associated matrix $B_{\omega'_1,\dots,\omega_{deg(\omega)}}$. We claim that it exchanges the two chosen entries equal to 1 and -1.

Consider each distinguished pair of leaf edges through which we passed during our construction of $(p_1, \dots, p_{m-1})$. If we did not stop at a pair (l_1, l_2) each network a_1 and a_2 assigns the same value to l_1 and l_2 — otherwise we would have stopped because of Case (ii). On these leaf edges ω'_1 and ω'_2 agree with ω_2 and ω_1 respectively. Hence, they also assign the same value to l_1 and l_2. In particular, both $B_{\omega_1,\dots,\omega_{deg(\omega)}}$ and $B_{\omega'_1,\dots,\omega_{deg(\omega)}}$ have zeros in first two rows in the column indexed by (l_1, l_2). In fact, the only four entries on which $B_{\omega_1,\dots,\omega_{deg(\omega)}}$ and $B_{\omega'_1,\dots,\omega_{deg(\omega)}}$ might possibly differ are the entries in first two rows in the columns indexed by (e'_1, e''_1) or (e'_m, e''_m), where p_m is the last path.

Let us exclude the possibility that the construction stopped in Case (i). In this case the last leaf edge is not paired, hence we only change entries in the column indexed by (e'_1, e''_1). Since both ω'_1 and ω'_2 agree on

e'_1 and e''_1, we have that $B_{\omega'_1,\ldots,\omega_{deg(\omega)}}$ has two zeros, whereas $B_{\omega_1,\ldots,\omega_{deg(\omega)}}$ had 1 and -1. This contradicts the choice of $B_{\omega_1,\ldots,\omega_{deg(\omega)}}$.

Suppose now that the construction terminated in Case (ii). We consider two sub-cases.

1) The edges $e'_m \neq e'_1$ are distinct. We will exclude this case. We change four entries in two columns. The two entries in the column indexed by (e'_1, e''_1) are changed from 1 and -1 to zero. We know that matrix $B_{\omega'_1,\ldots,\omega_{deg(\omega)}}$ has at most as many zero entries as $B_{\omega_1,\ldots,\omega_{deg(\omega)}}$. Hence the two entries in the column indexed by (e'_m, e''_m) must be changed from two zeros to two non-zeros. Having two zeros in $B_{\omega_1,\ldots,\omega_{deg(\omega)}}$ in those entries contradicts the assumptions of Case (ii).

2) The edges $e'_m = e'_1$ are equal. In this case $e''_m = e''_1$, so we only exchange two entries in the column indexed by (e'_1, e''_1). This means that we have exchanged 1 and -1, which proves the lemma. □

16.3. **Proof of decomposability.** We are ready to prove the theorem on the upper bound of the degree of minimal generators of $\tau(\mathcal{G})$.

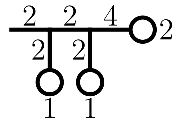
Proof of Theorem 14.1. Let us consider an element ω of degree $\deg(\omega) > g+1$ in $\tau(\mathcal{G})$. We consider any tree $\mathcal{T}$ associated with the graph $\mathcal{G}$. Let us choose a decomposition of ω in $\tau(\mathcal{T})$, such that the associated matrix $B_{\omega_1,\ldots,\omega_{deg(\omega)}}$ has as many zero entries as possible. First we want to find a subset of rows of the matrix $B_{\omega_1,\ldots,\omega_{deg(\omega)}}$ such that the sum of entries in each column is even. We reduce the entries of $B_{\omega_1,\ldots,\omega_{deg(\omega)}}$ modulo 2 obtaining the matrix C_ω with entries from $\mathbb{Z}_2$. We consider rows of C_ω as vectors of the g dimensional vector space over the field $\mathbb{Z}_2$. We have $\deg(\omega) > g+1$ such vectors. Hence we can find a *strict* subset of linearly dependent vectors. As we are working over $\mathbb{Z}_2$ we see that we have a strict subset of vectors summing to 0. The same subset R of rows in matrix $B_{\omega_1,\ldots,\omega_{deg(\omega)}}$ sums to even numbers in each column.

The element ω is in $\tau(\mathcal{G})$. Hence the sum of entries in each column of the matrix $B_{\omega_1,\ldots,\omega_{deg(\omega)}}$ is zero. Suppose that the sum of entries in the rows from R is non-zero in a column. Using Lemma 16.2 we can exchange the entries, changing the sum by 2 until it is equal to zero. In this way we get a decomposition of ω such that the rows from R sum to zero in each column. Hence, the sum of networks corresponding to rows from R is in $\tau(\mathcal{G})$. The sum of the remaining networks is in $\tau(\mathcal{G})$ too. Thus we obtain a non-trivial decomposition of ω. □

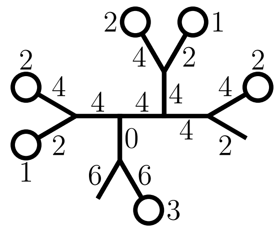
17. Examples on small graphs

We conclude this part with some examples of indecomposable elements for special cases of graphs with low first Betti number g.

The first one is an indecomposable element of degree 4 on the 3-caterpillar graph. It proves that in the case $g = 3$, the upper bound of Theorem 14.1 is attained.



The second example is a degree 6 indecomposable element on a graph with 6 loops and one leaf.



The following table presents the numbers of generators of $\tau(\mathcal{G})$ in each degree, where $\mathcal{G}$ is the g-caterpillar graph, and $g \leq 5$.

deg	$g=1$	$g=2$	$g=3$	$g=4$	$g=5$
all	3	15	163	2708	49187
1	2	4	8	16	32
2	1	7	37	175	781
3		4	64	704	6624
4			54	1701	35190
5				112	6560

Part 3. **Derived categories**

Dans la dernière partie, nous étudions la structure de la catégorie dérivée des faisceaux cohérents des variétés toriques lisses. Dans un travail commun avec Michał Lasoń [LM11], nous construisons une collection fortement exceptionnelle complète de fibrés en droites pour une grande classe de variétés toriques complètes lisses dont le nombre de Picard est égal á trois. De nombreuses questions concernant le type de collections auxquelles on peut s'attendre sur les variétés toriques de certains types sont encore ouvertes. A ce titre, nous prouvons que $\mathbb{P}^n$ éclaté en deux points ne possède pas de collection fortement exceptionnelle complète de fibrés en droites pour n assez grand. Ceci fournit une collection infinie de contre-exemples à la conjecture de King 19.2. Le premier contre-exemple est dû à Hille et Perling [HP06]. Récemment, des contre-exemples ont également été trouvés par Efimov [Efi] dans le cadre des variétés de Fano.

18. Introduction

18.1. **Definition of the derived category.** Let X be a smooth variety over the field $\mathbb{C}$. Let us briefly recall the construction of the derived category of X. We encourage the reader to consult first chapters of [Huy06] for precise definitions, examples and most important theorems. A well-motivated, relatively short introduction to derived categories can be found in [Căl05]. A much longer, classical reference is [GM03].

We start the construction with the category Kom of complexes of coherent sheaves on X. Let us introduce the homotopy category K of complexes. The objects in K are the same as in Kom. We identify morphisms that are homotopically equivalent.

Definition 18.1 (Homotopically equivalent morphisms of complexes, Definition 2.12 [Huy06])**.** *Let us consider two complexes*

$$\cdots \longrightarrow A_i \xrightarrow{\delta_i} A_{i+1} \xrightarrow{\delta_{i+1}} \cdots$$

$$\cdots \longrightarrow B_i \xrightarrow{\delta'_i} B_{i+1} \xrightarrow{\delta'_{i+1}} \cdots$$

and two morphisms f, g between them with components given by $f_i, g_i : A_i \to B_i$. We say that f and g are homotopically equivalent if and only if there exists a collection of morphisms $h_i : A_i \to B_{i-1}$ such that

$$f_i - g_i = h_{i+1} \circ \delta_i - \delta'_{i-1} \circ h_i.$$

The relation of being homotopically equivalent is an equivalence relation. A morphism in the category K is an equivalence class of morphisms up to this relation.

Recall that a morphism f between complexes $A, B \in Kom$ induces a morphism in cohomology

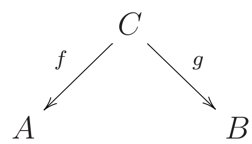
$$H^i(f) : \frac{\operatorname{Ker} \delta_i}{\operatorname{Im} \delta_{i-1}} =: H^i(A) \to H^i(B).$$

Moreover if f and g are homotopically equivalent, then $H^i(f) = H^i(g)$. Hence given a morphism in K we have the well-defined induced morphism on cohomologies.

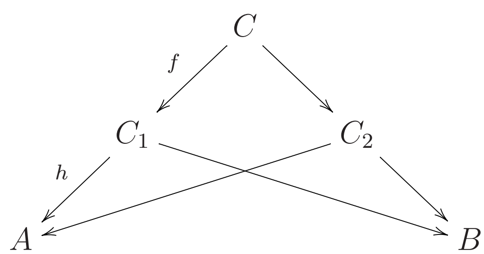
Definition 18.2 (Quasi isomorphism)**.** *A morphism between complexes (in Kom or K) is called a quasi-isomorphism if the induced morphism on cohomologies is an isomorphism.*

The objects of the derived category $D(X)$ will be complexes of coherent sheaves. However the morphisms in the derived category are defined differently.

Definition 18.3 (Morphism in the derived categories)**.** A roof *(between A and B) is the following diagram:*



where A, B, C are complexes, f, g are morphisms in K and f is a quasi-isomorphism. Two roofs between A and B are called equivalent if they can be dominated in K by a common roof. More precisely consider two roofs for which the domains of the morphisms are given respectively by C_1 and C_2. These roofs are equivalent if and only if there exists the following commutative diagram in K:



with $h \circ f$ a quasi-isomorphism. A morphism in the derived category $D(X)$ is an equivalence class of roofs. In particular, one can show that a composition of roofs is also given by a roof that dominates them.

The construction seems, and indeed is, quite technical. In [Căl05] the author motivates the construction by topology, especially the theorem of Whitehead. One of the aims of the construction is to make quasi-isomorphisms, real isomorphisms. The process of adding inverse morphisms to the category is called localisation[14]. However, for the localization process to work well one should pass from the category of complexes Kom to the category K. Indeed, the derived category can be regarded as the smallest possible category obtained from Kom by adding inverses of quasi-isomorphisms. Formally this can be characterized by a universal property [Huy06, Theorem 2.10] that the derived category satisfies.

We will be mostly interested in the bounded derived category $D^b(X)$. To define it one can repeat the construction of $D(X)$ starting not from

[14]Indeed it is similar to the algebraic localization where we add formal inverses of elements.

the category Kom, but a subcategory of complexes with only a finite number of nonzero elements. The category $D^b(X)$ is equivalent to a full subcategory of $D(X)$ containing complexes with only a finite number of nonzero cohomology.

Details of category theory are beyond the scope of this thesis. However, we should mention that the derived category *is not an abelian category*. Thus one cannot speak about exact sequences. Still, some diagrams, called *the distinguished triangles* exist and play a role similar to short exact sequences. This gives the derived category the structure of the *triangulated category*. When we say that two derived categories are equivalent, we assume that the exact triangles are preserved. Formally by an equivalence we mean *an exact equivalence*. Details and definitions can be found in any book on derived categories or homological algebra.

On the one hand, the derived categories give a uniform language that allows to state many definitions, useful from the point of view of algebraic geometry. On the other hand, the structure of the derived category can be extremely complicated and is an object of intensive studies. In some cases one can consider a collection of objects from the derived category that plays a role of the "basis" of the derived category. The following sections investigate when such special collections exist.

We have to note, that the derived category of an algebraic manifold does not fully characterize it. Indeed, the questions how subtile the derived category is as an invariant, is one of the most important one in the domain. Let us present two well-known results.

Theorem 18.4 (Bridgeland [Bri02]). *Any two birational Calabi-Yau threefolds have got equivalent derived categories.* □

Theorem 18.5 (Bondal, Orlov [BO01]). *Let X and Y be smooth projective varieties and assume that the (anti-)canonical bundle of X is ample. If there exists an equivalence $D^b(X) \simeq D^b(Y)$, then X and Y are isomorphic.* □

18.2. **Full, strongly exceptional collections.** The structure and properties of the derived category of an arbitrary variety X can be very complicated and they are an object of many studies. One of the approaches to understand the derived category uses the notion of exceptional objects. Let us introduce the following definitions (see also [GR87]):

Definition 18.6 (Strongly exceptional collection).

*(i) An object of the bounded derived category $D^b(X)$ of X is **exceptional** if $Hom(F,F) = \mathbb{K}$ and $Ext^i_{\mathcal{O}_X}(F,F) = 0$ for $i \neq 0$.*

(ii) *An ordered collection* $(F_0, F_1, \ldots, F_m)$ *of objects is an* ***exceptional collection*** *if each sheaf* F_i *is exceptional and* $Ext^i_{\mathcal{O}_X}(F_k, F_j) = 0$ *for* $j < k$ *and* $i \geq 0$.

(iii) *An exceptional collection* $(F_0, F_1, \ldots, F_m)$ *of objects is a* ***strongly exceptional collection*** *if* $Ext^i_{\mathcal{O}_X}(F_j, F_k) = 0$ *for* $j \leq k$ *and* $i \geq 1$.

(iv) *A (strongly) exceptional collection* $(F_0, F_1, \ldots, F_m)$ *of objects is a* ***full, (strongly) exceptional collection*** *if it generates the bounded derived category* $D^b(X)$ *of* X *i.e. the smallest triangulated category containing* $\{F_0, F_1, \ldots, F_n\}$ *is equivalent to* $D^b(X)$.

For an exceptional collection $(F_0, \ldots, F_m)$ one may define an object $F = \oplus_{i=0}^m F_i$ and an algebra $A = \mathrm{Hom}(F, F)$. Such an object gives us a functor G_F from $D^b(X)$ to the derived category $D^b(A - mod)$ of right finite-dimensional modules over the algebra A. Bondal proved in [Bon89], that if X is smooth and (F_i) is a full, strongly exceptional collection, then the functor G_F gives an equivalence of these categories. In further sections we will be mostly interested in the strongly exceptional collections.

19. Toric varieties and exceptional collections

In the whole section we assume that X is a smooth toric variety. In particular X is normal, thus given by a fan.

19.1. **Known results and counterexamples.** As the structure of derived categories is complicated it is natural to look at examples of toric varieties. In particular, exceptional collections for toric varieties have been an object of studies. The strongest positive result is due to Kawamata [Kaw06].

Theorem 19.1 ([Kaw06])**.** *For any smooth, projective toric variety* X*, the bounded derived category* $D^b(X)$ *is generated by an exceptional collection of coherent sheaves.*

Due to the fact that line bundles have got a particularly nice description for toric varieties one could ask whether "coherent sheaves" in previous theorem can be replaced by line bundles [Huy06, Remark 8.38]. This is an open problem. However, there exists a counter example to the following conjecture of King.

Conjecture 19.2 (King's)**.** *For any smooth, complete toric variety* X *there exists a full, strongly exceptional collection of line bundles.*

The first known counterexample was presented by Hille and Perling in [HP06]. They gave an example of a smooth, complete toric surface which does not have a full, strongly exceptional collection of line bundles. Further results gave a full description of the two dimensional case [HP]. In Section 19.7 we show that $\mathbb{P}^n$ blown up in two points for n large enough are also counterexamples to King's conjecture. The conjecture was reformulated by Miró-Roig and Costa (stated also in [BH09]):

Conjecture 19.3. *For any smooth, complete Fano toric variety there exists a full, strongly exceptional collection of line bundles.*

This conjecture has an affirmative answer when the Picard number of X is less than or equal to two [CMR04] or the dimension of X is at most three [BH09], [Bon89], [BT09]. Recently it was disproved by Efimov in [Efi]. In the same paper the author states the following conjecture, suggested by D. Orlov.

Conjecture 19.4 ([Efi])**.** *For any smooth projective toric Deligne–Mumford stack Y, the derived category $D^b(Y)$ is generated by a strongly exceptional collection.*

We will often make use of the construction of a collection of line bundles due to Bondal. The construction is described in Section 19.3. Using this, one obtains a full collection of line bundles in $D^b(X)$. In some cases Bondal's collection of line bundles is a strongly exceptional collection (see also [Bon06]), but it is not true in general. Often one can find a subset of this collection and order it in such a way that it becomes strongly exceptional and remains full. This approach was well described in [CMRb] for a class of toric varieties with Picard number three.

19.2. **Toric varieties with Picard number three.** Smooth, complete toric varieties with Picard number three have been classified by Batyrev in [Bat91] according to their primitive relations. Let Σ be a fan in $N = \mathbb{Z}^n$ and let R be the set of rays of Σ.

Definition 19.5. *We say that a subset $P \subset R$ is a primitive collection if it is a minimal subset of R which does not span a cone in Σ.*

In other words a primitive collection is a subset of ray generators, such that all together they do not span a cone in Σ but if we remove any generator, then the rest spans a cone that belongs to Σ. To each primitive collection $P = \{x_1, \dots, x_k\}$ we associate a primitive relation. Let $w = \sum_{i=1}^{k} x_i$. Let $\sigma \in \Sigma$ be the cone of the smallest dimension that contains w and let $y_1, \dots, y_s$ be the ray generators of this cone. The toric variety of Σ was assumed to be smooth, so there are unique positive integers $n_1, \dots, n_s$ such that

$$w = \sum_{i=1}^{s} n_i y_i.$$

Definition 19.6. *For each primitive collection $P = \{x_1, \dots, x_k\}$ let n_i and y_i be as described above. The linear relation:*

$$x_1 + \dots + x_k - n_1 y_1 - \dots - n_s y_s = 0$$

is called the primitive relation (associated to P).

Using the results of [Grü03] and [OP91] Batyrev proved in [Bat91] that for any smooth, complete n dimensional fan with $n+3$ generators its set of ray generators can be partitioned into l non-empty sets $X_0, \dots, X_{l-1}$ in such a way that the primitive collections are exactly sums of $p+1$ consecutive sets X_i (we use a circular numeration, that is we assume that $i \in \mathbb{Z}/l\mathbb{Z}$), where $l = 2p+3$. Moreover l is equal to 3 or 5. The number l is of course the number of primitive collections. In the case $l = 3$ the fan Σ is a splitting fan (that is any two primitive collections are disjoint). These varieties are well characterized, and we know much about full, strongly exceptional collections of line bundles on them. The case of five primitive collections is much more complicated and is our object of study. For $l = 5$ we have the following result of Batyrev.

Theorem 19.7 ([Bat91, Theorem 6.6])**.** *Let $Y_i = X_i \cup X_{i+1}$, where $i \in \mathbb{Z}/5\mathbb{Z}$,*

$$X_0 = \{v_1, \dots, v_{p_0}\}, \quad X_1 = \{y_1, \dots, y_{p_1}\}, \quad X_2 = \{z_1, \dots, z_{p_2}\},$$
$$X_3 = \{t_1, \dots, t_{p_3}\}, \quad X_4 = \{u_1, \dots, u_{p_4}\},$$

where $p_0 + p_1 + p_2 + p_3 + p_4 = n + 3$. Then any n-dimensional fan Σ with the set of generators $\bigcup X_i$ and five primitive collections Y_i can be described up to a symmetry of the pentagon by the following primitive relations with nonnegative integral coefficients $c_2, \dots, c_{p_2}, b_1, \dots, b_{p_3}$:

$$v_1 + \dots + v_{p_0} + y_1 + \dots + y_{p_1} - c_2 z_2 - \dots - c_{p_2} z_{p_2} - (b_1+1) t_1 - \dots - (b_{p_3}+1) t_{p_3} = 0,$$
$$y_1 + \dots + y_{p_1} + z_1 + \dots + z_{p_2} - u_1 - \dots - u_{p_4} = 0,$$
$$z_1 + \dots + z_{p_2} + t_1 + \dots + t_{p_3} = 0,$$
$$t_1 + \dots + t_{p_3} + u_1 + \dots + u_{p_4} - y_1 - \dots - y_{p_1} = 0,$$
$$u_1 + \dots + u_{p_4} + v_1 + \dots + v_{p_0} - c_2 z_2 - \dots - c_{p_2} z_{p_2} - b_1 t_1 - \dots - b_{p_3} t_{p_3} = 0.$$

□

In this case we may assume that

$$v_1, \dots, v_{p_0}, y_2, \dots, y_{p_1}, z_2, \dots, y_{p_2}, t_1, \dots, t_{p_3}, u_2, \dots, u_{p_4}$$

form a basis of the lattice N. The other vectors are given by

$$\begin{aligned} z_1 &= - z_2 - \cdots - z_{p_2} - t_1 - \cdots - t_{p_3} \\ y_1 &= - y_2 - \cdots - y_{p_1} - z_1 - \cdots - z_{p_2} + u_1 + \cdots + u_{p_4} \\ u_1 &= - u_2 - \cdots - u_{p_4} - v_1 - \cdots - v_{p_0} + c_2 z_2 + \cdots + c_{p_2} z_{p_2} \\ &\quad + b_1 t_1 + \cdots + b_{p_3} t_{p_3} \end{aligned} \tag{19.1}$$

19.3. **Bondal's construction and Thomsen's algorithm.** This section contains joint results with Michał Lasoń [LM11].

We start this section by recalling Thomsen's [Tho00] algorithm for computing the summands of the push forward of a line bundle by a Frobenius morphism. We do this because of two reasons.

First is that Thomsen in his paper assumes finite characteristic of the ground field and uses absolute Frobenius morphism. We claim that the arguments used apply also in case of geometric Frobenius morphism and characteristic zero.

Moreover by recalling all methods we are able to show that the results of Thomsen coincide with the results stated by Bondal in [Bon06]. Combining both methods enables us to deduce some interesting facts about toric varieties. A reader interested in a short proof and a method for the decomposition of the push forward of a line bundle by a Frobenius morphism is advised to consult [Ach].

Most of the results of this section are due to Bondal and Thomsen. We use the notation from [Tho00]. Let $\Sigma \subset N$ be a fan such that the toric variety $X = X(\Sigma)$ is smooth. Let us denote by $\sigma_i \in \Sigma$ the cones of our fan and by T the torus of our variety. If we fix a basis $(e_1, \dots, e_n)$ of the lattice N, then of course $T = \operatorname{Spec} R$, where $R = k[X_{e_1^*}^{\pm 1}, \dots, X_{e_n^*}^{\pm 1}]$.

In characteristic p we have got two p-th Frobenius morphisms $F : X \to X$. One of them is the absolute Frobenius morphism given as an identity on the underlying topological space and a p-th power on structure sheaves. Notice that on the torus it is given by a map $R \to R$ that is simply a p-th power map, hence it is not a morphism of k algebras (it is not an identity on k).

The other morphism is called the geometric Frobenius morphism and can be defined in any characteristic. Let us fix an integer m. Consider a morphism of tori $T \to T$ that associates t^m to a point t. This is a morphism of schemes over k that can be extended to the m-th geometric Frobenius morphism $F : X \to X$. What is important is that both of these morphisms can be considered as endomorphisms of open affine subsets associated to cones of Σ. We claim that in both cases Thomsen's algorithm works.

We begin by recalling the algorithm from [Tho00]. Let $v_{i1}, \dots, v_{id_i}$ be the ray generators of the d_i dimensional cone σ_i. As the variety was assumed to be smooth we may extend this set to a basis of N. Let A_i be a square matrix whose rows are vectors v_{ij} in the fixed basis of N. Let $B_i = A_i^{-1}$ and let w_{ij} be the j-th column of B_i. Of course the columns of B_i are ray generators (extended to a basis) of the dual cone $\sigma_i^* \subset M = N^*$.

Let us remind that $X(\Sigma)$ is covered by affine open subsets $U_{\sigma_i} = \operatorname{Spec} R_i$, where $R_i = k[X^{w_{i1}}, \dots, X^{w_{id_i}}, X^{\pm w_{id_i+1}}, \dots, X^{\pm w_{in}}]$. Here we use the notation $X^v = X_{e_1^*}^{v_1} \cdot \dots \cdot X_{e_n^*}^{v_n}$. Let also $X_{ij} = X^{w_{ij}}$. In this way the monomials $X_{i1}, \dots, X_{in}$ should be considered as coordinates on the affine subset U_{σ_i}, so we are able to think about monomials on U_{σ_i} as vectors: a vector v corresponds to the monomial X_i^v. Of course all of these affine subsets contain T, that corresponds to the inclusions $R_i \subset R$.

Using basic results from toric geometry, see [Ful93, p. 21], we know that $U_{\sigma_i} \cap U_{\sigma_j} = U_{\sigma_i \cap \sigma_j}$ and this is a principal open subset of U_{σ_i}. This means that there is a monomial M_{ij} such that $U_{\sigma_i \cap \sigma_j} = \operatorname{Spec}((R_i)_{M_{ij}})$.

We are interested in Picard divisors. A T invariant Picard divisor is given by a compatible collection $\{(U_{\sigma_i}, X_i^{u_i})\}_{\sigma_i \in \Sigma}$. Compatible means that the quotient of any two functions in the collection is invertible on the intersection of domains. This motivates the definition:

$$I_{ij} = \{v : X_i^v \text{ is invertible in } (R_i)_{M_{ij}}\}.$$

Given a monomial X_i^v, if we want to know how it looks in coordinates $X_{e_1^*}, \dots, X_{e_n^*}$ (obviously from the definition of X_i) we just have to multiply v by B_i: $X_i^v = X^{B_i v}$. We see that $X_i^v = X_j^{B_j^{-1} B_i v}$. That is why we define $C_{ij} = B_j^{-1} B_i$ and we think of C_{ij} as the matrices that translate the monomials in coordinates of one affine piece to another.

Now the compatibility in the definition of a Cartier divisor simply is equivalent to the condition $u_j - C_{ij} u_i \in I_{ji}$. We define $u_{ij} = u_j - C_{ij} u_i$ and think about them as transition maps. Of course a divisor is principal if and only if $u_{ij} = 0$ for all i, j (vector equal to 0 corresponds to a constant function equal to 1).

Let $P_m = \{v = (v_1, \dots, v_n) : 0 \leq v_i < m\}$. Later we will see that this set has got a description in terms of characters of the kernel of the Frobenius map between tori.

Using simple algebra Thomsen proves that the following functions are well defined (the only thing to prove is that the image of h is in I_{ji}).

Definition 19.8 (h^w_{ijm}, r^w_{ijm}). *Let us fix $w \in I_{ji}$ and a positive integer m. We define the functions*

$$h^w_{ijm} : P_m \to I_{ji}$$

$$r^w_{ijm} : P_m \to P_m,$$

for any $v \in P_m$ by the equation

$$C_{ij}v + w = mh^w_{ijm}(v) + r^w_{ijm}(v).$$

This is a simple division by m with the rest. Moreover r^w_{ijm} is bijective.

If we have any $v \in P_m$, a T-Cartier divisor $D = \{(U_{\sigma_i}, X_i^{u_i})\}_{\sigma_i \in \Sigma}$ and a fixed $\sigma_l \in \Sigma$ then Thomsen defines $t_i = h^{u_{li}}_{lim}(v)$. He proves that the collection $\{(U_{\sigma_i}, X_i^{t_i})\}_{\sigma_i \in \Sigma}$ is a T-Cartier divisor D_v. This is of course independent on the representation of D up to linear equivalence. The choice of l corresponds to "normalizing" the representation of D on the affine subset U_{σ_l}. Although the definition of D_v may depend on l, the vector bundle $\oplus_{v \in P_m}\mathcal{O}(D_v)$ is independent on l. Moreover Thomsen proves that in case of p-th absolute Frobenius morphism and characteristic $p > 0$ this vector bundle is a push forward of the line bundle $\mathcal{O}(D)$. The proof uses only the fact that the Frobenius morphism can be considered as a morphism of affine pieces U_{σ_i}, so can be extended to the case of geometric Frobenius morphism and arbitrary characteristic. One only has to notice that the basis of free modules obtained by Thomsen in [Tho00, Section 5, Theorem 1] are exactly the same in all cases.

Now let us remind that there is an exact sequence 2.1:

$$0 \to M \to D_T \to Pic \to 0,$$

where D_T are T invariant divisors. Let (g_j) be the collection of ray generators of the fan Σ and D_{g_j} a divisor associated to the ray generator g_j. The morphism from M to D_T is given by $v \to \sum_j v(g_j)D_{g_j}$. Such a map may be extended to a map from $M_\mathbb{R} = M \otimes_\mathbb{Z} \mathbb{R}$ by $f : v \to \sum_j [v(g_j)]D_{g_j}$. Notice that this is no longer a morphism, however if $a \in M$ and $b \in M_\mathbb{R}$, then $f(a + b) = f(a) + f(b)$. We obtain a map $\mathbf{T} := \frac{M_\mathbb{R}}{M} \to Pic$, where $\mathbf{T}$ is a real torus (do not confuse with T). We also fix the notation for an $\mathbb{R}$-divisor $D = \sum_j a_j D_{g_j}$:

$$[D] := \sum_j [a_j]D_{g_j}.$$

Let G be the kernel of the m-th geometric Frobenius morphism between the tori T. By acting with the functor $\mathrm{Hom}(\cdot, \mathbb{C}^*)$ we obtain an exact sequence:

$$0 \to M \to M \to G^* \simeq \frac{M}{mM} \to 0.$$

We also have a morphism:

$$\frac{1}{m} : G^* \simeq \frac{M}{mM} \to \mathbf{T},$$

that simply divides the coordinates by m. By composing it with the morphism from $\mathbf{T} \to Pic$ we get a morphism from G^* to Pic. It can be also described as follows:

We fix $\chi \in G^*$ and arbitrarily lift it to an element $\chi_M \in M$. Now we use the morphism $M \to Div_T$ to obtain a T invariant principal divisor D_χ. The image of χ in Pic is simply equal to $[\frac{D_\chi}{m}]$. Of course for different lifts of χ to M we get linearly equivalent divisors. Now we prove one of the results stated by Bondal in [Bon06]:

Proposition 19.9. *Let $L = \mathcal{O}(D)$ by any line bundle on a smooth toric variety X. The push forward $F_*(\mathcal{O}(D))$ is equal to $\oplus_{\chi \in G^*} \mathcal{O}([\frac{D+D_\chi}{m}])$.*

Remark 19.10. *The characters of G play the role of $v \in P_m$ in Thomsen's algorithm. Notice also that it is not clear that $\oplus_{\chi \in G^*} \mathcal{O}([\frac{D+D_\chi}{m}]$ is independent on the representation of L by D. If we prove that this is equal to the push forward then this fact will follow, but in the proof we have to take any representation of L and we cannot change D with a linearly equivalent divisor.*

Proof. Let $D = \{(U_{\sigma_i}, X_i^{u_i})\}$ and let us fix $\chi \in G^*$. We have to prove that $\mathcal{O}([\frac{D+D_\chi}{m}])$ is one of $\mathcal{O}(D_v)$ for $v \in P_m$ and that this correspondence is one to one over all $\chi \in G^*$. We already know that $[\frac{D_\chi}{m}]$ is independent on the choice of the lift of χ, so we may take such a lift, that $v = \chi_M + u_l$ is in the P_m. Here l is an index of a cone, but we may assume that its ray generators form a standard basis of N, so $A_l = Id$. Of course such a matching between $\chi \in G^*$ and $v \in P_m$ is bijective.

Let us compare the coefficients of $[\frac{D+D_\chi}{m}]$ and D_v. We fix a ray generator $r = (r_1, \dots, r_n) \in \sigma_j$. Let k be such that this ray generator is the k-th row of matrix A_j. We compare coefficients of D_r. Let $\chi_M = (a_1, \dots, a_n)$. We see that:

$$\left[\frac{D + D_\chi}{m}\right] = \dots + \left[\frac{(u_j)_k + \sum_{w=1}^n a_w r_w}{m}\right] D_r + \dots.$$

Here of course $(u_j)_k$ is not a transition map u_{jk}, but the k-th entry of vector u_j that is of course the coefficient of D_r of the divisor D. From Thomsen's algorithm described above we know that

$$C_{lj}(\chi + u_l) + u_{lj} = mt_j + r,$$

where $r \in P_m$. We see that

$$t_j = \left[\frac{C_{lj}(\chi + u_l) + u_{lj}}{m}\right].$$

Now $A_l = Id$ and from the definition of u_{lj} we have $C_{lj}u_l + u_{lj} = u_j$, so:

$$t_j = \left[\frac{A_j\chi + u_j}{m}\right].$$

This gives us:

$$D_v = \dots + \left[\frac{\sum_{w=1}^n a_w r_w + (u_j)_k}{m}\right] D_r + \dots$$

what completes the proof. □

From [Bon06] we know that the image B of $\mathbf{T}$ in Pic is a full collection of line bundles. Of course B is a finite set (the coefficients of divisors associated to ray generators are bounded). Moreover the image of rational points of $\mathbf{T}$ contains the whole image of $\mathbf{T}$ (a set of equalities and inequalities with rational coefficients has got a solution in $\mathbb{R}$ if and only if it has got a solution in $\mathbb{Q}$). This means that for sufficiently large m the split of the push forward of the trivial bundle by the m-th Frobenius morphism coincides with the image of $\mathbf{T}$ and hence is full.

Let us consider an example of $\mathbb{P}^2$. Let v_1, v_2 and $v_3 = -v_1 - v_2$ be the ray generators of the fan. We fix a basis (v_1, v_2) of N. The image of the torus $\mathbf{T}$ is equal to the set of all divisors of the form $[a]D_{v_1} + [b]D_{v_2} + [-a-b]D_{v_3}$ for $0 \leq a, b < 1$. We see that the image of the torus $\mathbf{T}$ is $\mathcal{O}, \mathcal{O}(-1), \mathcal{O}(-2)$. This is a full collection. Notice however that it is not true that if we have a line bundle L then there exists an integer m_0 such that the push forward of L by the m-th Frobenius morphism for $m > m_0$ is a direct sum of line bundles from B. For example the push forward of $\mathcal{O}(-3)$ always contains in the split $\mathcal{O}(-3)$ that is not an element of B. However, as we will see only minor differences from the set B are possible.

Definition 19.11. *Let us fix a natural bijection between points of* $\boldsymbol{T}$ *and elements of* $M_{\mathbb{R}}$ *with entries from* $[0, 1)$ *in some fixed basis. Now each element of* B *has got a natural representative in* Div_T *as sum of* D_{g_j} *with integer coefficients. Let* $B_0 \subset Div_T$ *be the set of these representatives. We define the set* B' *as the set of all divisors* D *in* Pic *for which there exists an element in $b \in B_0$, such that there exists a representation of D whose coefficients differ by at most one from the coefficients of b.*

In other words we take (some fixed) representations of all elements of B, we take all other representations whose coefficients differ by at most one and we take the image in Pic to obtain B'.

Let us look once more at the example of $\mathbb{P}^2$. With previous notation B is equal to 0, $-D_{v_3}$, $-2D_{v_3}$. The set B' would be equal to $\pm D_{v_1} \pm D_{v_2} \pm D_{v_3}$, $\pm D_{v_1} \pm D_{v_2} \pm D_{v_3} - D_{v_3}$, $\pm D_{v_1} \pm D_{v_2} \pm D_{v_3} - 2D_{v_3}$. This gives us $\mathcal{O}(3)$,$\mathcal{O}(2)$,$\mathcal{O}(1)$, $\mathcal{O}$, $\mathcal{O}(-1)$, $\mathcal{O}(-2)$, $\mathcal{O}(-3)$, $\mathcal{O}(-4)$, $\mathcal{O}(-5)$.

Proposition 19.12. *For any smooth toric variety and any line bundle there exists an integer m_0 such that the push forward by the m-th Frobenius morphism for any $m > m_0$ splits into the line bundles from B'.*

Proof. From 19.9 we know that the line bundles from the split are of the form $[\frac{D}{m} + \frac{D_\chi}{m}]$, where $L = \mathcal{O}(D)$ is a fixed representation of L. Of course for sufficiently large m all coefficients of $\frac{D}{m}$ belong to the interval $(-1, 1)$. Hence the coefficients of $[\frac{D}{m} + \frac{D_\chi}{m}]$ differ by at most one from the coefficients of $[\frac{D_\chi}{m}]$ that is in B. This shows that $[\frac{D}{m} + \frac{D_\chi}{m}] \in B'$. □

This combined with the result of Thomsen [Tho00] that the push forward and the line bundle are isomorphic as sheaves of abelian groups gives us the following result:

Corollary 19.13. *There exists a finite set, namely B', such that each line bundle is isomorphic as a sheaf of abelian groups to a direct sum of line bundles from B'. In particular their cohomologies agree.* □

19.4. **Techniques of counting homology.** This section contains joint results with Michał Lasoń [LM11]. Our aim will be to describe line bundles on toric varieties with vanishing higher cohomologies, that we call acyclic. Later, we will use this characterization to check if $Ext^i(L, M) = H^i(L^\vee \otimes M)$ is equal to zero for $i > 0$. We start with general remarks on cohomology of line bundles on smooth, complete toric varieties.

Let Σ be a fan in $N = \mathbb{Z}^n$ with rays $x_1, ..., x_m$ and let $\mathbb{P}_\Sigma$ denote the variety constructed from the fan Σ. For $I \subset \{1, \dots, m\}$ let C_I be a simplicial complex generated by sets $J \subset I$ such that $\{x_i : i \in J\}$ generate a cone in Σ. For $r = (r_i : i = 1, \dots, m)$ let us define $Supp(r) := C_{\{i:\, r_i \geq 0\}}$.

The proof of the following well known fact can be found in the paper [BH09]:

Proposition 19.14. *The cohomology $H^j(\mathbb{P}_\Sigma, L)$ is isomorphic to the direct sum over all $r = (r_i : i = 1, \dots, m)$ such that $\mathcal{O}(\sum_{i=1}^m r_i D_{x_i}) \cong L$ of the $(n-j)$-th reduced homology of the simplicial complex $Supp(r)$.*

Definition 19.15. *We call a line bundle L on $\mathbb{P}_\Sigma$ acyclic if $H^i(\mathbb{P}_\Sigma, L) = 0$ for all $i \geq 1$.*

Definition 19.16. *For a fixed fan Σ we call a proper subset I of $\{1, \dots, m\}$ a forbidden set if the simplicial complex C_I has nontrivial reduced homology.*

From Proposition 19.14 we have the following characterization of acyclic line bundles.

Proposition 19.17. *A line bundle L on $\mathbb{P}_\Sigma$ is acyclic if it is not isomorphic to any of the following line bundles*

$$\mathcal{O}(\sum_{i \in I} r_i D_{x_i} - \sum_{i \notin I} (1 + r_i) D_{x_i})$$

where $r_i \geq 0$ and I is a proper forbidden subset of $\{1, \dots, m\}$.

Hence to determine which bundles on $\mathbb{P}_\Sigma$ are acyclic it is enough to know which sets I are forbidden.

In our case $C_I = \{J \subset I : \widehat{Y}_i := \{j : x_j \in Y_i\} \nsubseteq J \text{ for } i = 1, \dots, 5\}$, since Y_i are primitive collections. We call sets $\widehat{Y}_i$ also primitive collections. The only difference between sets $\widehat{Y}_i$ and Y_i is that the first one is the set of indices of rays in the second one, so in fact they could be even identified.

In case of a simplicial complex S on the set of vertices V we also define a primitive collection as a minimal subset of vertices that do not form a simplex. Complex S is determined by its primitive collections, namely it contains simplexes (subsets of V) that contain none of the primitive collections.

We describe a very powerful method of counting homologies of simplicial complexes which are given by their primitive collections (as in our case). To a simplicial complex S one can associate a complex C of vector spaces with the border map defined in the usual way. The objects in the complex C are indexed by nonnegative integers. Each object indexed by i is a direct sum of one dimension vector spaces, each corresponding to an i dimensional simplex in S. We assume that in C there is a one dimensional vector space indexed by 0 that corresponds to the empty set. Of course one can count cohomologies of any complex C of vector spaces, not necessarily coming from a simplicial complex. We transform the complex C so that the homologies remain unchanged. The method is due to Mrozek and Batko [MB09]. We will be removing some simplices from S. In particular, after some steps it will be no longer true that all faces of a simplex from a complex are in the complex. In this case the border map takes its values only in the simplices that are in the complex. This is a special example of the so-called S-complexes - for details see [MB09].

Example 19.18. Suppose that a one dimensional simplex P_1P_2 is in the complex S. The usual border map would be

$$\partial(P_1P_2) = P_1 - P_2.$$

However if we suppose that P_2 does not belong to S then

$$\partial(P_1P_2) = P_1.$$

Definition 19.19 (Reductive pair)**.** *Suppose that in a complex C the there exist simplices Z and B such that either*

$$\partial Z = B \text{ or } \partial Z = -B.$$

Then we call the pair (Z, B) a reductive pair*.*

We use the result of Mrozek and Batko [MB09]:

Lemma 19.20. *A reductive pair can be removed from a chain complex without changing the homology.*

Example 19.21. Consider a simplicial complex consisting of

$$\{\emptyset, P_1, P_2, P_3, P_1P_2, P_1P_3, P_2P_3\}.$$

(i) We remove the reductive pair $(P_1, \emptyset)$.
(ii) We remove the reductive pair (P_1P_2, P_2).
(iii) We remove the reductive pair (P_1P_3, P_3).

We are left with one simplex P_2P_3 and all border maps equal to zero.

For more information we advise the reader to consult [MB09, Section 6].

Definition 19.22. *Let X be a simplicial complex defined by its set of primitive collections P on the set of vertices V. We say that simplicial complex X' on the set of vertices $V \setminus P$ is obtained from X by deleting a primitive collection P if the set of primitive collections of X' is equal to the family of sets in $\{Q \cap (X \setminus P) : Q \in \mathrm{P}\}$ that are minimal with respect to inclusion.*

Lemma 19.23. *Let X be a simplicial complex and suppose that there exists an element x which belongs to exactly one primitive collection*

P. Let $m = |P|$ and let X' be a simplicial complex obtained from X by delating P, then

$$h^i(X) = h^{i-m+1}(X').$$

Proof. Using Lemma 19.20 we will be removing inductively on dimension reductive pairs (Z, B) such that $x \in Z$. We start from $(\{x\}, \emptyset)$. One can see that in each dimension we can take all $(Z, Z \setminus \{x\})$ for Z containing x as reductive pairs. Let us consider all simplexes of X that do not contain $P \setminus \{x\}$. One can prove by induction on dimension that we will remove all of them:

Let D be a simplex. If it contains x, than it will be removed as a first element of a reductive pair. If it does not, then $D \cup \{x\}$ is also a simplex of X and we will remove $(D \cup \{x\}, D)$.

We see that our simplicial complex can be reduced to a complex with simplexes containing $P \setminus \{x\}$. Now one immediately sees that such a complex is isomorphic to a complex X' (with a degree shifted by $|P \setminus \{x\}| = m - 1$). □

The same method allows us to easily compute homologies when there are few primitive collections and many points. The idea is that we can glue together points that are in exactly the same primitive collections.

Definition 19.24. *Let X be a simplicial complex defined by its set of primitive collections P on the set of vertices V. Suppose that there exist two points $x, y \in X$ such that they belong to the same primitive collections. We say that a simplicial complex X' on the set of vertices $V \setminus \{y\}$ is obtained from X by gluing points x and y if the set of primitive collections of X' is equal to $\{Q \setminus \{y\} : Q \in \mathrm{P}\}$. We can think of it like x was in fact two points x, y.*

Proposition 19.25. *Let X be a simplicial complex and suppose that there exist two points $x, y \in X$ such that they belong to the same primitive collections. Let X' be a simplicial complex obtained from X by gluing points x and y, then*

$$h^i(X) = h^{i-1}(X').$$

Proof. In both complexes we will be removing reductive pairs of the form (Z, B) with $x \in Z$ just as in Lemma 19.23. In both situations all that is left are simplexes that contain a set of a form $P \setminus \{x\}$, where P is a primitive collection containing x. In this situation all of simplexes of X that are left contain y and they can be identified with simplexes of X' that are left, the maps are exactly the same what finishes the proof. □

Corollary 19.26. *Let X be a simplicial complex on the set of vertices V. Let X' be a simplicial complex obtained from X by gluing equivalence classes of the relation $\sim$ that identifies elements that are in exactly the same primitive collections. Suppose $|V| - |V/\sim| = m$, then*

$$h^i(X) = h^{i-m}(X').$$

Proof. We use 19.25 for pairs of points in the equivalence classes. □

Corollary 19.27. *In the situation of Lemma 19.23 and Corollary 19.26 X is acyclic if and only if X' is acyclic.*

With these tools we are ready to determine forbidden subsets. In general we have got two following Lemmas:

Lemma 19.28. *If a nonempty subset I is not a sum of primitive collections, then it is not forbidden.*

Proof. There exists $a \in I$ such that a does not belong to any primitive collection which is contained in I. Using Lemma 19.20 we can remove subsequently on dimension reductive pairs (Z, B) such that $a \in Z$. We start from $(\{a\}, \emptyset)$. One can see that in this way we remove all of simplexes and as a consequence the chain complex is exact. □

Lemma 19.29. *A primitive collection is a forbidden subset.*

Proof. Using Lemma 19.23 we can remove this primitive collection and get a complex consisting of the empty set only that has nontrivial reduced homologies.

This can be also seen from the fact that the considered complex topologically is a sphere. □

The following Lemmas apply to the case when the Picard number is three and we have five primitive collections as in Batyrev's classification. Let us remind that primitive collections of simplicial complex in this case are $\widehat{Y_i} := \{j : x_j \in Y_i\}$, for our convenience we define also $\widehat{X_i} := \{j : x_j \in X_i\}$.

Lemma 19.30. *A sum of two consecutive primitive collections is a forbidden subset.*

Proof. Using Lemma 19.23 we remove one primitive collection and get a situation of Lemma 19.29. □

Lemma 19.31. *A sum of three consecutive primitive collections $\widehat{Y_i}$, $\widehat{Y_{i+1}}$, $\widehat{Y_{i+2}}$ is not a forbidden subset.*

Proof. First we can remove primitive collection $\widehat{Y_i}$. The image of $\widehat{Y_{i+2}}$ contains the image of $\widehat{Y_{i+1}}$, so in fact we are left with just one primitive collection P which is an image of $\widehat{Y_{i+1}}$. We can remove P and obtain a nonempty full simplicial complex which is known to have trivial homologies. □

The above lemmas match together to the following:

Theorem 19.32. *The only forbidden subsets are primitive collections, their complements and the empty set.*

This gives us that in our situation

Corollary 19.33. *A line bundle L is acyclic if and only if it is not isomorphic to any of the line bundles*

$$\mathcal{O}(\alpha_1^1 D_{v_1} + \cdots + \alpha_2^1 D_{y_1} + \cdots + \alpha_3^1 D_{z_1} + \cdots + \alpha_4^1 D_{t_1} + \cdots + \alpha_5^1 D_{u_1} + \cdots)$$

where exactly $2, 3$ or 5 consecutive $\alpha_i := (\alpha_i^1, \cdots, \alpha_i^{p_i})$ are negative.

Proof. It is an immediate consequence of Proposition 19.17 and Theorem 19.32 □

Corollary 19.34. *If all of the coefficients b and c are zero in the primitive relations from Theorem 19.7 then a line bundle L is acyclic if and only if it is not isomorphic to any of the line bundles*

$$\mathcal{O}(\alpha_1 D_v + \alpha_2 D_y + \alpha_3 D_z + \alpha_4 D_t + \alpha_5 D_u)$$

where exactly $2, 3$ or 5 consecutive α_i are negative and if $\alpha_i < 0$ then $\alpha_i \leq -|X_i|$.

Proof. Since all divisors corresponding to elements of the set X_i are linearly equivalent we match them together and as a consequence α_i is the sum of all of their coefficients. □

19.5. **Large family of smooth toric varieties with Picard number 3.** This section contains joint results with Michał Lasoń [LM11]. We give an explicit construction of a full, strongly exceptional collection of line bundles in the derived category $D^b(X)$ for a large family of smooth, complete toric varieties X with Picard number three. Namely for varieties X whose sets X_1, X_3 and X_4 from Batyrev's classification presented in Theorem 19.7 have only one element. We will use results from Section 19.4.

19.5.1. *Our setting.* In this subsection we establish a family of varieties which we consider in this section and we also fix notation.

From now on for the whole Section let X be a smooth, complete toric variety with Picard number three, which using the notation from Theorem 19.7 has $|X_1| = |X_3| = |X_4| = 1$.

Let $r = |X_2|$. Then of course $|X_0| = n - r$. We allow arbitrary non-negative integer parameters $b := b_1,\ c_2, \ldots, c_r$. This family generalizes one considered in [DLM09] (there, the case $r = 1$ was considered) and [CMRa] (there the case $b = c_1 = \cdots = c_r = 0$ was considered).

Remark 19.35. *A variety of this type is Fano if and only if*

$$n - r > \sum_{i=2}^{r} c_r + b.$$

In what follows we do not restrict to the Fano case.

Let $e_1, \ldots, e_n$ be a basis of the lattice N. Let us write what are the coordinates of the ray generators in the considered situation:

$$\begin{aligned}
&v_1 = e_1, v_2 = e_2, \ldots, v_{n-r} = e_{n-r}\\
&y = -e_1 - \cdots - e_{n-r} + c_2 e_{n-r+2} + \cdots + c_r e_n - (b+1)(e_{n-r+1} + \cdots + e_n)\\
&z_1 = e_{n-r+1}, \ldots, z_r = e_n\\
&t = -e_{n-r+1} - \cdots - e_n\\
&u = -e_1 - \cdots - e_{n-r} + c_2 e_{n-r+2} + \cdots + c_r e_n - b(e_{n-r+1} + \cdots + e_n)
\end{aligned} \tag{19.2}$$

Let D_w be the divisor associated to the ray generator w. One can easily see that the divisors $D_{v_1}, \ldots, D_{v_{n-r}}$ are all linearly equivalent. Let D_v be any their representant in the Picard group. The other equivalence relations that generate all the relations in the Picard group are:

$$\begin{aligned}
D_v &\simeq D_u + D_y\\
D_{z_1} &\simeq D_t + bD_u + (b+1)D_y\\
D_{z_i} &\simeq D_t + (b - c_i)D_u + (b - c_i + 1)D_y \quad 2 \le i \le r
\end{aligned} \tag{19.3}$$

From these relations we can easily deduce:

Proposition 19.36. *The Picard group of the variety X is isomorphic to $\mathbb{Z}^3$ and is generated by D_t, D_y, D_v.*

We introduce two sets of divisors. We claim that these sets can be ordered in such a way that line bundles corresponding to divisors from these sets form a strongly exceptional collection.

$$\begin{aligned} Col_1 = \{ -sD_t - sD_y + (-(n-r) - bs + q)D_v : \\ 0 \le s \le r, 0 \le q \le n-r\} \\ Col_2 = \{ -sD_t - (s-1)D_y + (-(n-r) - bs + q)D_v : \\ 1 \le s \le r, 0 \le q \le n-r-1\} \end{aligned} \tag{19.4}$$

Definition 19.37. *Let* $Col = Col_1 \cup Col_2$.

Remark 19.38. *Let us notice that* $|Col_1| = (r+1)(n-r+1)$ *and* $|Col_2| = r(n-r)$, *so* $|Col| = 2rn - 2r^2 + n + 1$.

We calculate the number of maximal cones in the fan defining the variety X. In order to obtain a maximal cone we have to choose n ray generators that do not contain a primitive collection. This is equivalent to removing three ray generators in such a way that the rest do not contain a primitive collection. First let us notice that we can remove at most one element from each group X_i because otherwise the rest would contain a primitive collection. We have the following possibilities:

1) We remove one element from X_0 and X_2. Then we have to remove one element from X_3 or X_4. We have got $2(n-r)r$ such possibilities.

2) We remove one element from X_0 and none from X_2. We have got $n-r$ such possibilities.

3) We remove one element from X_2 and none from X_0. We have got r such possibilities.

4) We do not remove any elements from X_0 and from X_2. We have got 1 such possibility.

All together we see that we have $2rn - 2r^2 + n + 1$ maximal cones. From the general theory we know that the rank of the Grothendieck group is the same. Let us notice that from Remark 19.38 our set Col is of the same number of elements.

19.5.2. *Acyclicity of differences of line bundles from* Col. In this Subsection we order the set Col and prove that line bundles corresponding to divisors from Col form a strongly exceptional collection.

Let us first check that $\mathrm{Ext}^i_{\mathcal{O}_X}(\mathcal{O}(D_1), \mathcal{O}(D_2)) = 0$ for any divisors D_1, D_2 from the set Col and for any $i > 0$. We know that

$$\mathrm{Ext}^i_{\mathcal{O}_X}(\mathcal{O}(D_1), \mathcal{O}(D_2)) = H^i(\mathcal{O}(D_1)^\vee \otimes \mathcal{O}(D_2)) = H^i(\mathcal{O}(D_2 - D_1)).$$

This means that we have to show that all line bundles associated to differences of divisors from Col are acyclic.

Definition 19.39. *Let $Diff$ be the set of all divisors of the form $D_1 - D_2$, where $D_1, D_2 \in Col$.*

Proposition 19.40. *The set $Diff$ is the union of the sets $Diff_1$, $Diff_2$, $Diff_3$, where:*

$$Diff_1 = \{sD_t + sD_y + (bs+q)D_v :$$

$$-r \leq s \leq r, r-n \leq q \leq n-r\}$$

$$Diff_2 = \{sD_t + (s-1)D_y + (bs+q)D_v :$$

$$-r+1 \leq s \leq r, r-n+1 \leq q \leq n-r\}$$

$$Diff_3 = \{sD_t + (s+1)D_y + (bs+q)D_v :$$

$$-r \leq s \leq r-1, r-n \leq q \leq n-r-1\}.$$

Proof. The set $Diff_1$ is equal to the set of all possible differences of two divisors from Col_1 and this set contains all possible differences of two divisors from Col_2. The set $Diff_2$ is the set of all possible differences of the form $D_1 - D_2$, where $D_1 \in Col_1, D_2 \in Col_2$. The set $Diff_3$ is equal to $-Diff_2$ and so it is equal to the set of all differences of the form $D_2 - D_1$, where $D_1 \in Col_1, D_2 \in Col_2$. These are of course all possible differences of two elements from Col. □

From the Corollary 19.33 we know that it is enough to prove that elements of $Diff$ are not of the form

$$\alpha_1 D_v + \alpha_2 D_y + \alpha_3^1 D_{z_1} + \alpha_3^2 D_{z_2} + \cdots + \alpha_3^r D_{z_r} + \alpha_4 D_t + \alpha_5 D_u,$$

where exactly two, three or five consecutive α_i's are negative (we call a number positive when it is nonnegative and consider only two signs positive and negative) and:

1) if $\alpha_1 < 0$, then $\alpha_1 \leq -(n-r)$ (α_1 is in fact sum of all the coefficients of D_{v_i}, which have to be of the same sign),

2) if any $\alpha_3^i < 0$ then $\alpha_3^j < 0$ (all parameters α_3^j are treated as one group and have the same sign).

From now on we assume that these conditions on α_i's are satisfied.

Using the relations 19.3 we obtain:

$$\alpha_1 D_v + \alpha_2 D_y + \alpha_3^1 D_{z_1} + \alpha_3^2 D_{z_2} + \cdots + \alpha_3^r D_{z_r} + \alpha_4 D_t + \alpha_5 D_u =$$

$$(\alpha_4 + \sum_{j=1}^{r} \alpha_3^j) D_t + (\alpha_2 - \alpha_5 + \sum_{j=1}^{r} \alpha_3^j) D_y +$$

$$(\alpha_1 + b\alpha_3^1 + \sum_{j=2}^{r} (b - c_j)\alpha_3^j + \alpha_5) D_v \tag{19.5}$$

Lemma 19.41. *If the elements α_3^j are negative then the divisors from $Diff$ are not of the form 19.5.*

Proof. If α_4 was negative, then the coefficient of D_t would be less than or equal to $-r-1$ and none of the divisors from $Diff$ has got such a coefficient, so α_4 has to be positive. Since α_3 is negative and α_4 is positive, then α_2 has to be negative and α_5 has to be positive. This means that the coefficient of D_y is less then or equal to $-r-1$. The divisors from $Diff$ are not of this form. □

From now on we may assume that α_3 is positive.

Lemma 19.42. *The divisors from $Diff_1$ are not of the form (19.5).*

Proof. Suppose that a divisor from $Diff_1$ can be written in a form (19.5). We have:

$$\alpha_4 + \sum_{j=1}^{r} \alpha_3^j = \alpha_2 - \alpha_5 + \sum_{j=1}^{r} \alpha_3^j,$$

so $\alpha_4 + \alpha_5 = \alpha_2$. But α_2, α_4 and α_5 cannot be of the same sign, so α_4 and α_5 have to have different signs. As α_3 was positive we see that α_4 is positive, so α_5 and α_1 are negative. Let us notice that:

$$\alpha_1 + b\alpha_3^1 + (\sum_{j=2}^{r} (b - c_j)\alpha_3^j) + \alpha_5 \leq$$

$$-n + r + b(\sum_{j=1}^{r} \alpha_3^j) - 1 \leq$$

$$-n + r - 1 + b(\alpha_4 + \sum_{j=1}^{r} \alpha_3^j)$$

This shows precisely that the coefficient of D_v is less than or equal to $-n+r-1$ plus b times the coefficient of D_t. Let s be the coefficient of D_t. From the definition of $Diff_1$ the coefficient of D_v is at least $-n+r+bs$. This gives us a contradiction. □

Lemma 19.43. *The divisors from $Diff_3$ are not of the form (19.5).*

Proof. Suppose that a divisor from $Diff_3$ can be written in a form (19.5). We have:

$$\alpha_4 + \sum_{j=1}^{r} \alpha_3^j = \alpha_2 - \alpha_5 - 1 + \sum_{j=1}^{r} \alpha_3^j,$$

so $\alpha_4 + \alpha_5 = \alpha_2 - 1$. The rest of the proof is identical to the proof of Lemma 19.66. □

Lemma 19.44. *The divisors from $Diff_2$ are not of the form (19.5).*

Proof. Suppose that a divisor from $Diff_2$ can be written in a form (19.5). We have:

$$\alpha_4 + \sum_{j=1}^{r} \alpha_3^j = \alpha_2 - \alpha_5 + 1 + \sum_{j=1}^{r} \alpha_3^j,$$

so $\alpha_4 + \alpha_5 = \alpha_2 + 1$. But α_2, α_4 and α_5 cannot be of the same sign, so we have two possible cases:

1) The coefficients α_4 and α_5 have different signs. In this case the proof is the same as in Lemmas 19.66 and 19.43.

2) We have $\alpha_4 = \alpha_5 = 0$ and $\alpha_2 = -1$. In this case α_1 has to be negative, because α_3 was positive. Let $s = \alpha_4 + \sum_{j=1}^{r} \alpha_3^j$ be the coefficient of D_t. We have:

$$\alpha_1 + b\alpha_3^1 + \sum_{j=2}^{r} (b - c_j)\alpha_3^j + \alpha_5 \leq -n + r + bs,$$

so the coefficient of D_v is less than or equal to $-n+r+bs$. But from the definition of $Diff_2$ we know that the coefficient of D_v is at least $bs+r-n+1$ what gives us a contradiction. □

Now we only have to order the line bundles corresponding to divisors from Col in such a way that

$$0 = \operatorname{Ext}^0_{\mathcal{O}_X}(\mathcal{O}(D_1), \mathcal{O}(D_2)) = H^0(\mathcal{O}(D_1)^\vee \otimes \mathcal{O}(D_2)) = H^0(\mathcal{O}(D_2 - D_1)).$$

for any divisors $D_1 > D_2$.

Let us define the order by: $L_{s,q} < L'_{s,q} < L_{s,q+1}$, $L_{s+1,q_1} < L_{s,q_2}$ where

$$L_{s,q} = \mathcal{O}(-sD_t - sD_y + (q - bs - (n-r))D_v)$$

for $s = 0, \dots, r$ and $q = 0, \dots, n-r$ and

$$L'_{s,q} = \mathcal{O}(-sD_t - (s-1)D_y + (q - bs - (n-r))D_v)$$

for $s = 1, \dots, r-1$ and $q = 0, \dots, n-r-1$. It is easy to see that zero cohomology of appropriate differences vanish.

19.5.3. *Generating the derived category.* We prove that the strongly exceptional collection from Subsection 19.5.1 is also full. First we show that it generates all line bundles. Due to [BH06, Corollary 4.8] the collection generates the derived category. In order to generate all line bundles we need several lemmas. Our first aim is to generate line bundles of the type $sD_tsD_y + qD_v$ and $sD_t + (s+1)D_y + qD_v$. We first do it for fixed s and any q – the result is in Lemma 19.49. The idea is to generate the line bundles inductively on q. We will be doing this using the Koszul complexes for families of divisors for different primitive collections. As the ray generators corresponding to divisors of a primitive collection do not form a cone, we obtain indeed the exact sequences given by Koszul complexes.

Lemma 19.45. *Let s and k be any integers. Line bundles $L_q = \mathcal{O}(-sD_t - sD_y + (k+q)D_v)$ for $q = 0, \dots, n-r$ and $L'_q = \mathcal{O}(-sD_t - (s-1)D_y + (k+q)D_v)$ for $q = 0, \dots, n-r-1$ generate $\mathcal{O}(-sD_t - (s-1)D_y + (n-r+k)D_v)$ in the derived category.*

Proof. We consider the Koszul complex for $\mathcal{O}(D_y), \mathcal{O}(D_{v_1}), \dots, \mathcal{O}(D_{v_{n-r}})$:

$$0 \to \mathcal{O}(-D_y - (n-r)D_v) \to \cdots \to \mathcal{O}(-D_v)^{n-r} \oplus \mathcal{O}(-D_y) \to \mathcal{O} \to 0.$$

By tensoring it with $\mathcal{O}(-sD_t - (s-1)D_y + (k+n-r)D_v)$ we obtain:

$$0 \to \mathcal{O}(-sD_t - sD_y + kD_v) \to \cdots \to \mathcal{O}(-sD_t - (s-1)D_y + (k+n-r-1)D_v)^{n-1}$$

$$\oplus \mathcal{O}(-sD_t - sD_y + (k+n-r)D_v) \to \mathcal{O}(-sD_t - (s-1)D_y) + (k+n-r)D_v) \to 0.$$

All sheaves that appear in this exact sequence, apart from the last one, are exactly $\mathcal{O}(-sD_t - sD_y + kD_v), \dots, \mathcal{O}(-sD_t - sD_y + (k+n-r)D_v), \mathcal{O}(-sD_t - (s-1)D_y + kD_v), \dots, \mathcal{O}(-sD_t - (s-1)D_y + (k+n-r-1)D_v)$, so indeed we can generate $\mathcal{O}(-sD_t - (s-1)D_y + (k+n-r)D_v)$. □

Lemma 19.46. *Let s and k be any integers. Line bundles $L_q = \mathcal{O}(-sD_t - sD_y + (k+q)D_v)$ for $q = 0, \dots, n-r$ and $L'_q = \mathcal{O}(-sD_t - (s-1)D_y + (k+q)D_v)$ for $q = 1, \dots, n-r$ generate $\mathcal{O}(-sD_t - (s-1)D_y + kD_v)$ in the derived category.*

Proof. The proof is similar to the last one. We deduce assertion from the same exact sequence of sheaves. □

Lemma 19.47. *Let s and k be any integers. Line bundles $L_q = \mathcal{O}(-sD_t - sD_y + (k+q)D_v)$ for $q = 1, \dots, n-r$ and $L'_q = \mathcal{O}(-sD_t - (s-1)D_y + (k+q)D_v)$ for $q = 0, \dots, n-r$ generate $\mathcal{O}(-sD_t - sD_y + (n-r+k+1)D_v)$ in the derived category.*

Proof. The proof is similar to the first one. We have to consider the Koszul complex for line bundles $\mathcal{O}(D_u), \mathcal{O}(D_{v_1}), \dots, \mathcal{O}(D_{v_{n-r}})$:

$$0 \to \mathcal{O}(-D_u - (n-r)D_v) \to \dots \to \mathcal{O}(-D_v)^{n-r} \oplus \mathcal{O}(-D_u) \to \mathcal{O} \to 0$$

we dualize it and we tensor it with $\mathcal{O}(-sD_t - (s-1)D_y + kD_v)$. □

Lemma 19.48. *Let s and k be any integers. Line bundles $L_q = \mathcal{O}(-sD_t - sD_y + (k+q)D_v)$ for $q = 1, \dots, n-r+1$ and $L'_q = \mathcal{O}(-sD_t - (s-1)D_y + (k+q)D_v)$ for $q = 1, \dots, n-r$ generate $\mathcal{O}(-sD_t - sD_y + kD_v)$ in the derived category.*

Proof. The proof is similar to the last one. We deduce assertion from the same exact sequence of sheaves. □

Lemma 19.49. *Let s and k be any integers. Line bundles $L_q = \mathcal{O}(-sD_t - sD_y + (k+q)D_v)$ for $q = 0, \dots, n-r$ and $L'_q = \mathcal{O}(-sD_t - (s-1)D_y + (k+q)D_v)$ for $q = 0, \dots, n-r-1$ generate in the derived category line bundles $\mathcal{O}(-sD_t - sD_y + q'D_v)$ and $\mathcal{O}(-sD_t - (s-1)D_y + q'D_v)$ for an arbitrary integer q'.*

Proof. We prove it by induction on $|q'|$. For $q' \geq k + n - r$ we use Lemmas 19.45 and 19.47, for $q' < k$ we use Lemmas 19.46 and 19.48. □

Next we generate all line bundles of the type $sD_t + sD_y + qD_v$ and $sD_t + (s+1)D_y + qD_v$ with no restrictions on s and q. The ideas are the same and the result is in Lemma 19.54.

Lemma 19.50. *Let k be any integer. Line bundles $L_{s,q} = \mathcal{O}(-sD_t - sD_y + qD_v)$ for $s = k, \dots, k+r$ and arbitrary q and $L'_{s,q} = \mathcal{O}(-sD_t - (s-1)D_y + qD_v)$ for $s = k, \dots, k+r-1$ and arbitrary q generate in the derived category line bundles $L'(k+r, q) = \mathcal{O}(-(k+r)D_t - (k+r-1)D_y + qD_v)$ with arbitrary q.*

Proof. Consider the Koszul complex for $\mathcal{O}(D_y), \mathcal{O}(D_{z_1}), \dots, \mathcal{O}(D_{z_r})$:

$$0 \to \mathcal{O}(-D_{z_1} - (r-1)D_{z_2} - D_y) \to \dots$$
$$\dots \to \mathcal{O}(-D_{z_1}) \oplus \mathcal{O}(-D_{z_2})^{r-1} \oplus \mathcal{O}(-D_y) \to \mathcal{O} \to 0.$$

After tensoring it with $\mathcal{O}(-(k-1)D_y + q'D_v)$ for appropriate q' we get the assertion. □

Lemma 19.51. *Let k be any integer. Line bundles $L_{s,q} = \mathcal{O}(-sD_t - sD_y + qD_v)$ for $s = k, \dots, k+r$ and arbitrary q and $L'_{s,q} = \mathcal{O}(-sD_t - (s-1)D_y + qD_v)$ for $s = k+1, \dots, k+r$ and arbitrary q generate in the derived category line bundles $L'(k,q) = \mathcal{O}(-kD_t - (k-1)D_y + qD_v)$ for arbitrary q.*

Proof. The proof is similar to the last one. We deduce assertion from the same exact sequence of sheaves. □

Lemma 19.52. *Let k be any integer. Line bundles $L_{s,q} = \mathcal{O}(-sD_t - sD_y + qD_v)$ for $s = k+1, \dots, k+r$ and arbitrary q and $L'_{s,q} = \mathcal{O}(-sD_t - (s-1)D_y + qD_v)$ for $s = k+1, \dots, k+r+1$ and arbitrary q generate in the derived category line bundles $L(k,q) = \mathcal{O}(-kD_t - kD_y + qD_v)$ for arbitrary q.*

Proof. Consider the Koszul complex for $\mathcal{O}(D_{z_1}), \dots, \mathcal{O}(D_{z_r}), \mathcal{O}(D_t)$:

$$0 \to \mathcal{O}(-D_{z_1} - (r-1)D_{z_2} - D_t) \to \dots$$
$$\dots \to \mathcal{O}(-D_{z_1}) \oplus \mathcal{O}(-D_{z_2})^{r-1} \oplus \mathcal{O}(-D_t) \to \mathcal{O} \to 0.$$

After tensoring it with $\mathcal{O}(-kD_y + q'D_v)$ for appropriate q' we get the assertion. □

Lemma 19.53. *Let k be any integer. Line bundles $L_{s,q} = \mathcal{O}(-sD_t - sD_y + qD_v)$ for $s = k, \dots, k+r$ and arbitrary q and $L'_{s,q} = \mathcal{O}(-sD_t - (s-1)D_y + qD_v)$ for $s = k+1, \dots, k+r$ and arbitrary q generate in the derived category line bundles $L'(k+r+1, q) = \mathcal{O}(-(k+r+1)D_t - (k+r)D_y + qD_v)$ for arbitrary q.*

Proof. The proof is similar to the last one. We deduce assertion from the same exact sequence of sheaves. □

Lemma 19.54. *Let k be any integer. Line bundles $L_{s,q} = \mathcal{O}(-sD_t - sD_y + qD_v)$ for $s = k, \dots, k+r$ and arbitrary q and $L'_{s,q} = \mathcal{O}(-sD_t - (s-1)D_y + qD_v)$ for $s = k, \dots, k+r-1$ and arbitrary q generate in the derived category line bundles $L(s,q) = \mathcal{O}(-sD_t - sD_y + qD_v)$ and $L'(s,q) = \mathcal{O}(-sD_t - (s-1)D_y + qD_v)$ for arbitrary s and q.*

Proof. We prove it by induction on $|s|$. For $s \geq k+n-r$ we use Lemmas 19.50 and 19.53, for $r < k$ we use Lemmas 19.51 and 19.64. □

Finally we proceed inductively on the difference of the coefficients of D_t and D_y.

Lemma 19.55. *Let k be any integer. Line bundles $\mathcal{O}(-sD_t - (s+k)D_y + qD_v)$ and $\mathcal{O}(-sD_t - (s+k+1)D_y + qD_v)$ for arbitrary s and q generate in the derived category line bundles $\mathcal{O}(-sD_t - (s+k+2)D_y + qD_v)$ for arbitrary s and q.*

Proof. Consider the Koszul complex for $\mathcal{O}(D_t), \mathcal{O}(D_u)$:

$$0 \to \mathcal{O}(-D_t - D_u) \to \mathcal{O}(-D_t) \oplus \mathcal{O}(-D_u) \to \mathcal{O} \to 0.$$

After tensoring it with $\mathcal{O}(-k'D_y + q')$ for appropriate k' and q' we get the assertion. □

Lemma 19.56. *Let k be any integer. Line bundles $\mathcal{O}(-sD_t - (s+k)D_y + qD_v)$ and $\mathcal{O}(-sD_t - (s+k+1)D_y + qD_v)$ for arbitrary s and q generate in the derived category line bundles $\mathcal{O}(-sD_t - (s+k-1)D_y + qD_v)$ for arbitrary s and q.*

Proof. Consider the Koszul complex for $\mathcal{O}(D_t), \mathcal{O}(D_u)$:

$$0 \to \mathcal{O}(-D_t - D_u) \to \mathcal{O}(-D_t) \oplus \mathcal{O}(-D_u) \to \mathcal{O} \to 0.$$

After tensoring it with $\mathcal{O}(-k'D_y + q')$ for appropriate k' and q' we get the assertion. □

Proposition 19.57. *Line bundles*

$$L_{s,q} = \mathcal{O}(-sD_t - sD_y + (q - bs - (n-r))D_v)$$

for $s = 0, \dots, r$ and $q = 0, \dots, n-r$ and

$$L'_{s,q} = \mathcal{O}(-sD_t - (s-1)D_y + (q - bs - (n-r))D_v)$$

for $s = 0, \dots, r-1$ and $q = 0, \dots, n-r-1$ generate in the derived category all line bundles.

Proof. We use Lemmas 19.49, 19.54, 19.55 and 19.56. □

Summarizing, we have proved:

Theorem 19.58. *Let X be a smooth, complete, n dimensional toric variety with Picard number three and the set of ray generators $X_0 \cup \dots \cup X_4$, where*

$$X_0 = \{v_1, \dots, v_{n-r}\},\ X_1 = \{y\},\ X_2 = \{z_1, \dots, z_r\},\ X_3 = \{t\},\ X_4 = \{u\},$$

primitive collections $X_0 \cup X_1$, $X_1 \cup X_2, \dots, X_4 \cup X_0$ and primitive relations:

$$v_1 + \dots + v_{n-r} + y - cz_2 - \dots - cz_r - (b+1)t = 0,$$
$$y + z_1 + \dots + z_r - u = 0,$$
$$z_1 + \dots + z_r + t = 0,$$
$$t + u - y = 0,$$
$$u + v_1 + \dots + v_{n-r} - c_2z_2 - \dots - c_rz_r - bt = 0,$$

where b and c are positive integers.

Then the ordered collection of line bundles

$$L_{s,q} = \mathcal{O}(-sD_t - sD_y + (q - bs - (n-r))D_v)$$

for $s = 0, \dots, r$ and $q = 0, \dots, n-r$ and

$$L'_{s,q} = \mathcal{O}(-sD_t - (s-1)D_y + (q - bs - (n-r))D_v)$$

for $s = 0, \dots, r-1$ and $q = 0, \dots, n-r-1$ where the order is defined by $L_{s,q} < L'_{s,q} < L_{s,q+1}$, $L_{s+1,q_1} < L_{s,q_2}$ is a full, strongly exceptional collection of line bundles.

Proof. From Subsection 19.5.2 we already know that this is a strongly exceptional collection. We have just checked the sufficient condition for fullness in Proposition 19.57. □

19.6. **The split of the push forward of the structural sheaf not containing a full, strongly exceptional collection.** This section contains joint results with Michał Lasoń [LM11].

19.6.1. *Example.* Let us consider the case when:

$$X_0 = \{v_1\}, \quad X_1 = \{y_1, \dots, y_k\}, \quad X_2 = \{z_1\},$$
$$X_3 = \{t_1, \dots, t_k\}, \quad X_4 = \{u_1, \dots, u_k\}$$

then we can take

$$v_1, y_2, \dots, y_k, t_1, \dots, t_k, u_2, \dots, u_k$$

to be a basis of the lattice $N = \mathbb{Z}^{3k-1}$. Other vectors are like in 19.1 with all coefficients b_i and c_i equal to zero. We have linear dependencies of divisors:

$$D_{v_1} = D_{u_1} + D_{y_1}, \quad D_{t_i} = D_{z_1} + D_{y_1}, \quad D_{y_i} = D_{y_1}, \quad D_{u_i} = D_{u_1}$$

Let B be the image of the real torus in the Picard group as described in the Subsection 19.3. One can easily see that:

$$B = \{\mathcal{O}([\sum_{i=1}^{k} -\alpha_t^i]D_{z_1} + [\sum_{i=2}^{k} -\alpha_u^i - \alpha_v^1]D_{u_1} + [-\alpha_v^1 + \sum_{i=2}^{k} -\alpha_y^i + \sum_{i=1}^{k} \alpha_t^i]D_{y_1}) :$$
$$0 \le \alpha_v^i, \alpha_y^i, \alpha_t^i, \alpha_u^i < 1\}.$$

So B is contained in the set:

$$S := \{\mathcal{O}(-aD_{z_1} - bD_{u_1} + (a-c)D_{y_1}) : a, b, c \in \{0, \dots, k\}\} =$$
$$= \{\mathcal{O}(-a(D_{z_1} - D_{y_1}) - bD_{u_1} - cD_{y_1}) : a, b, c \in \{0, \dots, k\}\}.$$

From Corollary 19.34 we know that line bundle is acyclic if and only if it is not isomorphic to any of the following line bundles

$$\mathcal{O}(\alpha_1 D_{v_1} + \alpha_2 D_{y_1} + \alpha_3 D_{z_1} + \alpha_4 D_{t_1} + \alpha_5 D_{u_1}) =$$
$$= \mathcal{O}((\alpha_3 + \alpha_4)(D_{z_1} - D_{y_1}) + (\alpha_1 + \alpha_2 + \alpha_3)D_{y_1} + (\alpha_1 + \alpha_5)D_{u_1}),$$

where exactly $2, 3$ or 5 consecutive α are negative and if $\alpha_2 < 0$ then $\alpha_2 \leq -k$, if $\alpha_4 < 0$ then $\alpha_4 \leq -k$ and if $\alpha_5 < 0$ then $\alpha_5 \leq -k$. Let us observe that line bundles from the set

$$R = \{\mathcal{O}(a(D_{z_1}-D_{y_1})+bD_{y_1}+cD_{u_1}) : (a, b, c) \in [\frac{k}{2}, k]\times[-k, -\frac{k}{2}-1]\times[0, k]\}$$

are not acyclic. Indeed fixing $\alpha_1 = -k, \alpha_3 = \frac{k}{2}$ and taking α_4, α_5 nonnegative and α_2 negative we can achieve all of them. Let us define the set of pairs

$$P := \{-(\frac{k}{2}+\frac{a}{2})(D_{z_1}-D_{y_1})-(\frac{k}{2}+\frac{b}{2})D_{y_1}-(\frac{k}{2}+\frac{c}{2})D_{u_1}, -(\frac{k}{2}-\frac{a}{2})(D_{z_1}-D_{y_1})-$$

$$-(\frac{k}{2}-\frac{b}{2})D_{y_1}-(\frac{k}{2}-\frac{c}{2})D_{u_1}) : (a, b, c) \in [\frac{k}{2}, k] \times [-k, -\frac{k}{2}-1] \times [0, k]\}.$$

It is easy to see that elements of these pairs are distinct and they belong to S. Difference in each pair is an element of R so it is not acyclic line bundle. Hence to have a strongly exceptional collection C in S we have to exclude at least one element from each pair. To have integer coefficients of divisors in P we should take $a \equiv b \equiv c \equiv k \pmod 2$, so we have to throw out at least $\frac{k^3}{32}$ elements among $(k+1)^3$ elements in S. Full, strongly exceptional collection has to have l elements, where l is the rank of the Grothendieck group $K^0(X)$ (for toric varieties it is isomorphic to $\mathbb{Z}^l$, where l is the number of maximal cones). In our case there are at least k^3 maximal cones, since each time we throw out one element from X_2, X_4 and X_5 we get different maximal cone (exact number is $k^3 + 2k^2 + 2k$). So we have proven the following:

Theorem 19.59. *For $k > 32$ there is no full, strongly exceptional collection contained in the set of line bundles that come from Bondal's construction.*

Proof. For $k > 32$ we have $(k+1)^3 - \frac{1}{32}k^3 < k^3 + 2k^2 + 2k$ so the proof follows from the discussion above. □

Remark 19.60. *Notice that the considered variety is Fano, so is expected to have a full, strongly exceptional collection.*

19.6.2. *Our case.* Let us consider the case from Subsection 19.5.1, but with all coefficients c_i equal to $c \leq b$. Let B be the image of the real torus in the Picard group as described in the Subsection 19.3. One can see that:

$$B = \{\mathcal{O}([\sum_{i=1}^{r} -\alpha_z^i]D_t + [\sum_{i=1}^{n-r} -\alpha_v^i + c\sum_{i=2}^{r} \alpha_z^i - (b+1)\sum_{i=1}^{r} \alpha_z^i)]D_y +$$

$$+[\sum_{i=1}^{n-r} -\alpha_v^i + c\sum_{i=2}^{r}\alpha_z^i - b\sum_{i=1}^{r}\alpha_z^i)]D_u) : 0 \le \alpha_v^i, \alpha_z^i < 1\}.$$

So B is contained in the set:

$$S := \{\mathcal{O}(-sD_t - sD_y + qD_v), \mathcal{O}(-sD_t - (s-1)D_y + qD_v) : s \in \{0, \dots, r\},$$

$$q \in \{-(n-r) - c - (b-c)s), \dots, (b-c)(-s+1)\}\}$$

Our collection defined in Subsection 19.5.1, or its torsion, is contained in the set S unless $cr \le b$. It can be also shown that if this inequality fails then there is no full strongly exceptional collection among line bundles that come from Bondal's construction.

19.7. $\mathbb{P}^n$ **blown up in two points.** The results of this section can be found in [Mic11a].

The varieties we consider are of Picard number 3. Using the classification of Theorem 19.7 $\mathbb{P}^n$ blown up in two points is given by $|X_0| = |X_2| = |X_3| = |X_4| = 1$ and $|X_1| = n-1$ with all other parameters equal to 0. Choosing the basis of the one parameter subgroups lattice N equal to $v_1, y_2, \dots, y_{n-1}, z_1$ the ray generators of the fan are the basis elements and vectors y_1, t_1, u_1 satisfying:

$$t_1 = -z_1, \quad y_1 = -y_2 - \dots - y_{n-1} - z_1 - v_1, \quad u_1 = -v_1.$$

The rank of the Grothendieck group is equal to the number of maximal cones that is $3n-1$. All divisors in a given X_i are linearly equivalent and, as before, are given by D_v, D_y, D_z, D_t, D_u respectively for $i = 0, 1, 2, 3, 4$. Divisors with nonzero higher cohomology will be called forbidden. The following classification of forbidden divisors is very easy to establish. In a general case of Picard number three this has been done in the previous section, but in this special case one can use arguments of elementary topology. The forbidden divisors in our case are $\alpha_1 D_v + \alpha_2 D_y + \alpha_3 D_z + \alpha_4 D_t + \alpha_5 D_u$, where exactly $2, 3$ or 5 consecutive (in a cyclic way, that is indices are considered modulo 5) α's are negative and if $\alpha_2 < 0$, then $\alpha_2 \le -n+1$.

We have $D_z = D_t + D_y$ and $D_v = D_u + D_y$. We choose the basis D_y, D_t, D_u, what gives us forbidden divisors $(\alpha_1 + \alpha_2 + \alpha_3)D_y + (\alpha_3 + \alpha_4)D_t + (\alpha_1 + \alpha_5)D_u$ with the conditions on α's as above. A divisor $aD_y + bD_t + cD_u$ will be denoted by (a, b, c) and we reserve precise letters for precise coordinates. A line bundle L_1 will be called compatible with L_2 if and only if they can both appear in a strongly exceptional collection, that is if and only if $L_1 - L_2$ and $L_2 - L_1 = -(L_1 - L_2)$ are not forbidden.

Let us fix a strongly exceptional collection E. We assume without loss of generality that $0 \in E$ and that all other divisors in E have nonnegative coefficient a.

Lemma 19.61. *The only divisors with $a = 0$ compatible with $(0, 0, 0)$ are:*

$$(0, -1, 0), (0, 0, -1), (0, 1, 0), (0, 0, 1), (0, -1, 1), (0, 1, -1).$$

Proof. If $b < -1$, then we take $\alpha_1 = 0$, $\alpha_2 = 1$, $\alpha_3 = -1$, α_4-negative to obtain b, α_5-any to obtain c. Analogously for $c < -1$, hence $-1 \leq b, c \leq 1$. Moreover $(0, -1, -1)$ is also bad (so also $(0, 1, 1)$). □

Corollary 19.62. *There can be at most 3 distinct line bundles with $a = 0$ in E. For a fixed a we can have only 3 line bundles in E.*

Proof. Follows by inspection. □

Lemma 19.63. *For $a > 0$ the only line bundles (a, b, c) that are not forbidden must satisfy $-1 \leq b \leq a$ and $-1 + a - b \leq c \leq a$ (and by symmetry $-1 + a - c \leq b \leq a$).*

Proof. For $b < -1$ we take $\alpha_1 = 0$, $\alpha_3 = -1$, α_2-positive to have a, α_4-negative to have b, α_5-any to have c. For $b > a$ we look at $(-a, -b, -c)$ and take $\alpha_3 = -a$, $\alpha_1 = \alpha_2 = 0$, α_4-negative to have $-b$, α_5-any to have $-c$. In the same way $-1 \leq c \leq a$[15]. So the only case that we have to exclude is $-1 \leq c < -1 + a - b$. In such a case we can take $\alpha_4 = -1$, $\alpha_3 = b+1$, $\alpha_2 = 0$, $\alpha_1 = a-b-1$, $\alpha_5 = c-a+b+1 < 0$. □

Lemma 19.64. *For three consecutive parameters a's there can be at most 8 line bundles in E.*

Proof. We assume without loss of generality $0 \leq a \leq 2$. If the lemma does not hold, then from the Corollary 19.62 we would have to have 3 line bundles for each a. For $a = 0$ we can have either:

Case 1: $(0, 0, 0), (0, -1, 0), (0, 0, -1)$ then for $a = 1$ there is only one compatible from the Lemma 19.63 namely $(1, 0, 0)$.

Case 2:$(0, 0, 0), (0, 1, 0), (0, 0, 1)$ then for $a = 1$ the compatible line bundles are $(1, 1, 1), (1, 1, 0), (1, 0, 1)$. If we choose all of them then the only one compatible for $a = 2$ is $(2, 1, 1)$ from the Lemma 19.63.

Case 3: $(0, 0, 0), (0, -1, 0), (0, -1, 1)$; $(0, 0, 0), (0, 0, -1), (0, 1, -1)$; $(0, 0, 0), (0, 1, 0), (0, 1, -1)$; $(0, 0, 0), (0, 0, 1), (0, -1, 1)$. All these possibilities are cases 1 or 2 after subtracting a divisor from all three considered divisors. □

[15]The parameters b and c are in symmetry.

Definition 19.65. *Line bundles in the collection E with $a > n$ are called high. Others are called low.*

Lemma 19.66. *A high line bundle is forbidden unless either $b = 1$ (high bundles of type 1)or $c = 1$ (high bundles of type 2).*

Proof. Suppose that $b = 0$ or $b = -1$. We show that $(-a, -b, -c)$ is forbidden. Take $\alpha_1 = -1$, $\alpha_2 = -a + 1$, $\alpha_3 = 0$, $\alpha_4 = -b$, α_5-any to obtain $-c$. So $b \geq 1$ and analogously $c \geq 1$. If both coefficients were strictly greater than 1 we would obtain $(-a, -b, -c)$ by taking all α's negative. □

Lemma 19.67. *We cannot have high line bundles of both types in E.*

Proof. From the Lemma 19.63 a high line bundle must have the coordinate different from 1 greater or equal to $n - 1$. If we subtract two high line bundles of different types we can assume that the first coordinate is positive and one of the others will be less or equal to $-n + 2$ what contradicts the Lemma 19.63 for $n > 3$. □

From now on without loss of generality we assume that we only have high line bundles of type 1 in E. Let us project all high line bundles from E onto the first coordinate obtaining a subset of $\mathbb{N}$. Suppose that this subset has got k elements, that is high line bundles can have k different parameters a. We obtain:

Lemma 19.68. *There are at most $k + 2$ high line bundles in E.*

Proof. We assumed that $0 \in E$, so the high line bundles in E must not be forbidden. We know that for each high line bundle in E we have $b = 1$, so from the Lemma 19.63 we know that $0 \leq a - c \leq 2$. Let us notice that the difference $a - c$ cannot decrease when a increases for high line bundles in E. Indeed suppose that we have two high line bundles in E of the form $(a_1, 1, c_1)$, $(a_2, 1, c_2)$ with $a_2 > a_1$ and $a_2 - c_2 < a_1 - c_1$. By subtracting these two line bundles we obtain $(a_2 - a_1, 0, c_2 - c_1)$ that is forbidden by the Lemma 19.63.

Notice that each time we have more than one line bundle for a fixed a then the difference $a - c$ strictly increases. This means that we can have one line bundle for each a plus possibly two more as $a-c$ increases from 0 to 2. This gives us in total $k + 2$ line bundles.

□

Proposition 19.69. *There are at most $\frac{8}{3}(n - 1) + 6$ low line bundles (from the Lemma 19.64), so $k > 0$ for $n > 13$.*

Remark 19.70. *Of course k is at most $n + 1$. Otherwise we would have two high line bundles in E with the difference that is high. By* *the Lemma 19.67 the difference would have $b = 0$, hence by the Lemma 19.66 it would have $c = 1$ and would be forbidden by the Lemma 19.63.*

From the definition of k we know that there is a line bundle $L = (a, 1, c)$ in E, with $a \geq n + k$. Now we investigate line bundles with $a < k$, that are called very low.

Lemma 19.71. *Each very low line bundle in E must have $b = 0$.*

Proof. Let B be a very low line bundle. $L - B$ is high, so from the Lemma 19.66 either the second or third coordinate is 1. The third one is $c_L - c_B \geq a_L - 2 - a_B > n + k - 2 - k = n - 2 > 1$, for $n > 3$. We see that $b_L - b_B = 1$. As $b_L = 1$ the Lemma follows. □

For very low line bundles in E the parameter c is either a or $a - 1$ by the Lemma 19.63 and the Lemma 19.71. Reasoning analogously to the proof of the Lemma 19.68, we see that there are at most $k+1$ very low line bundles (the difference $a - c$ cannot decrease).

Theorem 19.72. *The sequence E can have at most: $k+1+\frac{8}{3}(n-k-1)+6+k+2 \leq \frac{8}{3}n - \frac{2}{3}k + \frac{19}{3} < 3n - 1$ for $n > 20$.*

Remark 19.73. *The bounds on n can be easily improved. For example by considering separately the case $k = 1$ one can decrease the bound to $n > 18$. We concentrated rather on brevity of the proof than sharp bounds.*

References

[Ach] Piotr Achinger, *A note on the frobenius morphism on toric varieties*, arXiv:1012.2021v2.

[APRS11] Elizabeth S. Allman, Sonja Petrović;, John A. Rhodes, and Seth Sullivant, *Identifiability of two-tree mixtures for group-based models*, IEEE/ACM Transactions on Computational Biology and Bioinformatics **8** (2011), 710–722.

[AR06] Elizabeth S. Allman and John A. Rhodes, *The identifiability of tree topology for phylogenetic models, including covarion and mixture models*, J. Comput. Biol. **13** (2006), no. 5, 1101–1113.

[AR08] ———, *Phylogenetic ideals and varieties for the general Markov model*, Advances in Applied Mathematics **40(2)** (2008), 127–148.

[Bat91] Victor V. Batyrev, *On the classification of smooth projective toric varieties*, Tohoku Math. J. (2) **43** (1991), no. 4, 569–585.

[BBKM10] Weronika Buczyńska, Jarosław Buczyński, Kaie Kubjas, and Mateusz Michalek, *Degrees of generators of phylogenetic semigroups on graphs*, arXiv:1004.1183v1 [math.AG] (2010).

[BDW09] Weronika Buczyńska, Maria Donten, and Jarosław A. Wiśniewski, *Isotropic models of evolution with symmetries*, Contemporary Mathematics **496** (2009), 111–132.

[BH06] Lev A. Borisov and R. Paul Horja, *On the K-theory of smooth toric DM stacks*, Snowbird lectures on string geometry, Contemp. Math., vol. 401, Amer. Math. Soc., Providence, RI, 2006, pp. 21–42.

[BH09] Lev Borisov and Zheng Hua, *On the conjecture of King for smooth toric Deligne-Mumford stacks*, Adv. Math. **221** (2009), no. 1, 277–301.

[BHV01] Louis J. Billera, Susan P. Holmes, and Karen Vogtmann, *Geometry of the space of phylogenetic trees*, Adv. in Appl. Math. **27** (2001), no. 4, 733–767.

[BO01] Alexei Bondal and Dmitri Orlov, *Reconstruction of a variety from the derived category and groups of autoequivalences*, Compositio Math. **125** (2001), no. 3, 327–344.

[BO10] D. J. Bates and L. Oeding, *Toward a salmon conjecture*, ArXiv e-prints (2010).

[Bon89] A. I. Bondal, *Representations of associative algebras and coherent sheaves*, Izv. Akad. Nauk SSSR Ser. Mat. **53** (1989), no. 1, 25–44.

[Bon06] A.I. Bondal, *Derived categories of toric varieties*, Oberwolfach Report **5** (2006), 284–286.

[Bri02] Tom Bridgeland, *Flops and derived categories*, Invent. Math. **147** (2002), no. 3, 613–632.

[BT09] Alessandro Bernardi and Sofia Tirabassi, *Derived categories of toric Fano 3-folds via the Frobenius morphism*, Matematiche (Catania) **64** (2009), no. 2, 117–154.

[Buc12] Weronika Buczyńska, *Phylogenetic toric varieties on graphs*, Journal of Algebraic Combinatorics **35(3)** (2012), 421–460.

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

[Căl05] Andrei Căldăraru, *Derived categories of sheaves: a skimming*, Snowbird lectures in algebraic geometry, Contemp. Math., vol. 388, Amer. Math. Soc., Providence, RI, 2005, pp. 43–75.

[CFS08] M. Casanellas and J. Fernandez-Sanchez, *Geometry of the Kimura 3-parameter model*, Advances in Applied Mathematics **41(3)** (2008), 265–292.

[CLS] D. Cox, J. Little, and H. Schenck, *Toric varieties*, preprint at http://www.cs.amherst.edu/ dac/toric.html.

[CMRa] L. Costa and R.M. Miró-Roig, *Derived category of toric varieties with small picard number*, (preprint).

[CMRb] ———, *Frobenius splitting and derived category of toric varieties*, arXiv:1006.5315v1.

[CMR04] L. Costa and R. M. Miró-Roig, *Tilting sheaves on toric varieties*, Math. Z. **248** (2004), no. 4, 849–865.

[CP07] J. Chifman and S. Petrović, *Toric ideals of phylogenetic invariants for the general group-based model on claw trees $k_{1,n}$*, Proceedings of the 2nd international conference on Algebraic biology (2007), 307–321.

[DBM] Maria Donten-Bury and Mateusz Michałek, *Phylogenetic invariants for group-based models*, arXiv:1011.3236v1.

[DK09] Jan Draisma and Jochen Kuttler, *On the ideals of equivariant tree models*, Mathematische Annalen **344(3)** (2009), 619–644.

[DK11] ———, *Bounded-rank tensors are defined in bounded degree*, arXiv:1103.5336v2 (2011).

[DLM09] Arijit Dey, Michał Lasoń, and Mateusz Michałek, *Derived category of toric varieties with Picard number three*, Matematiche (Catania) **64** (2009), no. 2, 99–116.

[Efi] A. Efimov, *Maximal lengths of exceptional collections of line bundles*, arXiv 1010.3755v2.

[EKS11] Alexander Engstrom, Thomas Kahle, and Seth Sullivant, *Multigraded commutative algebra of graph decompositions*, arXiv:1102.2601v2 (2011).

[ERSS04] N. Eriksson, K. Ranestad, B. Sturmfels, and S. Sullivant, *Phylogenetic algebraic geometry*, Projective Varieties with Unexpected Properties; Siena, Italy (2004), 237–256.

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

[ES96] David Eisenbud and Bernd Sturmfels, *Binomial ideals*, Duke Math. J. **84** (1996), no. 1, 1–45.

[Fel04] Joseph Felsenstein, *Inferring phylogenies*, 2004.

[Ful93] William Fulton, *Introduction to toric varieties*, Annals of Mathematics Studies, vol. 131, Princeton University Press, Princeton, NJ, 1993, The William H. Roever Lectures in Geometry.

[GAP] *GAP — Groups, Algorithms, and Programming, Version 4.4.10*, The GAP Group, (http://www.gap-system.org), 2007.

[GJ00] Ewgenij Gawrilow and Michael Joswig, *polymake: a Framework for Analyzing Convex Polytopes*, Polytopes — Combinatorics and Computation (Gil Kalai and Günter M. Ziegler, eds.), Birkhäuser, 2000, pp. 43–74.

[GM03] Sergei I. Gelfand and Yuri I. Manin, *Methods of homological algebra*, second ed., Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2003.

[GR87] A. L. Gorodentsev and A. N. Rudakov, *Exceptional vector bundles on projective spaces*, Duke Math. J. **54** (1987), no. 1, 115–130.

[Grü03] Branko Grünbaum, *Convex polytopes*, second ed., Graduate Texts in Mathematics, vol. 221, Springer-Verlag, New York, 2003, Prepared and with a preface by Volker Kaibel, Victor Klee and Günter M. Ziegler.

[GS] Daniel R. Grayson and Michael E. Stillman, *Macaulay2, a software system for research in algebraic geometry*, Available at http://www.math.uiuc.edu/Macaulay2/.

[Har77] Robin Hartshorne, *Algebraic geometry*, Springer-Verlag, New York, 1977, Graduate Texts in Mathematics, No. 52.

[HP] L. Hille and M. Perling, *Exceptional sequences of invertible sheaves on rational surfaces*.

[HP06] Lutz Hille and Markus Perling, *A counterexample to King's conjecture*, Compos. Math. **142** (2006), no. 6, 1507–1521.

[HS11] Christopher J. Hillar and Seth Sullivant, *Finite groebner bases in infinite dimensional polynomial rings and applications*, arXiv:0908.1777v2 (2011).

[Huy06] D. Huybrechts, *Fourier-Mukai transforms in algebraic geometry*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, Oxford, 2006.

[Ibe09] Oliver C. Ibe, *Markov processes for stochastic modeling*, Elsevier/Academic Press, Amsterdam, 2009.

[Ilt10] Nathan Ilten, *Deformations of Rational Varieties with Codimension-One Torus Action*, Doctoral Thesis, FU Berlin, 2010.

[JC69] T. Jukes and C. Cantor, *Evolution of protein molecules.*, Manmmalian Protein Metabolism (1969), 21–132.

[JW92] Lisa C. Jeffrey and Jonathan Weitsman, *Bohr-Sommerfeld orbits in the moduli space of flat connections and the Verlinde dimension formula*, Comm. Math. Phys. **150** (1992), no. 3, 593–630.

[Kaw06] Y. Kawamata, *Derived categories of toric varieties*, Michigan Math. J. **54** (2006), 517–535.

[Kim80] M. Kimura, *A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences*, Journal of Molecular Evolution **16** (1980), 111–120.

[Kim81] ______, *Estimation of evolutionary distances between homologous nucleotide sequences*, Proceedings of the National Academy of Sciences **78** (1981), no. 1, 454–458.

[KKMSD73] G. Kempf, Finn Faye Knudsen, D. Mumford, and B. Saint-Donat, *Toroidal embeddings. I*, Lecture Notes in Mathematics, Vol. 339, Springer-Verlag, Berlin, 1973.

[Kub10] Kaie Kubjas, *Hilbert polynomial of the kimura 3-parameter model*, arXiv:1007.3164 (2010).

[LM11] Michał Lasoń and Mateusz Michałek, *On the full, strongly exceptional collections on toric varieties with picard number three*, Collectanea Mathematica **62** (2011), 275–296.

[Man09] C. Manon, *The algebra of Conformal Blocks*, arXiv:0910.0577 (2009).

[Man11] ______, *Coordinate rings for the moduli of $sl_2(\mathbb{C})$ quasi-parabolic principal bundles on a curve and toric fiber products*, arXiv:1105.2045v1 (2011).

[MB09] Marian Mrozek and Bogdan Batko, *Coreduction homology algorithm*, Discrete Comput. Geom. **41** (2009), no. 1, 96–118.

[Mic11a] Mateusz Michałek, *Family of counterexamples to King's conjecture*, C. R. Math. Acad. Sci. Paris **349** (2011), no. 1-2, 67–69.

[Mic11b] Mateusz Michałek, *Geometry of phylogenetic group-based models*, Journal of Algebra **339** (2011), 339–356.

[Ney71] Jerzy Neyman, *Molecular studies of evolution: A source of novel statistical problems*, Statistical decision theory and related topics (Proc. Sympos., Purdue Univ., Lafayette, Ind., 1970), Academic Press, New York, 1971, pp. 1–27.

[Oda87] T. Oda, *Convex bodies and algebraic geometry*, Springer, 1987.

[OP91] Tadao Oda and Hye Sook Park, *Linear Gale transforms and Gel′fand-Kapranov-Zelevinskij decompositions*, Tohoku Math. J. (2) **43** (1991), no. 3, 375–399.

[PS05] Lior Pachter and Bernd Sturmfels (eds.), *Algebraic statistics for computational biology*, Cambridge University Press, New York, 2005.

[SS03] C. Semple and M. Steel, *Phylogenetics*, Oxford University Press, 2003.

[SS05] Bernd Sturmfels and Seth Sullivant, *Toric ideals of phylogenetic invariants*, J. Comput. Biology **12** (2005), 204–228.

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

[Stu96] Bernd Sturmfels, *Gröbner bases and convex polytopes*, University Lecture Series, vol. 8, American Mathematical Society, 1996.

[Stu98] Bernd Sturmfels, *Polynomial equations and convex polytopes*, Amer. Math. Monthly **105** (1998), no. 10, 907–922.

[Sul07] Seth Sullivant, *Toric fiber products*, Journal of Algebra **316** (2007), no. 2, 560 – 577, Computational Algebra.

[SV10] Bernd Sturmfels and Mauricio Velasco, *Blow-ups of* $\mathbb{P}^{n-3}$ *at* n *points and spinor varieties*, J. Commut. Algebra **2** (2010), no. 2, 223–244.

[SX10] Bernd Sturmfels and Zhiqiang Xu, *Sagbi bases of Cox-Nagata rings*, Journal of the European Mathematical Society **12** (2010), 429–459.

[Tho00] Jesper Funch Thomsen, *Frobenius direct images of line bundles on toric varieties*, J. Algebra **226** (2000), no. 2, 865–874.

[tt] 4ti2 team, *4ti2—a software package for algebraic, geometric and combinatorial problems on linear spaces*, www.4ti2.de.

[Zwi] Piotr Zwiernik, *Toric geometry of statistical binary graph models*, Master Thesis, Warsaw University.