UNIVERSITÉ DE NEUCHÂTEL

Faculté des Sciences
Institut de mathématiques
Rue Emile-Argand 11, 2000 Neuchâtel

Some applications of combinatorics: from group theory to mathematical ecology

Thèse

présentée à la Faculté des Sciences
pour l'obtention du grade de Docteur ès Sciences en mathématiques

par

Laura Grave de Peralta Gonzalez

Acceptée sur proposition du jury :

Prof. Aleksandr Kolpakov, Université de Neuchâtel, directeur de thèse
Prof. Alain Valette, Université de Neuchâtel, rapporteur
Prof. Laura Ciobanu, Heriot–Watt University, rapporteuse
Prof. Christian Mazza, Université de Fribourg, rapporteur
Prof. Roman Nedela, University of West Bohemia, rapporteur

Soutenue le 23 août 2021

UNIVERSITÉ DE NEUCHÂTEL

Faculté des Sciences
Secrétariat-décanat de Faculté
Rue Emile-Argand 11
2000 Neuchâtel – Suisse
Tél : + 41 (0)32 718 21 00
E-mail : secretariat.sciences@unine.ch

IMPRIMATUR POUR THESE DE DOCTORAT

La Faculté des sciences de l'Université de Neuchâtel autorise l'impression de la présente thèse soutenue par

Madame Laura GRAVE DE PERALTA

Titre:

"Some applications of combinatorics: from group theory to mathematical ecology"

sur le rapport des membres du jury composé comme suit:

- Prof. ass. Aleksandr Kolpakov, directeur de thèse, Université de Neuchâtel, Suisse
- Prof. Alain Valette, Université de Neuchâtel, Suisse
- Prof. Laura Ciobanu, Heriot-Watt University, Edinburgh, United Kingdom
- Prof. Christian Mazza, Université de Fribourg, Suisse
- Prof. Roman Nedela, University of West Bohemia, Czech Republic

Neuchâtel, le 16 septembre 2021

Le Doyen, Prof. A. Bangerter

Remerciements

Sasha, I feel really lucky to having done my PhD with you. You were an awesome director. You provided many fruitful insights and guidance but also gave me the freedom to embark in unrelated projects. We discussed many hours about math but also about careers and the future. It has been a real pleasure working with you and I hope you don't regret hiring me as your PhD student.

Alain, je ne saurai te remercier pour ta confiance. J'ai pris beaucoup de plaisir à discuter avec toi et te suis très reconnaissante pour tout le temps que tu m'as accordé et ton soutien. Merci de m'avoir partagé un peu de ta sagesse et pour ta bienveillance à mon égard !

Thanks to the members of the jury,Laura Ciobanu, Christian Mazza, Roman Nedela and Alain Valette, for the kind comments and the various remarks on how to improve this manuscript and the time you spent reading it!

Mis amores, se merecerian paginas y paginas. Pero se supone que escriba de matematicas y no un manuscrito detallando cuanto y porque los quiero. Largas horas pasamos conversando y discutiendo de matematicas y de mucho mas. Gracias por su apoyo, los quiero.

Victor, merci pour ton temps, ta patience et ta tendresse sans égale. Et tes TikZ, sans quoi, cette thèse n'aurait aucune figure. Rien n'aurait été pareil sans toi.

Cass, Juls, Ophé, Steve, Yas, merci d'être toujours là pour fêter comme pour pleurer. Si les années à venir sont aussi droles que celles déjà passées ensemble, je suis comblée.

Terence, Léo, Sylvain, Livio, Lucie, Lucia, Donik, Flavien, Gaëtan, Vass, je suis heureuse que notre amitié aie perduré après l'AEM. Vous êtes les meilleurs.

ZumZum, mon compagnon de galère, merci pour ton accueil. Tu as fait de l'institut une deuxième maison et cette thèse aurait été beaucoup moins drôle sans ton accent et tes irruptions dans mon bureau.

Alex, Antoine, Jérémy, Joe, Laurent, Léo, Maël, Thomas, merci pour ces mercredis soirs de folie. Cette dernière année avec vous a été splendide, un peu trop de damassine, mais splendide!

Merci Rémi pour ta patience au moment de m'expliquer de l'informatique et pour tes intuitions toujours éclairantes ! J'ai beaucoup apprécié nos moments ragots et plaintes ainsi que nos discussions plus sérieuses.

Christine merci pour votre bonne humeur. Vos conseils et votre soutien ont été précieux pour le projet de médiation et pour bien plus. L'institut à de la chance de vous avoir!

Merci à Matthieu de m'avoir initié à la biologie ! J'ai beaucoup apprécié nos discussions et nos échanges et me réjouis de continuer à travailler ensemble et à faire des dessins incompréhensibles tout en s'exprimant en des termes absurdes pour le commun des mortels.

Thanks to the members of the jury for the kind comments, the various remarks on how to improve this manuscript and the time spent reading it.

Abstract

This thesis is composed of four apparently different chapters that each showcase applications of algebraic combinatorics and algebraic graph theory to the study of diverse subjects. The problems tackled here range from enumeration of combinatorial and algebraic structures to mathematical biology.

In the first chapter, we investigate *dessins d'enfants*, meaning "children's drawings". This name was given by Grothendieck. Nowadays they are called combinatorial maps and in the first chapter we focus on them and their generalizations. We prove that such combinatorial objects can have any finite group of automorphisms, and we describe an algorithm to construct such a combinatorial structure with given symmetry group.

The second chapter looks at enumerating subgroups of Hecke groups. By trying to enumerate graphs and using the powerful tools of generating functions, we give asymptotic estimates regarding the number of finite index subgroups of such groups.

The third chapter looks at a space of graphs as a combinatorial model for pants decomposition of surfaces. Closely related to Teichmüller space, this combinatorial model focuses on the so-called *wide part of the moduli space*. We look at the expansion properties of this graph of graphs.

Finally, the last chapter leaps further away from maths and looks at graphs as a model for trophic interactions and uses machinery from graph theory to try to detect vulnerability in ecological networks modelling microbiomes. We develop a new approach to detect and characterize important species in an ecological network, taking into account both their position in the ecological network and their trophic position and contribution within the biogeochemical cycles driving the ecosystem's stability.

Keywords: combinatorics, graph theory, group theory, geometry, ecology, combinatorial maps, subgroup growth, Hecke groups, pants decomposition, expansion,

Résumé

Cette thèse comporte quatre chapitres qui, à première vue, peuvent sembler n'avoir pas grand-chose à faire ensemble. Cependant, un fil rouge est toujours présent : la théorie des graphes et la combinatoire. Chaque chapitre applique des outils de combinatoire, de la théorie algébrique des graphes en passant par les fonctions génératrices, nous appliquons des méthodes combinatoires pour attaquer divers problèmes. Au fil des chapitres, nous passerons de la théorie géométrique des groupes à l'écologie théorique.

Dans le premier chapitre, nous nous penchons sur les "dessins d'enfants" (surnommés ainsi par Grothendieck) ou également appelés cartes combinatoires (et leurs généralisations). Nous prouvons que ces objets peuvent admettre tout groupe fini comme groupe d'automorphismes. En prouvant cela, nous exhibons un algorithme permettant de générer une grande variété de cartes combinatoires ayant pour groupe de symétries un groupe fini donné.

Le second chapitre traite de l'énumération asymptotique des sous-groupes des groupes de Hecke. Nous associons à chaque sous-groupe d'un groupe de Hecke un graphe de manière canonique. Ensuite, en utilisant la théorie combinatoire des espèces de structures, nous énumérons les graphes et donc ainsi les sous-groupes.

Le troisième chapitre se consacre à l'étude de l'espace des graphes 3-réguliers. Ces graphes sont un modèle pour les décompositions en pantalons de surfaces et donc l'espace de tous ces graphes, reliés par des transformations appelées mouvements de Whitehead, sont un modèle d'une partie de l'espace de Teichmüller d'une surface de genre g. Nous prouvons dans ce chapitre que cet espace n'est pas expanseur, pour une définition d'expansion adaptée au cas d'une famille des graphes dirigés dont le degré est croissant.

Finalement, le dernier chapitre est une incursion dans le monde de l'écologie théorique. Nous étudions les réseaux trophiques et les réseaux de cooccurrence et proposons une nouvelle approche pour détecter des nœuds dont le rôle est crucial pour la stabilité et la cohésion de l'écosystème en question. Notre approche prend en compte à la fois le réseau trophique dans son ensemble ainsi que la position trophique de chacune des espèces considérées.

Mots clés : combinatoire, théorie des graphes, théorie des groupes, géométrie, écologie, maps combinatoires, énumération de sous-groupes, groupes de Hecke, décompositions en pantalons, expansion

Contents

Introduction

Graph theory has a long history. It dates back to the works of Euler in 1736 who introduced graphs in order to solve the so-called *seven bridges of Königsberg* problem. The great power of graphs is that as simple as they may seem, they can still model a wide range of problems. Each chapter of this thesis will use and exploit some particular type of graph that we will briefly present hereunder.

Intuitively, when we think of graphs we usually mean a set of vertices, usually graphically represented as points, and a set of edges, represented by lines joining pairs of points. Graphs can also be endowed with additional structure. Graphs can be oriented, we call them *directed graphs*. In this case, the edges have a direction and are thus represented by pointing arrows. Graphs can be labelled, we call them *labelled graphs*, or coloured graphs, that is every edge is labelled by a letter or drawn in a specific colour. Graphs can have multiple edges or loops. If they don't, we call them simple. We can also combine these structures and for example work with labelled oriented simple graphs.

In this introduction, we will present for each subject the types of graph we considered (whether they are oriented, simple, labelled, etc.) and how they can be used to model the problems we are trying to solve.

Combinatorial maps and their symmetries

In Chapter 1, we study Schreier (coset) graphs of subgroups of free products of cyclic groups. We present a method for constructing such graphs with a given finite group of automorphisms, thus proving that combinatorial maps and their generalizations admit any finite group as symmetries.

Combinatorial maps can be seen as graphs where edges are constructed by gluing two half-edges (called darts) together.

Formally, a combinatorial (oriented, labelled) map is a triple $M = (D; R, L)$ where D is a non-empty finite set (called the set of darts) and R and L are two permutations of D with $L^2 = \mathrm{id}$. The orbits of R describe how the darts are glued around each vertex and the orbits of L describe how the darts are joined two by two to form the edges.

One can exhibit a one-to-one correspondence between combinatorial maps and subgroups of free products of cyclic groups. The rooted maps on n darts (where the root is always supposed to be marked 1) are in a one-to-one correspondence with index n free subgroups of $\Delta^+ = \mathbb{Z} * \mathbb{Z}_2$. See Section 1.1 for more details.

The isomorphisms classes of all combinatorial maps (and their generalizations) correspond to the conjugacy classes of free subgroups of Δ^+ [JS78, Theorem 3.7]. The symmetries (i.e. unrooted self-isomorphisms) of a map M corresponding to a subgroup $H < \Delta^+$ form a group isomorphic to $N(H)/H$, where $N(H) = \{g \in \Delta^+ \mid gHg^{-1} = H\}$ is the normalizer of H in Δ^+ [JS78, Theorem 3.8].

To study symmetries of maps, one can focus on studying their corresponding subgroups and one can do so via their Schreier graphs. Given a group G, a set of generators S and a subgroup H, one can construct a labeled, oriented graph. The vertices correspond to the cosets gH, and the edges are labelled by the generators in S. Schreier graphs provide a combinatorial setting in which several algebraic properties can be described, e.g. being of finite index or being free. In particular, we develop a method to construct subgroups of Δ^+ using "LEGO–like" blocks and in a manner that we have control over the automorphism group of the graph.

From these ideas, we can prove our main result,Theorem 1.3.7, that can be summarized to :

Theorem (Theorem 1.3.7)**.** *For any finite group* Γ*, there exists a combinatorial map with symmetry group* Γ.

Our proof is constructive and thus provides tools for the enumeration of combinatorial maps with given symmetry group. Theorem 1.6.1 provides asymptotic bounds on the number of combinatorial maps. More precisely, it states that:

Theorem. *Let T be a finite free product of cyclic groups, different from $\mathbb{Z}_2 * \mathbb{Z}_2$. Then for any finite group Γ, there exist constants $A > 1$, $B > 0$ and $M \in \mathbb{N}$ such that for all $d \geq M$ the set $F(T,\Gamma,d) = \{$free subgroups $H < T$ of index $\leq d$ with $N_T(H)/H \cong \Gamma$, up to conjugacy$\}$ has cardinality $\geq A^{Bd\log d}$.*

Chapter 1 is organized as follows. Section 1.1 is a general introduction to combinatorial maps and their generalizations. We present historical references as well as recent results from other authors. In Section 1.2, we state the preliminaries needed for the formulation and the proof of Theorem 1.3.7. These include results about Schreier graphs and their automorphism groups, as well as the correspondence between subgroups of free products of cyclic groups and their associated graphs. In Section 1.3, we state the main result and give the general sketch of the proof. The main idea of the proof is described, and we explain how the "LEGO–like" pieces are constructed. Section 1.4 deals with the technical details of the proof. The reader interested only in the ideas may skip this part. Section 1.5 states some extensions of the result. Finally, Section 1.6 gives some asymptotic estimates about the number of combinatorial objects we can construct given a group of symmetries and Section 1.7 apply and generalize our results to other combinatorial structures.

The results presented in Chapter 1 can be found in [BGdPK20].

Enumeration of graphs and generating functions applied to subgroup growth

In Chapter 2 we study the asymptotic behaviour of the number of subgroups of finite index n of $\Delta_q = \mathbb{Z}_2 * \mathbb{Z}_q$, often called Hecke groups. To do so, we establish a correspondence between finite index subgroups of Δ_q and finite, labelled graphs, and we use the machinery of generating functions to establish the asymptotic behaviour of the number of subgroups of finite index n of Δ_q.

One can always associate the Schreier (or coset) graph to a subgroup H of G. The Schreier graph is finite only if the subgroup is of finite index. For free groups, Stallings introduced in [Sta83] the Stallings graph of a subgroup. This graph is a subgraph of the Schreier graph which is finite if and only if the subgroup is finitely generated. Stallings graph provide a combinatorial and finite description of all finite subgroups of a free group.

For $H \leq F(X)$, the Stallings graph is the core of the Schreier graph in a very literal sense of the term. You start from the Schreier graph and you prune its ends and remove the tree-like parts of the graph. If H is finitely generated, you are left with a finite graph.

This approach can be generalized to more than just free groups and was done partly in [KMW17]. In particular, one can associate to any finitely generated subgroup of Δ_q a finite labelled graph that is minimal in the sense that any superfluous part of the Schreier graph is pruned and you are left only with the core which contains much information about your subgroup.

The Stallings graph encodes information, such as if the subgroup is of finite index or is free.

For subgroups of Δ_q, the Stallings graph of a finitely generated subgroup is labelled by the two generators a, b. The b edges decompose the vertices of the Stallings graphs into $1-$ and $q-$ sets (that is sets containing 1 or q elements) and the a edges into $1-$ or $2-$sets. To these decompositions of the vertices, we associate two generating functions T_a and T_b. The computation of the coefficients of these exponential generating functions is key to the subgroup growth problem. The main result of Chapter 2 is the following result.

Theorem (Corollary 2.7.9)**.** *The number λ_n of index n subgroups of Δ_q for $q \geq 3$ prime satisfies*

$$\lambda_n \sim \frac{\sqrt{2\pi n}\cdot e^{-1/4}}{2\pi n\sqrt{2q}} n^{n-n/q-n/2}\, exp\left(\left(\frac{1}{2}+\frac{1}{q}-1\right)\cdot n + n^{1/2} + \left(1+\frac{1}{q}-\frac{1}{q^2}\right) n^{1/q}\right)$$

It is important to note that this result is not new. However, it was originally proved using group actions and group representations. The goal of this chapter is to provide a new method, based in graph theory, to make these computations. Moreover, the methods used to prove Corollary 2.7.9 can be applied to enumerating other subgroups, not only the ones of finite index, and also to refine the computations: for example, determining the number of free finite index subgroups of Δ_q.

Chapter 2 is organized as follows. Section 2.2 introduces the general setting of Schreier graphs and the formalism necessary for the subsequent sections. Section 2.3 introduces Stallings graphs in the particular context of free products of cyclic groups. After doing so we can describe more precisely for Δ_q what combinatorial conditions must be satisfied to be the core graph of some subgroup of Δ_q. This is done in Section 2.4 and what properties does the core graph encore are described in Section 2.5. Finally, Section 2.6 contains the computations to derive the main result which is stated in Section 2.7.

Pants decomposition of surfaces and trivalent graphs

An orientable surface Σ_g of genus $g \geq 2$ can be equipped with a hyperbolic metric. However, the metric is not unique. The Teichmüller space of Σ_g is precisely the topological space of all non-equivalent hyperbolic or flat metrics on Σ_g, up to isometries isotopic to the identity and rescaling.

Hyperbolic orientable manifolds can be obtained by gluing *pairs of pants* along their boundaries. A pant is a sphere with 3 boundary components and can be obtained by identifying the three alternating sides of a right angled hexagon.

If the boundaries of the pants are short enough, the geometry of the surface is determined by the combinatorics of the pants decomposition.

In this chapter we study a graph where each vertex represents a cubic graph on $2n$ vertices and two graphs g_1, g_2 are adjacent in Γ_n if g_1 can be obtained from g_2 by performing a Whitehead move, which is a local transformation of the graph corresponding to the change of the gluing of two labelled pants. This graph of graphs gives a combinatorial model for the study of pants decompositions of surfaces.

The main result of Chapter 3 is Theorem 3.2.2, where $\phi_{out}(\Gamma_n)$ is a slight modification of the notion of conductance.

The conductance of a graph is defined as the minimal ratio between the boundary of a subset of vertices and some notion of volume of the subset (usually we define $\mathrm{vol}(A) = \sum_{v \in A} \deg(v)$). This notion is a slight generalization of the isoperimetric constant and satisfies Cheeger-like inequalities. The conductance, and the isoperimetric constant, are related to the idea of measuring how hard it is in general to escape a subset of the graph and thus is related to random walks on the space considered.

Theorem (Theorem 3.2.2)**.** *Let Γ_n be the graph of cubic graphs on $2n$ ($n \geq 1$) vertices connected by Whitehead moves. Then $\phi_{out}(\Gamma_n) \to 0$, as $n \to \infty$.*

Here, we adapt the notion of conductance to a directed graph with constant out-degree. We modify it by taking a new notion of volume where we consider the out-degree of the vertices for normalization. By doing so, the constant we work with still describes geometrically how easy it is to escape a subset.

Chapter 3 is organised as follows. In Section 3.1 we give an introduction to the graph of graphs and define the tools we will use. In Section 3.2 we state the main theorem about the expansion properties of the graph of graph. This result is then proved in Section 3.3. Finally, we state some open questions related to the subject in Section 3.4.

Networks as modelling interactions, the specific case of trophic networks

Ecologists are interested in communities and ecosystems. Briefly, an ecosystem is a set of species living together in some precise geographical setting. In a given ecosystem, the species can interact with each other and each has its role in the global system.

A particularly interesting question for ecologists is understanding the importance of each species in the ecosystem and how the disappearance of a given species will affect the whole system. The interactions of species can be modelled using differential equations. However, these models require to specify the interactions for each pair of species and can depend on various parameters.

Another approach that has proved to be fruitful is to consider a graph where the nodes represent the species and two species are connected by an edge if they interact in some way.

Here we are interested in trophic networks, also called food webs. In this case, the edges represent the relation of one species preying on another. However, when ecologists study microbial ecosystems, they do not have access to such information and use correlation or co-occurence networks (two species are linked if they co-occur enough in the samples collected) as proxies for the trophic network.

One can now ask whether from the trophic network one can infer the importance, or vulnerability, of a given species. Several approaches exist, most of them rely on some notion of centrality, or a combination of multiple centrality measures. These approaches use the network but don't take into account the trophic position of each species in the network.

In Chapter 4, we present a method to identify pivotal species in the network. To do so, we construct a metric on the space of species and extract clusters of similar species. We then focus on the species connecting those clusters. Indeed, species at the intersection of several clusters are keystone, or pivotal, and essential for the stability of the network.

Chapter 4 is organized as follows. In Section 4.1 we give a thorough introduction to the general context. In Section 4.2, we give the necessary biological preliminaries to formalize the question. Section 4.3 presents the algorithm we used to generate coherent trophic networks. The main section, Section 4.4, presents in detail our method. Finally, in Section 4.5 we provide two examples of our method applied to randomly generated trophic networks to make more explicit the previous section.

1 Combinatorial maps

This chapter is based on the results of [BGdPK20].

1.1 Introduction

The objects we study in this chapter were first introduced by Grothendieck and he called them *dessins d'enfants* (children's drawings) because of their apparent simplicity. Here we focus on a slight modificaiton of the original object introduced by Grothendieck and other generalisaitons of said objects.

A combinatorial (oriented, labelled) map is a triple $M = (D; R, L)$ where D is a non-empty finite set (called the set of darts) and R and L are two permutations of D with $L^2 = \mathrm{id}$. The orbits of L are conventionally called the edges of M, the orbits of R are its vertices, and the orbits of $R^{-1}L$ are its faces. The map M is called connected if the group $\langle R, L\rangle$ acts transitively on D. Unless otherwise stated, we shall assume all maps to be connected.

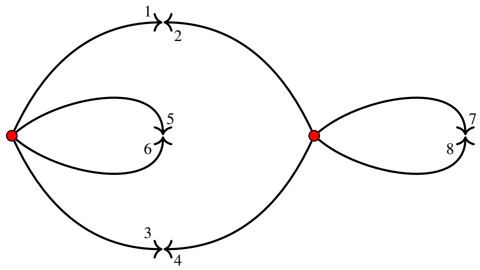


Figure 1.1: An example of a connected combinatorial map with defining permutations $L = (1\ 2)(3\ 4)(5\ 6)(7\ 8)$ and $R = (1\ 5\ 6\ 3)(2\ 7\ 8\ 4)$ on the set of darts $D = \{1, 2, ..., 8\}$.

A topological (oriented) map $M = (\Sigma_g; \Gamma)$ is an oriented (connected) genus $g \geq 0$ surface with an embedded graph Γ such that the complement $\Sigma_g \setminus \Gamma$ is a collection of disjoint topological discs. By providing a labelling on the half-edges of Γ (thus defining its labelled darts), and thus obtaining a labelled topological map, one can recover the permutations L and R, so that L encodes the identification of half-edges into edges, and R encodes the positive cyclic order of half-edges around each vertex. Vice versa, provided a combinatorial map, one can recover its corresponding topological labelled counterpart by creating the faces (which are discs) by following the cycles of $R^{-1}L$, and then identifying their boundaries by using L. For more details, cf. [CM92, JS78, Tut73].

One can define a more elaborate class of combinatorial objects (and the corresponding topological objects) such as hypermaps [Cor75]. A triple $H = (D; R, L)$, where D is a non-empty finite set of darts and R, L are permutation of D, is called an (oriented, labelled) hypermap. The orbits of L are called the hyper-edges of H, the orbits of R are its hyper-faces. A hypermap H is connected whenever the group $\langle R, L\rangle$ acts transitively on D (which will be our standing assumption).

A hypermap also naturally appears in the setting of an orientable genus g surface Σ_g and a graph Γ embedded in Σ_g that satisfies the following properties:

1) the complement $\Sigma_g \setminus \Gamma$ is a union of topological discs called faces,

2) the faces are properly two-colourable (e.g. into black and white), i.e. faces of the same colour intersect only at vertices of Γ, and

3) the corners of the white faces are labelled with the numbers $1, 2, 3, \ldots$ in some fashion, and a black face corner label is equal to the adjacent white face corner label, when moving clockwise around their common vertex.

Then $H = (\Sigma_g; \Gamma)$ is an oriented labelled topological hypermap.

The correspondence between the topological and combinatorial definitions is as follows:

1) each disjoint cycle of R is obtained from recording the corner labels of a white face in a counter-clockwise direction,

2) each disjoint cycle of L is obtained from recording the corner labels of a black face in a counter-clockwise direction,

3) each disjoint cycle of $R^{-1}L$ is obtained from recording the labels around a vertex in a counter-clockwise direction,

We remark that condition (3) above is a consequence of (1) and (2).

The set of face labels becomes the set of darts of H, the white faces become hyper-faces of H and the black faces become hyper-edges of H. Thus, the combinatorial and topological definitions of an oriented labelled hypermap agree.

If $L^2 = \mathrm{id}$, then each bigon in the hypermap $H = (D; R, L)$ can be interpreted as a pair of darts pointing in opposite directions, and thus H becomes a map, as defined above.

We say that two oriented labelled (hyper-)maps $M_1 = (D; R_1, L_1)$ and $M_2 = (D; R_2, L_2)$ are isomorphic if, in the combinatorial setting, there exists a permutation T of D such that $TR_1 = R_2T$ and $TL_1 = L_2T$. In the topological setting, two oriented labelled (hyper-)maps $M_1 = (\Sigma_g; \Gamma_1)$ and $M_2 = (\Sigma_g; \Gamma_2)$ are isomorphic if there exists an orientation-preserving homeomorphism $\tau : \Sigma_g \to \Sigma_g$ such that $\tau(\Gamma_1) = \Gamma_2$ and the labelling of the corresponding half-edges is respected.

A rooted isomorphism will require only the root (a dedicated labelled dart) of one (hyper-)map to be carried to the root of another.

Finally, an isomorphism is not required to respect the dart labelling, nor the roots.

The above definition allows us to generalise the setting of maps to higher-dimensional objects, the so-called pavings. Namely, as defined in [AK89], a three-dimensional oriented combinatorial map or, simply, a (combinatorial) paving, is a quadruple $P = (D; R, L, V)$, where D is a non-empty set of darts and R, L, V are permutations of D such that $H_P = (D; R, L)$ is a map (not necessarily connected), and

1) the product LV is an involution,

2) the product VR^{-1} is an involution,

3) none of the above involutions have fixed points.

A paving P is connected if the group $\langle L, R, V\rangle$ acts transitively on D. The notion of (labelled, rooted) isomorphism for oriented combinatorial pavings is analogous to the one for combinatorial maps.

We may also think of P as a quadruple $P = (D; L, S, T)$, where D is the set of darts and L, S, T are its involutions without fixed points. In this case it is easy to see that letting $V = LS$ and $R = TLS$ produces the initial definition. As in the case of two-dimensional maps, a combinatorial paving P has a topological realisation which, however, is not always a three-dimensional manifold (however, it's always a pseudo-manifold).

In order to assemble an oriented cellular complex M_P, as described in [Spe91], we first produce its underlying map $H_P = (D; R, L)$, and realise each connected component of H as a topological map, i.e. as a surface Σ^i with an embedded graph Γ^i, $i = 1, 2, \ldots, m$, having labelled half-edges. Each surface Σ^i represents the boundary of a handle-body B^i, and then the handle-bodies B^i become identified along their boundaries in order to produce a labelled oriented cellular complex representing P topologically. Indeed, the faces of Σ^i's defined by the permutation $R^{-1}L$ are identified in accordance with the permutation V, and the conditions (1), (2), and (3) above ensure that one face cannot be identified to multiple disjoint counterparts (implied by (1) and (2)), and edges or faces cannot bend onto themselves (implied by (3)). Also, conditions (1) and (2) ensure that M_P is an orientable topological space.

There are other generalisations of maps, hypermaps and pavings, such as constellations, cf. the monograph [LZ04] for more information and references.

One of the basic questions is understanding possible symmetries, or automorphisms (i.e. unrooted self-isomorphisms), of any of the above defined objects. Those can be understood by means of building a one-to-one correspondence between a class of rooted (hyper-)maps $\mathcal{M}$ (or isomorphism classes of maps) on a set of darts D and (usually, torsion-free) subgroups (or their conjugacy classes) of a given single group Δ^+. This correspondence will associate to each map $M \in \mathcal{M}$ a subgroup $H_M \subset \Delta^+$ of index $|D|$. The origins of this technique draw back to the paper by Jones and Singerman [JS78], and have been developed more in the recent works by Breda, Mednykh and Nedela [BdMN10], Mednykh and Nedela [MN06, MN10] for the purpose of solving Tutte's problem of (hyper-)map classification, cf. also [CK17, BCK17].

Let us consider the case of maps, as described in [BdMN10, JS78]. Namely, the rooted maps on n darts (where the root is always supposed to be marked 1) are in a one-to-one correspondence with index n free subgroups of $\Delta^+ = \mathbb{Z} * \mathbb{Z}_2$. Indeed, each free subgroup $H < \Delta^+$ of index n produces a set of cosets $D = \Delta^+/H$ of cardinality n, which can be considered as a set of darts. The root dart here is the identity coset. A subgroup of Δ^+ is torsion-free if and only if it is free, as a consequence of Kurosh's theorem. Thus $\Delta^+ = \mathbb{Z} * \mathbb{Z}_2 \cong \langle\sigma\rangle * \langle\alpha\rangle$ acts on D transitively, and its generators σ and α give rise to permutations R and L acting transitively on D. Thus, we obtain a map M_H corresponding to a free subgroup $H < \Delta^+$. Vice versa, given a map $(D; R, L)$, we have a homomorphism $S : \Delta^+ \to \langle R, L\rangle$ by setting $S(\sigma) = R$, $S(\alpha) = L$. The homomorphism S defines an action of Δ^+ on D, and the subgroup corresponding to M is $H_M = Stab(1) < \Delta^+$.

The above correspondence between the free subgroups of Δ^+ and rooted maps can be extended to the case of hypermaps (with $\Delta = \mathbb{Z} * \mathbb{Z} \cong F_2$), or the so-called (p, q)-hypermaps (with $\Delta^+ = \mathbb{Z}_p * \mathbb{Z}_q$ [CK17]), or 3-dimensional maps (also called pavings [AK89, Spe91], with $\Delta^+ = \mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2$ [BCK17]).

The isomorphisms classes of all aforementioned objects correspond to the conjugacy classes of free subgroups of Δ^+ [JS78, Theorem 3.7]. The symmetries (i.e. unrooted self-isomorphisms) of a (hyper-)map M corresponding to a subgroup $H < \Delta^+$ form a group isomorphic to $N(H)/H$, where $N(H) = \{g \in \Delta^+ \,|\, gHg^{-1} = H\}$ is the normaliser of H in Δ^+ [JS78, Theorem 3.8].

In the sequel we shall study a more abstract question, namely the property of free products of cyclic groups being "telescopic", cf. Definition 1.2.1. Such a free product T being telescopic allows us to realise any finite group Γ as the "symmetry group" $N(H)/H$ of a suitable finite-index subgroup $H \leq T$. Thus, one of our main results is the following statement, cf. Theorem 1.3.7.

Theorem (Theorem 1.3.7)**.** *Any free product of at least two non-trivial cyclic groups is freely telescopic, except for the infinite dihedral group* $D_\infty \cong \mathbb{Z}_2 * \mathbb{Z}_2$.

If we allow the index of H to be sufficiently large, depending on the cardinality of Γ, then a great deal of same index subgroups H with $N(H)/H \cong \Gamma$ can be obtained, cf. Theorem 1.6.1. More precisely, the following holds.

Theorem (Theorem 1.6.1)**.** *Let T be a finite free product of cyclic groups, different from $\mathbb{Z}_2 * \mathbb{Z}_2$. Then for any finite group Γ, there exist constants $A > 1$, $B > 0$ and $M \in \mathbb{N}$ such that for all $d \geq M$ the set $F(T, \Gamma, d) = \{$free subgroups $H \leq T$ of index $\leq d$ with $N_T(H)/H \cong \Gamma$, up to conjugacy$\}$ has cardinality $\geq A^{Bd \log d}$.*

Finally, we translate our group-theoretic statements into the combinatorial language of (hyper-) maps and pavings, cf. Theorems 1.7.1–1.7.3. Such a transition from combinatorics to groups, to combinatorics again is an integral part of our approach. First, we want to obtain some information about symmetries of a sufficiently complicated combinatorial object. Next, we translate our questions about symmetries into a question about the existence of (torsion-free) subgroups of a free product of cyclic groups with some condition on their normalizers. This condition is formulated in terms of combinatorial automorphisms of the subgroup's Schreier graph, by analogy to the approach introduced in [KM02, Wel17]. The symmetries of the corresponding Schreier graphs appear more amenable to combinatorial analysis, which finally provides us with the desired results both in group-theoretic and combinatorial terms.

Remark 1.1.1. Soon after a draft of this paper appeared on the arXiv, the authors were notified by Gareth A. Jones that his paper [Jon18] contains similar results for a wider class of groups. In particular, by [Jon18, Theorem 3], all hyperbolic (extended) triangle groups are shown to be "finitely abundant", which is equivalent to being telescopic for the non-compact ones among them. Also, all subgroups produced in [Jon18, Theorem 3] are, in fact, torsion-free. The methods used in [Jon18] and in our approach differ substantially, as well as the emphasis in our work on the quantitative aspects, such as counting of combinatorial objects with given symmetries.

1.2 Preliminaries

We first establish the necessary notation and provide some basic definitions. Let G be a group, and H be a subgroup of G. Let $N_G(H) = \{g \in G \;:\; gHg^{-1} = H\}$ denote the normaliser of H in G.

Definition 1.2.1. We say that a group T is *telescopic* if for every finite group Γ there exists a finite-index subgroup $H \leq T$ such that $N_T(H)/H \cong \Gamma$.

Definition 1.2.2. If in the above definition we can always choose H to be a free subgroup of T, we say that T is *freely telescopic*.

Definition 1.2.3. A (di-) graph is a tuple $(V, E, \iota : E \to V, \tau : E \to V)$, where V is the set of *vertices*, E is the set of directed *edges* and ι, resp. τ, assigns to each edge e its initial vertex (or origin) $\iota(e)$, resp. its terminal vertex (or terminus) $\tau(e)$. We shall write $\mathrm{St}_+ v := \{e \in E : \iota(e) = v\}$ and $\mathrm{St}_- v := \{e \in E : \tau(e) = v\}$. A morphism of graphs $\phi : (V_1, E_1, \iota_1, \tau_1) \to (V_2, E_2, \iota_2, \tau_2)$ consists of a pair of maps $\phi_V : V_1 \to V_2$ and $\phi_E : E_1 \to E_2$ such that $\iota_2 \phi_E = \phi_V \iota_1$ and $\tau_2 \phi_E = \phi_V \tau_2$. If $\mathcal{G}$ is a (di-)graph, then $V\mathcal{G}$ will denote its set of vertices, and $E\mathcal{G}$ will be its set of edges.

A graph is *labelled* by the elements of a set S if a map $\mu : E \to S$ is given. Let us write S-digraph for an S-labelled digraph. A labelled S-digraph $\mathcal{G}$ is called *folded* (cf. [Sta83]) if, for any $v \in V\mathcal{G}$, the restrictions of μ to $\mathrm{St}_+ v$ and $\mathrm{St}_- v$ are injective, and $\mathcal{G}$ is called *regular* if they are bijective. A morphism of S-digraphs is a morphism of digraphs satisfying $\mu_2 \phi_E = \mu_1$.

Definition 1.2.4. Let G be a group generated by a set S, and $H \leq G$ is a subgroup of G. Then the *(right) Schreier graph* $\mathrm{Sch}_{G,S}(H)$ of H is an S-labelled regular graph having vertex set the right cosets of H, and an edge $H \cdot g \xrightarrow{s} H \cdot gs$ from $H \cdot g$ to $H \cdot gs$, labelled s, for each element $s \in S$ and each coset $H \cdot g$.

Then the Cayley graph $\mathrm{Cay}(G, S)$ of G is defined as the Schreier graph of the trivial subgroup $\{\mathrm{id}\}$ of G. Both Cayley and Schreier graphs are generally considered with a basepoint: the coset $H \cdot e = H$. When graphs with basepoints are considered, their morphisms are assumed to send basepoints to basepoints.

If a graph $\mathcal{G}$ has label set $S \subset G$, for some group G, and a vertex $v \in V\mathcal{G}$ is specified, then the set

$$L(\mathcal{G}, v) := \{\text{labels of loops at } v, \text{ evaluated in } G\}$$

forms a group, called the language of $\mathcal{G}$ at v (an empty loop gives the identity of G, paths concatenation corresponds to taking products, and reversing paths corresponds to taking inverses). Let $\mathrm{ev} : \{\text{words in } S \sqcup S^{-1}\} \to G$ denote the evaluation map, so that $L(\mathcal{G}, v) = \mathrm{ev} \circ \mu\,(\text{"loops at } v\text{"})$.

We also present a few key results that we make use of, most of which can be found in [KM02, Sta83, Wel17]. First, recall Kurosh's theorem.

Theorem 1.2.5 (Kurosh's Subgroup Theorem)**.** *Let $G_1, \ldots, G_n$ be groups, and $G := *_{i=1}^{n} G_i$ be their free product. Then, any subgroup $H \leq G$ of G has the form:*

$$(*_{i=1}^{n} *_{j=1}^{m_i} w_{ij} H_{ij} w_{ij}^{-1}) * F(X)$$

where each H_{ij} is a subgroup of G_i, $F(X)$ is a free subgroup generated by a subset X of G, and w_{ij} is an element of G, for $1 \leq i \leq n$ and $1 \leq j \leq m_i$.

Proof. See, for instance, the monograph [Ser03, pp. 56 – 57]. □

Lemma 1.2.6 ([KM02, Lemma 7.5])**.** *If $\mathcal{G}$ is an S-digraph with S a subset of a group G, then for any vertices v_1, v_2 connected by a path $p : v_1 \rightsquigarrow v_2$ with $g := \mathrm{ev} \circ \mu(p)$:*

$$L(\mathcal{G}, v_1) = gL(\mathcal{G}, v_2)g^{-1}.$$

Proof. If l is a loop at v_2, then plp^{-1} is a loop at v_1, and $\mathrm{ev}\,\mu(plp^{-1}) = g\,\mathrm{ev}\,\mu(l)g^{-1}$, so that the conjugation by g maps elements of $L(\mathcal{G}, v_2)$ to elements of $L(\mathcal{G}, v_1)$, and thus we have the inclusion $gL(\mathcal{G}, v_2)g^{-1} \subseteq L(\mathcal{G}, v_1)$. Symmetrically, $g^{-1}L(\mathcal{G}, v_1)g \subseteq L(\mathcal{G}, v_2)$, and the result follows. □

Lemma 1.2.7 ([KM02, Lemma 4.2])**.** *If $\mathcal{A}, \mathcal{B}$ are S-digraphs and $\mathcal{B}$ is folded, then for any vertices $v \in V\mathcal{A}$ and $u \in V\mathcal{B}$, there exists at most one morphism of S-digraphs $\phi : \mathcal{A} \to \mathcal{B}$ satisfying $\phi(v) = u$.*

Proof. Follow paths; cf. [KM02, Lemma 4.2] or [Sta83, 5.1 (c)]. □

Let $\mathbb{Z}_{p_i} = \langle s_i \mid s_i^{p_i} \rangle$ denote the cyclic group of order $p_i \in \mathbb{N} \cup \{\infty\}$, for $p_i \geq 2$, while setting $p_i = \infty$ yields $\mathbb{Z}$ (which will be our standard notation for the rest of the paper).

Lemma 1.2.8 ([Sto78, Theorem 1.2])**.** *Let $G := \mathbb{Z}_{p_1} * \cdots * \mathbb{Z}_{p_n}$ be a free product of cyclic groups with generators $S = \{s_1, \ldots, s_n\}$, assuming $\mathbb{Z}_{p_i} = \langle s_i \rangle$. There is a bijection between the sets*

$A :=$ *"subgroups of G" and*

$B :=$ *"connected, regular S-digraphs with a basepoint, such that for any s_i, the edges labelled by s_i form cycles of length dividing $p_i < \infty$, up to isomorphism".*

Moreover, free (equivalently, torsion-free) subgroups correspond to graphs with s_i-labelled cycles of length exactly $p_i < \infty$, and under this equivalence, the index of a subgroup equals the number of vertices in the corresponding graph.

If, in a labelled digraph as in the lemma above, a cycle labelled by s_i has length a proper divisor of $p_i < \infty$, such a cycle will be called degenerate, following [Sto78]. An element of B is a (G, S)-Schreier graph, which is called *non-degenerate*, if it contains no degenerate cycle.

Sketch of proof. The Schreier graph of the quotient $\mathrm{Sch}_{G,S}(\cdot)$ provides one direction of the equivalence, while the language at the root $L(\cdot)$ proves the other. In order to show the equivalence "free subgroup" ⇔ "no degenerate cycle", one uses the fact that a torsion element in G must be conjugate to an element of one of the factors, and vice versa. □

The equivalence of Lemma 1.2.8 can actually be generalised to arbitrary finitely generated groups if one does not care about freeness: this is done (using a slightly different language) in [Wel17, Theorem 3.5].

Proposition 1.2.9. *Let $\mathcal{S}(H) := \mathrm{Sch}_{G,S}(H)$. Then $\mathrm{Aut}(\mathcal{S}(H)) \cong N_G(H)/H$.*

In the above statement, Aut denotes the group of automorphisms of a labelled digraph without basepoint.

Proof. Let $N := N_G(H)$, and let us consider the following map:

$$\begin{aligned} \Phi : N &\longrightarrow \mathrm{Aut}(\mathcal{S}(H)) \\ n &\longmapsto (\phi_n : Hg \mapsto Hng) \end{aligned}$$

Then, Φ is a well-defined surjective group homomorphism with kernel exactly H.

It is routine to check that Φ is well-defined and is a homomorphism: this fact depends on N being the normaliser of H. Let us verify its surjectivity.

Let ϕ be an element of $\mathrm{Aut}(\mathcal{S}(H))$ and let Hg be the image of H under ϕ. Since ϕ is an automorphism, we have that

$$H = L(\mathcal{S}(H), He) = L(\phi(\mathcal{S}(H)), \phi(He)) = L(\mathcal{S}(H), Hg),$$

Also, we know that $L(\mathcal{S}(H), Hg) = g^{-1}L(\mathcal{S}(H), He)g$, since changing the basepoint changes the language by conjugation, as in Lemma 1.2.6. This implies $H = g^{-1}Hg$, and thus $g \in N$. Since $\mathcal{S}(H)$ is a regular graph, there is a unique morphism sending H to Hg, so that $\phi = \phi_g$. Therefore, Φ is surjective.

Finally, let us verify that $\ker\Phi = H$. Let n be an element of $\ker\Phi$. Then $\phi_n = \mathrm{id}_{\mathcal{S}(H)}$, which implies that $Hn = H$, and $n \in H$. Conversely, for any $h \in H$, ϕ_h is readily seen to be the identity map, since $Hh = H$. Therefore, $\ker\Phi = H$ and the claim follows by the first isomorphism theorem. □

1.3 Free products of cyclic groups

In this section we show that any free product of (non-trivial) cyclic groups with at least two factors is freely telescopic, with the obvious exception of $D_\infty \cong \mathbb{Z}_2 * \mathbb{Z}_2$, the infinite dihedral group.

From now on, let Γ denote a finite group with generating set S. Let T be a finite free product of cyclic groups $\mathbb{Z}_{p_1} * \cdots * \mathbb{Z}_{p_n}$, and X be the natural choice of its generators (one per factor). We always assume that $p_i \geq 2$ and, if $T = \mathbb{Z}_p * \mathbb{Z}_q$, also that $p \geq q \geq 2$, while $p \geq 3$.

We will proceed as follows, in order to prove that any T as above is freely telescopic, or equivalently, that any finite group Γ is isomorphic to a quotient $N_T(H)/H$ for H a free finite-index subgroup of T.

1.3.1 Plan of proof

1. Some algebraic arguments (Lemmas 1.3.5 – 1.3.6) using Kurosh's Subgroup Theorem reduce the problem to free products of the form $T = \mathbb{Z}_p * \mathbb{Z}_q$ ($p \geq 3, q \geq 2$) and $T = \mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2$.

2. Then, Lemma 1.2.8 translates the question of finding a free subgroup H of T into finding a non-degenerate (T, X)-graph $\mathcal{G}$.

3. By Lemma 1.2.9, the condition that $N_T(H)/H$ be isomorphic to Γ is equivalent to the condition that the automorphism group of $\mathcal{G}$ be isomorphic to Γ. Hence, the initial problem effectively reduces to finding a non-degenerate (T, X)-graph with a given automorphism group.

4. Starting with the Cayley graph $\mathrm{Cay}(\Gamma, S)$ of the finite group Γ, we replace its edges and vertices by certain pieces of non-degenerate (T, X)-graphs (defined in Section 1.4.3), so that the automorphism group is preserved, while obtaining a valid Schreier graph for a finite-index free subgroup H of T.

1.3.2 Basic cases

We start first by proving that $\mathbb{Z}_p * \mathbb{Z}_q$, with $p \geq 3$, $q \geq 2$, and $\mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2$ are freely telescopic. These are the "base cases" for the general statement that follows in Theorem 1.3.7.

Proposition 1.3.1. *The free product $\mathbb{Z}_p * \mathbb{Z}_q$ is freely telescopic for any $p \geq 3$ and $q \geq 2$.*

Proposition 1.3.2. *The free product $\mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2$ is freely telescopic.*

The proofs of these two results rely on a LEGO-like construction using pieces of non-degenerate (T, X)-graphs, that we produce below. Once this is done, and the necessary properties of the construction hold, the proofs will follow easily, cf. Section 1.3.2.

Vertex splitting and gluing

Let us consider a Schreier graph $\mathcal{S} = \mathrm{Sch}_{T,X}(H)$ of a subgroup $H \leq T$, where T has generating set X. If $Y \subsetneq X$, a vertex v of $\mathcal{S}$ is *split* along Y if v is replaced by two vertices v_Y, v_{X-Y}, where v_Y keeps the Y-coloured edges of v, and v_{X-Y} keeps its $(X-Y)$-coloured edges, as shown in Figure 1.2. We shall call v_Y a *dangling* Y-coloured vertex. Observe that splitting vertices breaks the X-regularity of the graph. If u_Y and v_{X-Y} are, respectively, Y- and $(X-Y)$-coloured dangling vertices we say that u_Y and v_{X-Y} are *complementary*, and were we to identify them, we would gain regularity back at the newly created vertex. With this idea in mind, dangling vertices are seen as "connection points" for our graphs: an Y-coloured dangling vertex can only be connected to an $(X-Y)$-dangling vertex, and once all dangling vertices of a graph are connected, the resulting graph is X-regular. Let us call the identification of complementary vertices *gluing*.

Finally, if $\mathcal{S}$ is a Schreier graph of a free subgroup of T, it has no degenerate cycles. Since the operations of splitting vertices and gluing complementary ones do not change the lengths of cycles of any given colour, as soon as regularity is gained back by gluing all dangling vertices of some split graph, one gets the Schreier graph of a *free* subgroup once again.

The following is now essentially obvious from the above considerations.

Proposition 1.3.3. *Let us choose $Y \subsetneq X$ and a finite number of Schreier graphs $\mathcal{S}_i = \mathrm{Sch}_{T,X}(H_i)$ of free, finite-index subgroups H_i of T. Consider their disjoint union $\sqcup \mathcal{S}_i$, in which we split a certain number of vertices along Y, and glue them, in complementary pairs, so that the resulting graph is connected. Then we obtain a Schreier graph for a free, finite-index subgroup of T.*

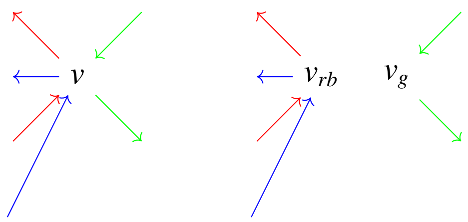


Figure 1.2: Before and after splitting an X-regular graph, with $X = \{$**r**ed, **g**reen, **b**lue$\}$, at a vertex v along $Y = \{r, b\}$.

Sketch of the construction

Below we explain the main idea of the construction. Let T be either $\mathbb{Z}_p * \mathbb{Z}_q$ or $\mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2$, with X its natural set of generators, and Γ be a finite group generated by a set S. Let us choose two X-coloured graphs, say $\mathcal{L}_e$ and $\mathcal{L}_v$ obtained by splitting, respectively, one and two vertices in the Schreier graph of a finite-index free subgroup of T, i.e. a non-degenerate (T, X)-graph. Call $\mathcal{L}_e$ an *edge-link* and $\mathcal{L}_v$ a *vertex-link*. We shall also choose an easily identifiable and *unique*, as we shall see in the sequel, vertex of $\mathcal{L}_e$ to be its *root*, denoted $r(\mathcal{L}_e)$.

We shall connect edge-links by gluing complementary vertices, so as to connect them into chains, and call the result *edge-graphs*. Similarly, we shall connect vertex-links but in a way to produce cycles of them, and call the result *vertex-graphs*. Finally, in the Cayley graph $C := \mathrm{Cay}(\Gamma, S)$ of Γ, we shall replace the edges by edge-graphs and the vertices by vertex-graphs, following the procedure of Section 1.4.3. Our construction will ensure that

(i) the automorphism group Aut C^* of the resulting graph C^* is the same as the automorphism group of C, which is exactly Γ (by Proposition 1.4.1); and

(ii) the graph C^* is actually the Schreier graph of a finite-index free subgroup of T (by Proposition 1.3.3).

From now on, if v is a vertex of a *folded* X-coloured graph, we write $v \cdot x_1 \dots x_n$ for the terminus of the unique path labelled $x_1, \dots, x_n$ and starting at v, if it exists. Then, an equality of the form $v \cdot x_1 \dots x_n = u \cdot y_1 \dots y_n$ holds if and only if both paths in question exist and their termini are equal.

Constructing the links

Our construction differs slightly for $T = \mathbb{Z}_p * \mathbb{Z}_q$ and $\mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2$. In the former case, one uses a relatively generic construction, while the latter is mostly ad-hoc.

Case $T = \mathbb{Z}_p * \mathbb{Z}_q$

Let us fix $p \geq 3$ and $q \geq 2$, let the corresponding generators of each free factor of T be **r**ed and **c**yan and consider q copies of a p-cycle with edge labels r and vertices $v_{0,i}, \dots, v_{p-1,i}$ for each i-th copy, where $0 \leq i \leq q-1$. Next, add two "special" q-cycles labelled c:

$$v_{0,0} \to v_{1,0} \to v_{1,1} \to v_{1,2} \to \dots \to v_{1,q-2} \to v_{0,0},$$

and

$$v_{0,1} \to v_{0,2} \to \dots \to v_{0,q-1} \to v_{1,q-1} \to v_{0,1}.$$

Then, for each $j \geq 2$, draw an extra q-cycle labelled c:

$$v_{j,0} \to v_{j,1} \to \dots \to v_{j,q-1} \to v_{j,0}.$$

In the case $q = 2$, the extra q-cycles have the form

$$v_{0,0} \to v_{1,0} \to v_{0,0} \text{ and } v_{0,q-1} \to v_{1,q-1} \to \dots \to v_{0,q-1}.$$

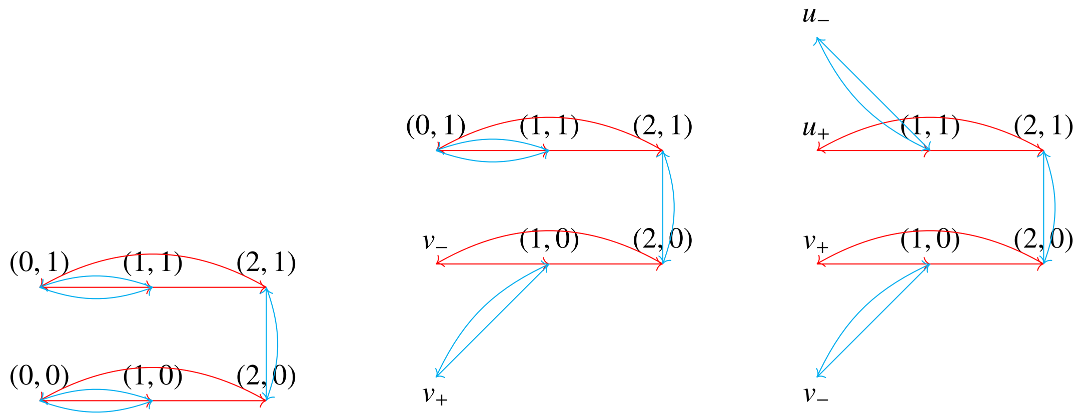
Let $\mathcal{G}_{p,q}$ denote the resulting non-degenerate $(\mathbb{Z}_p * \mathbb{Z}_q, \{r, c\})$-graph.

Now, split the vertex $v_{0,0}$ in order to produce an *edge-link* denoted $\mathcal{L}^e_{p,q}$ and, subsequently, split the vertex $v_{0,q-1}$ to get a *vertex-link* $\mathcal{L}^v_{p,q}$. The vertices obtained by splitting $v_{0,0}$ will be denoted v_+, v_-, and those obtained by splitting $v_{0,q-1}$ will be called u_+, u_-. In the sequel, we swap the assignment of v_+ and v_- vertices for edge- and vertex-links, as shown in Figures 1.3-1.5. This allows us to keep a consistent and clear notation for all associated objects.

In the edge-link $\mathcal{L}^e_{p,q}$, the vertex $v_{0,q-1}$ is unique in the following sense: this is the only vertex $v \in V\mathcal{L}^e_{p,q}$ that satisfies $v \cdot r = v \cdot c$. Observe that the assumption $p \geq 3$ is important here: if $p = q = 2$, then once $v \cdot r = v \cdot c$, the vertex $w = v \cdot r$ also satisfies $w \cdot r = w \cdot c$. Let then $r(\mathcal{L}^e_{p,q}) := v_{0,q-1}$ be called the *root*[1] of $\mathcal{L}^e_{p,q}$.

Some examples of "product graphs" and their splitting at one and two vertices that generate edge- and vertex-link graphs are depicted in Figures 1.3–1.4.

[1]In the case of an edge- or vertex-link the notion of a root is practically opposed to the notion of a root in a graph (map, hypermap, etc.). Indeed, the former is intrinsic to the respective combinatorial structure, while the latter is a matter of choice and can be assigned arbitrarily



(a) The "product" graph $\mathcal{G}_{3,2}$ (b) Splitting one vertex of $\mathcal{L}^e_{3,2}$ (c) Splitting two vertices of $\mathcal{L}^v_{3,2}$

Figure 1.3: The "link" graphs for $T = \mathbb{Z}_3 * \mathbb{Z}_2$: (a) the original graph; (b) the result of splitting $(0,0)$; (c) the result of splitting $(0,0)$ and $(0,1)$. The root of the edge link is $(0,1)$.

Case $T = \mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2$.
Let the generators of the group T be **r**ed, **g**reen and **b**lue. Let $\mathcal{G}_{2,2,2}$ denote the graph (a) in Figure 1.5. In $\mathcal{G}_{2,2,2}$, first split the vertex v_7 to get an edge-link that we call $\mathcal{L}^e_{2,2,2}$, and then split v_0 to get a vertex-link called $\mathcal{L}^v_{2,2,2}$. The vertices obtained by splitting the vertex v_7 will be denoted v_+, v_-, and those obtained by splitting the vertex v_0 will be called u_+, u_-. Observe that the vertices v_0 and v_1 are unique in $\mathcal{L}^e_{2,2,2}$ in the following sense: they are the only vertices $v \in V\mathcal{L}^e_{2,2,2}$ satisfying $v \cdot b = v \cdot g$. Furthermore, one can distinguish v_0 from v_1 as follows: while $v_1 \cdot rgbgrgr = v_1$, it is not the case for v_0. In other words, the path labelled $rgbgrgr$ and starting at v_1 is a loop, but the one identically labelled and starting at v_0 is not. Let then $r(\mathcal{L}^e_{2,2,2}) := v_1$ be the *root* of $\mathcal{L}^e_{2,2,2}$.

Let us write $\mathcal{L}^v$ for any of $\mathcal{L}^v_{p,q}$ and $\mathcal{L}^v_{2,2,2}$, and $\mathcal{L}^e$ for any of $\mathcal{L}^e_{p,q}$ and $\mathcal{L}^e_{2,2,2}$. Since the following constructions do not depend on the exact nature of those graphs, but rather on their abstract properties, this ambiguity is harmless.

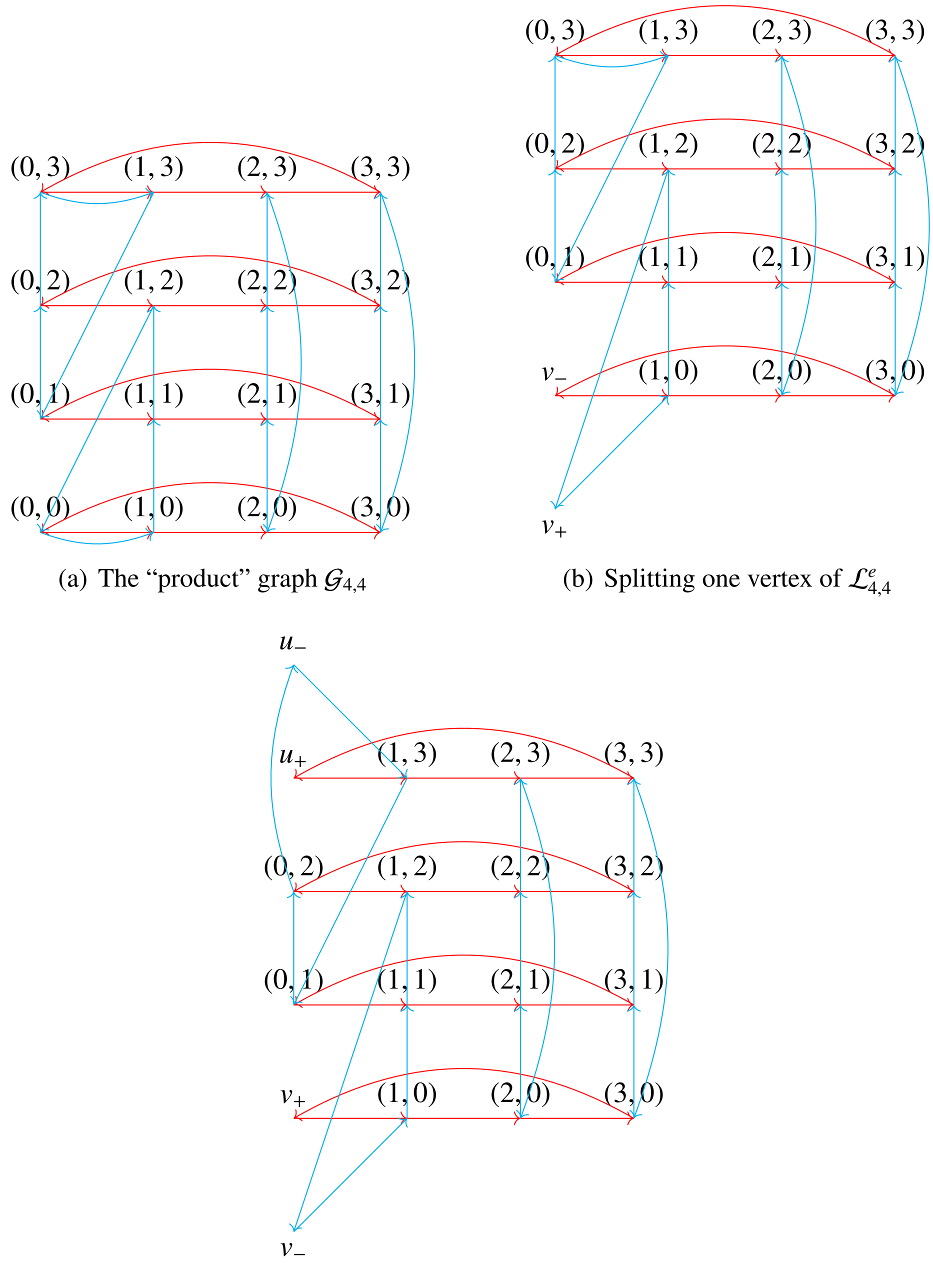
On the behaviour of roots. Let us define the following conditions

$$P_{p,q}(v) := \text{“}v \in V\mathcal{L}^e_{p,q} \text{ satisfies } v \cdot r = v \cdot c\text{”},$$
$$P_{2,2,2}(v) := \text{“}v \in V\mathcal{L}^e_{2,2,2} \text{ satisfies } v \cdot rgbgrgr = v \text{ and } v \cdot b = v \cdot g\text{”},$$

and observe that the following statements hold for $P = P_{p,q}$ or $P_{2,2,2}$, whichever is appropriate.

G.1 No vertex of a vertex-link $\mathcal{L}^v$ satisfies P, and exactly one vertex $r(\mathcal{L}^e)$ of an edge-link $\mathcal{L}^e$ does.

G.2 Gluing edge- and vertex-links together by identifying complementary dangling vertices does not create new vertices satisfying P, as long as the gluing is done on vertices with disjoint neighbourhoods (where the neighbourhood of a vertex v is the set of adjacent vertices.)



(a) The "product" graph $\mathcal{G}_{4,4}$

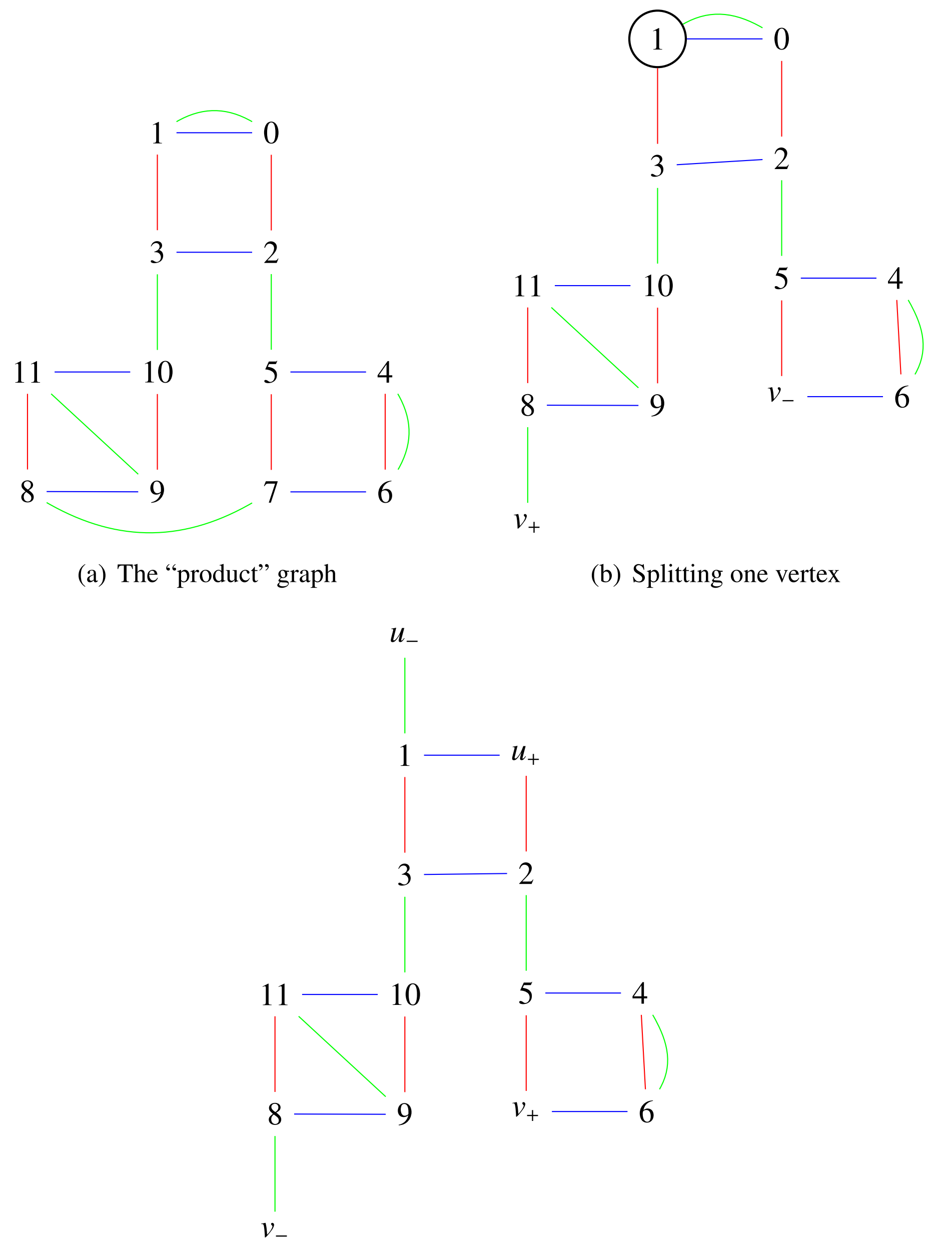
(b) Splitting one vertex of $\mathcal{L}^e_{4,4}$

(c) Splitting two vertices of $\mathcal{L}^v_{4,4}$

Figure 1.4: The "link" graphs for $T = \mathbb{Z}_4 * \mathbb{Z}_4$: (a) the original graph; (b) the result of splitting $(0, 0)$; (c) the result of splitting $(0, 0)$ and $(0, 3)$. The root of the edge link is $(0, 3)$.

G.3 If $\iota : \mathcal{A} \hookrightarrow \mathcal{B}$ is an embedding of folded graphs (in our case $\mathcal{A}$ and $\mathcal{B}$ will be obtained by gluing vertex- and edge-links), then the image of a vertex satisfying P also satisfies P.

Of the above, G.1 holds by construction of the vertex- and edge-links and G.3 is evident from the fact that "following the labels" and "passing to the image under an embedding" are commuting operations. Only G.2 is not as direct. Let u_Y and v_{X-Y} be complementary dangling vertices with disjoint neighbourhoods. Then, letting x be a label in $X - Y$, y be a label in Y, and



(a) The "product" graph

(b) Splitting one vertex

(c) Splitting two vertices

Figure 1.5: The "link" graphs for $T = \mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2$: (a) the original graph; (b) the result of splitting 7; (c) the result of splitting 0 and 7. The root of the edge link is encircled.

w be the result of gluing u_Y to v_{X-Y}, we get $w \cdot x = v_{X-Y} \cdot x$ and $w \cdot y = u_Y \cdot y$, which are distinct by the hypothesis. Since both conditions $P_{p,q}$ and $P_{2,2,2}$ involve equalities of the form $v \cdot x = v \cdot y$, it follows that no glued complementary vertices can satisfy them.

Constructing vertex- and edge-graphs

First of all, let us introduce some necessary notation, which will also be used in Section 1.4.3, later on.

A *vertex-graph* is a graph $\mathcal{V}$, along with an injection $\chi : S \times \{+,-\} \hookrightarrow V\mathcal{V}$, an example of which is depicted in Figure 1.6. Let the *boundary* of a vertex-graph be $\partial\mathcal{V} = \operatorname{im}\chi$.

An *edge-graph* for a label s is a graph $\mathcal{E}_s$, along with two distinguished vertices $h^+(\mathcal{E}_s)$ and $h^-(\mathcal{E}_s)$, as shown in Figure 1.7. We shall provide the general definitions of the graphs $\mathcal{V}$ and $\mathcal{E}_s$ below, which will be case-specific for different choices of T.

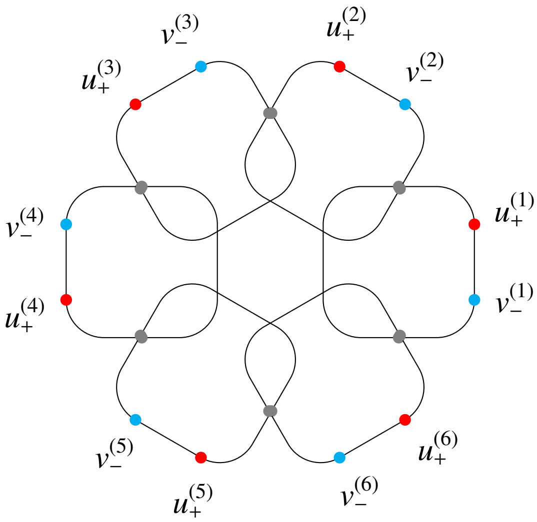


Figure 1.6: The vertex-graph $\mathcal{V}$ for $|S| = 6$. The squares represent the connected vertex-links $\mathcal{L}^v$, the gray dots represent the vertices glued along the way, while the cyan and red dots represent the remaining dangling vertices, which are exactly the images of χ.

In the Cayley graph $C := \operatorname{Cay}(\Gamma, S)$ of Γ, each vertex v will be replaced by an isomorphic copy $\mathcal{V}_v$ of $\mathcal{V}$, and each edge e labelled s will be replaced by an isomorphic copy $\mathcal{E}_e$ of $\mathcal{E}_s$. We shall also make the following identifications:

$$\mathcal{E}_e \ni h^+(\mathcal{E}_s) \sim \chi(s,+) \in \mathcal{V}_u, \qquad \mathcal{E}_e \ni h^-(\mathcal{E}_s) \sim \chi(s,-) \in \mathcal{V}_v,$$

whenever e has label s, origin u and terminus v, where $h^\pm(\mathcal{E}_s) \in \mathcal{E}_e$ means a copy of the vertex $h^\pm(\mathcal{E}_s)$ inside $\mathcal{E}_e$, and similarly for other instances of vertex- and edge-graphs.

Let C^* denote the resulting graph, and let $\iota_e : \mathcal{E}_s \hookrightarrow C^*$ and $\iota_v : \mathcal{V} \hookrightarrow C^*$ be the embeddings corresponding to an edge $e \in VC$ (with $\mu(e) = s$) and a vertex $v \in VC$, respectively. The image $\iota_e(\mathcal{E}_s)$ will be denoted by $\mathcal{E}_e$ and the image $\iota_v(\mathcal{V})$ will be called $\mathcal{V}_v$. Let $h^\pm(e)$ be the image of $h^\pm(\mathcal{E}_s)$ under ι_e.

Finally, if an ordering on the labels $s_1 < \cdots < s_n$ is given, then let

$$C^*_{(i)} := \left(\bigcup_{v \in VC} \mathcal{V}_v \right) \cup \left(\bigcup_{j \le i} \bigcup_{\substack{e \in EC \\ \text{s.t.} \mu(e) = s_j}} \mathcal{E}_e \right),$$

which means that $C^*_{(i)}$ is the subgraph of C^* consisting of the vertex-graphs and *only* the edge-graphs corresponding to the edges with labels s_j for $j \le i$.

From now on, assume that the generating set $S = \{s_1, \dots, s_n\}$ is ordered: $s_1 < s_2 < \cdots < s_n$. First, we consider the vertex- and edge-graphs that we use in the case $T = \mathbb{Z}_p * \mathbb{Z}_q$, with $p \ge q$, $p \ge 3$, $q \ge 2$.

Vertex-graphs. Let us take $|S|$ disjoint copies of the vertex-link $\mathcal{L}^v$ from Section 1.3.2, and call them $\mathcal{L}_v^{(0)}, \ldots, \mathcal{L}_v^{(|S|-1)}$, respectively. Observe that each graph $\mathcal{L}_v^{(i)}$ has four dangling vertices $v_+^{(i)}, v_-^{(i)}$, $u_+^{(i)}, u_-^{(i)}$. In order to create the vertex graph $\mathcal{V}$, identify each $v_+^{(i)}$ to $u_-^{((i+1) \mod |S|)}$, for $0 \le i \le |S|-1$. Let us also define $\chi : S \times \{+,-\} \hookrightarrow V\mathcal{V}$ as $\chi((s_i,+)) = u_+^{(i)}$ and $\chi((s_i,-)) = v_-^{(i)}$. For a sketch of the resulting graph, see Figure 1.6.

Note that in the case $|S| = 1$, only one vertex-link is used, and the fact that $v_+^{(0)}$ is glued to $u_-^{(0)}$, and not to $v_-^{(0)}$, ensures that the hypothesis G.2 holds, so that no extra root appears.

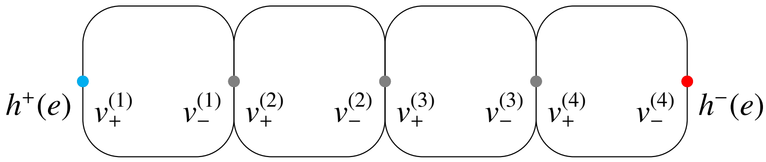


Figure 1.7: An edge-graph of length 4. The squares represent the connected edge-links $\mathcal{L}^e$, and the cyan and red dots represent the remaining dangling vertices, which are exactly $h^+(e)$ and $h^-(e)$.

Edge-graphs. For each $s_i \in S$, take i copies of the edge-link $\mathcal{L}^e$ from Section 1.3.2, and call them $\mathcal{L}_e^{(1)}, \ldots, \mathcal{L}_e^{(i)}$, respectively. Observe that each graph $\mathcal{L}_e^{(j)}$ has two dangling vertices $v_+^{(j)}, v_-^{(j)}$. Now, glue each $v_-^{(j)}$ to $v_+^{(j+1)}$, for $1 \le j \le i-1$. Let $\mathcal{E}_i$ denote the resulting edge-graph, a sketch of which is depicted in Figure 1.7. The vertex $v_+^{(1)}$ will be referred to as $h^+(e)$ and $v_-^{(i)}$ as $h^-(e)$.

Note that the exact same construction works for $T = \mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2$, when using the corresponding objects, i.e. when $\mathcal{L}_{p,q}^v$ is replaced by $\mathcal{L}_{2,2,2}^v$ and $\mathcal{L}_{p,q}^e$ is replaced by $\mathcal{L}_{2,2,2}^e$.

In both cases, letting P be the appropriate condition, i.e. either $P_{p,q}$ or $P_{2,2,2}$, each edge graph $\mathcal{E}_i$ contains exactly i vertices satisfying P: they are exactly the copies of the root vertex $r(\mathcal{L}^e)$ satisfying P in the edge-link $\mathcal{L}^e$. Let $r_1(s_i), \ldots, r_i(s_i)$, where $r_j(s_i) \in \mathcal{L}_e^{(j)}$, denote those roots. Observe that the vertex-graphs do not contain any vertex satisfying P and that $\mathrm{Aut}(\mathcal{E}_i) = \{\mathrm{id}\}$, where the former follows by construction and the latter follows from foldedness.

Substitution

Let $C = \mathrm{Cay}(\Gamma, S)$ be the Cayley graph of a finite group Γ with respect to a generating set $S = \{s_1, s_2, \ldots, s_n\}$. Our goal is to prove that $\mathrm{Aut}(C^*) \cong \mathrm{Aut}(C)$ which we can do by using the argument of Proposition 1.4.1 (cf. Section 1.4 for its proof). Thus, we only have to show that the following conditions hold:

S.1 $\{\phi \in \mathrm{Aut}\,\mathcal{V} : \phi|_{\partial\mathcal{V}} = \mathrm{id}_{\partial\mathcal{V}}\} = \{\mathrm{id}_{\mathcal{V}}\}$;

S.2 one can order the labels $s_1, \ldots, s_n$ in such a way that for any $i \in \{0, \ldots, n\}$, $C^*_{(i)}$ contains no subgraph isomorphic to $\mathcal{E}_{s_i}$ except for the subgraphs $\mathcal{E}_e$, for $e \in EC, \mu(e) = s_i$.

In the edge-link $\mathcal{L}^e$, choose a shortest path from v_+ to $r(\mathcal{L}^e)$ and from $r(\mathcal{L}^e)$ to v_-, and let w_1 and w_2 be their respective labels.

S.1 Recall that $\mathcal{V}$ is a folded graph. Let $\phi \in \mathrm{Aut}\,\mathcal{V}$ be an automorphism with $\phi|_{\partial\mathcal{V}} = \mathrm{id}_{\partial\mathcal{V}}$, and let $v \in \partial\mathcal{V}$ be a vertex in the boundary of $\mathcal{V}$. Then, since ϕ and $\mathrm{id}_{\mathcal{V}}$ agree at v, and $\mathcal{V}$ is folded, we have $\phi = \mathrm{id}_{\mathcal{V}}$. Thus **S.1** holds.

S.2 Fix $i \in \{0, \ldots, n\}$ and consider $C^*_{(i)}$. Our goal is to show that $C^*_{(i)}$ contains no other copies of $\mathcal{E}_i$ than those of the form $\mathcal{E}_e$, for e an edge with label $\mu(e) = s_i$. Let $r := r_1(s_i)$ be the root of the first edge-link in $\mathcal{E}_i$. Suppose that there exists an embedding $\iota : \mathcal{E}_i \hookrightarrow C^*_{(i)}$, and let us show that $\iota(\mathcal{E}_i)$ is equal to $\iota_e(\mathcal{E}_i)$ for some e with label $\mu(e) = s_i$.

We know that $\iota(r)$ lies in an edge-graph because $\iota(r)$ satisfies property P by G.3 and no vertex of a vertex-graph satisfies property P. Let $e \in \mathcal{E}_e$ be the edge such that $\iota(r)$ lies in $\mathcal{E}_e$, and $s_k := \mu(e)$. It is essential to notice that in $\mathcal{E}_i$, for any $1 \leq j \leq i-1$, we have $r_j(s_i)w_2w_1 = r_{j+1}(s_i)$. Similarly, in $\mathcal{E}_e$, for any $1 \leq j \leq k-1$, it holds that $r_j(e)w_2w_1 = r_{j+1}(e)$. A sketch of how the roots and boundary vertices are arranged within an edge graph is given in Figure 1.8.

We shall show that e has label s_i and $\iota(r) = r_1(e)$. This is enough to conclude that $\iota(\mathcal{E}_i)$ equals $\iota_e(\mathcal{E}_i)$, since, by assumption, the embeddings ι and ι_e agree at a vertex $r \in \mathcal{E}_i$, and thus must coincide by foldedness.

To this end, assume that e has label s_k with $k \leq i$, and $\iota(r) = r_j(e)$, necessarily with $1 \leq j \leq k$. Then, we prove that $k - j \geq i - 1$, which is equivalent, by using the constraints $k \leq i$ and $1 \leq j \leq k$, to $k = i$ and $j = 1$.

Let us assume, contrary to the above, that $k - j < i - 1$. Following the path labelled $(w_2w_1)^{k-j+1}$ and starting at $r = r_1(s_i)$ in $\mathcal{E}_i$, we reach the vertex $r_{k-j+2}(s_i)$ in $\mathcal{E}_i$, which is a root. We claim that $r_j(e)(w_2w_1)^{k-j+1}$ is *not* a root, which thus contradicts ι being an embedding, since $r_j(e)(w_2w_1)^{k-j+1} = \iota(r_{k-j+2}(s_i))$ is the image of a root.

First, let us consider the vertex $r_j(e)(w_2w_1)^{k-j} = r_k(e)$. This is the last root of $\mathcal{E}_e$, so that $r_k(e)w_2$ is in a vertex-graph: in fact, $r_k(e)w_2 = h^-(e)$. Now, if $\widehat{r} := r_j(e)(w_2w_1)^{k-j+1} = r_k(e)(w_2w_1)$ were a root somewhere in $C^*_{(i)}$, then $r_k(e)w_2 = \widehat{r}w_1^{-1}$ would either be in the interior of an edge-graph, or of the form $h^+(e')$ for some edge e'. Indeed, if $\widehat{r}$ were the first root $r_1(e')$ of some $\mathcal{E}_{e'}$, then $r_1(e')w_1^{-1} = h^+(e')$ would hold. If it were a subsequent root, then $\widehat{r}w_1^{-1}$ would lie in the interior of $\mathcal{E}_{e'}$. In neither case can it be equal to $r_k(e)w_2 = h^-(e)$, and the desired contradiction is reached.

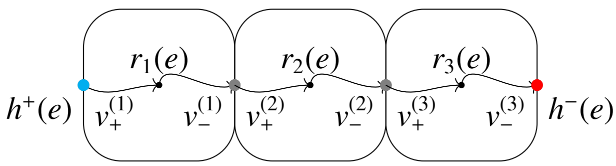


Figure 1.8: Edge-links assembled into an edge-graph, their respective roots and boundary vertices.

Free telescopicity of $\mathbb{Z}_p * \mathbb{Z}_q$ and $\mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2$

Proof of Propositions 1.3.1 and 1.3.2. Given T either $\mathbb{Z}_p * \mathbb{Z}_q$ ($p \geq q$, $p \geq 3$, $q \geq 2$) or $\mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2$ with its natural generating set X (one generator for each cyclic factor), and Γ a finite group with generators S, let us substitute the edges and vertices of $C := \mathrm{Cay}(\Gamma, S)$ by the vertex- and edge-graphs as described in Sections 1.3.2 – 1.3.2.

By Proposition 1.3.3, the resulting graph C^* is a non-degenerate (T, X)-graph with automorphism group

$$\mathrm{Aut}\, C^* \cong \mathrm{Aut}\, C \cong \Gamma,$$

as verified in Section 1.3.2.

Let H be a finite-index free subgroup of T with $\mathrm{Sch}_{T,X}(H) \cong C^*$. Then, by Lemma 1.2.9,

$$N_T(H)/H \cong \Gamma.$$

Since the procedure above can be performed for arbitrary Γ, both $\mathbb{Z}_p * \mathbb{Z}_q$ and $\mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2$ are freely telescopic. □

1.3.3 Adding factors

The following lemma allows us to pull back telescopicity by surjections. The only property not guaranteed in this case is freeness.

Lemma 1.3.4. *Let $f : A \longrightarrow B$ be an epimorphism. Then, if B is telescopic, A is also telescopic, although not necessarily freely.*

Proof. Let Γ be a finite group. Since B is telescopic, there exists a finite index subgroup $H \leq B$ such that $\frac{N_B(H)}{H} \cong \Gamma$. Consider $\widetilde{H} := f^{-1}(H) \leq A$, the preimage of H in A. Clearly, $\widetilde{H}$ is a finite index subgroup of A since it is a preimage of a finite index subgroup under an epimorphism.

We have that $N_A(\widetilde{H}) = f^{-1}(N_B(H))$. Indeed, let x be an element of $f^{-1}(N_B(H))$. By definition, this means that $f(x) \in N_B(H)$, and thus $f(x)H = Hf(x)$. By applying f^{-1} to both sides, we obtain $x\widetilde{H} = \widetilde{H}x$. The latter yields that x is an element of $N_A(\widetilde{H})$. The reverse inclusion is analogous.

Thus, ${N_A(\widetilde{H})}/{\widetilde{H}}$ equals ${f^{-1}(N_B(H))}/{f^{-1}(H)}$. It remains to check that the quotient group ${f^{-1}(N_B(H))}/{f^{-1}(H)}$ is isomorphic to ${N_B(H)}/{H} \cong \Gamma$. Let us consider the map $\widetilde{f} : f^{-1}(N_B(H)) \longrightarrow {N_B(H)}/{H}$ defined by $\widetilde{f}(x) = f(x) \cdot H$. Since $\widetilde{f}$ is a surjective homomorphism with kernel $\ker(\widetilde{f}) = f^{-1}(H)$, the desired result follows from the first isomorphism theorem. □

The following lemmas are key to making an inductive step and proving the main result stated as Theorem 1.3.7.

Lemma 1.3.5. *Assume that a product of the form $T * \mathbb{Z}_m$, $m \geq 2$, is freely telescopic. Then $T * \mathbb{Z}_a * \mathbb{Z}_b$ is also freely telescopic, for all $a, b \geq 2$ such that $lcm(a, b) = m$.*

Proof. Let $c := m/a$, $d := m/b$ and consider the following morphisms:

$$\begin{aligned} \iota_a : \mathbb{Z}_a &\to \mathbb{Z}_m, \\ [k]_a &\mapsto [ck]_m \end{aligned}$$

and

$$\begin{aligned} \iota_b : \mathbb{Z}_b &\to \mathbb{Z}_m \\ [k]_b &\mapsto [dk]_m \,. \end{aligned}$$

Since $\iota_a([1]_a) = [d]_m$ and $\iota_b([1]_b) = [c]_m$, and c and d are coprime, any element of $\mathbb{Z}_m$ can be written as a sum of elements in the images of ι_a and ι_b by Bézout's theorem. Let us consider the surjective morphism $\phi : T * \mathbb{Z}_a * \mathbb{Z}_b \to T * \mathbb{Z}_m$ induced by the maps:

$$\iota_a : \mathbb{Z}_a \to \mathbb{Z}_m, \quad \iota_b : \mathbb{Z}_b \to \mathbb{Z}_m, \quad \mathrm{id} : T \to T.$$

Indeed, ι_a, ι_b and id all extend to $T * \mathbb{Z}_m$, while the universal property of the free product yields ϕ.

Since the generator $[1]_m$ of $\mathbb{Z}_m$ and all of T are in the image of ϕ, the latter is surjective. Also, the restrictions of ϕ to each of the subgroups T, $\mathbb{Z}_a$, and $\mathbb{Z}_b$ are injective since they correspond exactly to the post-compositions of id, ι_a, and ι_b, respectively, with their inclusions in $T * \mathbb{Z}_m$.

Now, fix a finite group Γ to be realised as a quotient "normaliser/subgroup" of $T * \mathbb{Z}_a * \mathbb{Z}_b$. We know, by the hypothesis, that there exists a free subgroup $H \leq T * \mathbb{Z}_m$, of finite index, such that $N(H)/H \cong \Gamma$. Since ϕ is surjective, $\phi^{-1}(H)$ is the desired subgroup by Lemma 1.3.4, once we verify that it is free. To this end, assume that $\phi^{-1}(H)$ is not. Then, by Kurosh's theorem, $\phi^{-1}(H)$ contains a conjugate of a non-free subgroup of T, $\mathbb{Z}_a$ or $\mathbb{Z}_b$. In the first case, there exists $R \leq T$ non-free such that $\phi^{-1}(H)$ contains wRw^{-1}, for some w, as a free factor. Then

$$H \geq \phi(wRw^{-1}) = \phi(w)\phi(R)\phi(w^{-1}) = \phi(w)R\phi(w^{-1}),$$

where the third equality stems from the fact that ϕ restricts to "the identity + inclusion" on T, by definition. This implies that H contains a non-free subgroup, which is a contradiction.

Similarly, assume that $\phi^{-1}(H)$ contains a free factor of the form wCw^{-1}, with C a non-free subgroup of $\mathbb{Z}_a$. Then we have

$$H \geq \phi(wCw^{-1}) = \phi(w)\phi(C)\phi(w^{-1}) \cong \phi(w)\iota_a(C)\phi(w^{-1}),$$

where, once again, the third equality stems from the fact that ϕ, when restricted to $\mathbb{Z}_a$, is just ι_a, plus the inclusion of $\mathbb{Z}_m$ in $T * \mathbb{Z}_m$. Since ι_a is injective, and C is not free, then $\iota_a(C)$ is not free either, so that H contains a non-free subgroup, which is again a contradiction. An analogous reasoning applies if we assume that C is a non-free subgroup of $\mathbb{Z}_b$, and the lemma follows. □

Lemma 1.3.6. *Assume that a product of the form $T * \mathbb{Z}_m$, $m \geq 2$, is freely telescopic. Then $T * \mathbb{Z}$ is freely telescopic.*

Proof. As above, consider the morphisms $\mathrm{id} : T \to T$ and $q : \mathbb{Z} \to \mathbb{Z}_m$, together with the induced surjective "composite" morphism

$$\phi : T * \mathbb{Z} \to T * \mathbb{Z}_m.$$

In this case, the injectivity of ϕ on the $\mathbb{Z}$-factor is not required: we already know that any subgroup of $\mathbb{Z}$ is free. □

Theorem 1.3.7. *Any free product of at least two non-trivial cyclic groups is freely telescopic, except for $\mathbb{Z}_2 * \mathbb{Z}_2$.*

Proof. If $T = \mathbb{Z}_2 * \mathbb{Z}_2$, the infinite dihedral group, then the only Schreier graphs associated with finite-index free subgroups of T are cycles, and T cannot be freely telescopic (obviously, it cannot be telescopic at all). If $T = *_{i=1}^{n} \mathbb{Z}_{p_i}$ is a finite free product of non-trivial cyclic groups with $n \geq 2$ and $p_1 \geq 3$ or $n \geq 3$, the proof proceeds by induction on n.

If $n = 2$, then Proposition 1.3.1 yields that $T = \mathbb{Z}_{p_1} * \mathbb{Z}_{p_2}$ is freely telescopic, assuming that $p_1 \geq 3$ and $p_2 \geq 2$, without loss of generality. If $n = 3$, either each of p_1, p_2, p_3 equals 2, in which case Proposition 1.3.2 yields free telescopicity of $\mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2$, or, without loss of generality, we have $p_3 \neq 2$. In the latter case, $\mathbb{Z}_{p_1} * \mathbb{Z}_{\mathrm{lcm}(p_2,p_3)}$ is already freely telescopic, and Lemma 1.3.5 shows that $\mathbb{Z}_{p_1} * \mathbb{Z}_{p_2} * \mathbb{Z}_{p_3}$ is so, as well.

The inductive step towards $n \geq 4$ is made by using Lemma 1.3.5, with p_i's being finite. Setting $p_i = \infty$, for any i, also yields freely telescopic groups, by Lemma 1.3.6. This concludes the proof of the theorem. □

1.4 Graph substitution

In this section we shall always consider directed labelled graphs. Our goal is to define a reasonable condition that allows replacing vertices and edges of a Cayley graph by other "chunks of graphs" in a way that preserves the automorphism group. All edges having a fixed given label will be replaced by the same "chunk" of a suitable graph, for each label, and a similar procedure takes place for all vertices. In what follows, let Γ be a finite group generated by a finite set S, and let $C := \mathrm{Cay}(\Gamma, S)$ be its Cayley graph.

This section is necessary to complete the proof of the argument made in Section 1.3.2 when discussing the substitution of the edges and vertices of the Cayley graph of a finite group Γ by edge-links and vertex-links respectively. At that point, we verified two conditions, S.1 and S.2, but we had not proved yet that those conditions are sufficient to ensure that the automorphism group of the graph is preserved after the substitutions.

1.4.1 Vertex graphs

If v is a vertex in the Cayley graph $C = \mathrm{Cay}(\Gamma, S)$, let $\mathrm{Adj}\ v$ be the set of tuples consisting of edges adjacent to v and their orientations relative to v, i.e. $\mathrm{Adj}\ v = (\mathrm{St}_+\ v)\times\{+\}\sqcup(\mathrm{St}_-\ v)\times\{-\}$. Let $\Sigma := S\times\{+,-\}$, then each set $\mathrm{Adj}\ v$ is naturally in bijection with Σ: one can easily identify the corresponding pairs of edge labels and their orientations relative to v. Let $\tau_v : \mathrm{Adj}\ v \to \Sigma$ be this bijection, to which we shall refer as *the signature* of v.

If $\mathcal{V}$ is a connected graph and $\chi : \Sigma \hookrightarrow V\mathcal{V}$ is an injective map, then we call the pair $(\mathcal{V}, \chi)$ a *signed graph*. Looking ahead, the map χ will tell us how one should connect the edges adjacent to a vertex $v \in VC$ to the vertices of $\mathcal{V}$ when replacing v in C by $\mathcal{V}$. The injectivity of χ also comes useful later.

Let $\partial\mathcal{V} := \mathrm{im}\,\chi$, which we shall call the boundary of $\mathcal{V}$. Fix a signed graph $(\mathcal{V}, \chi)$, and let

$$\mathrm{Aut}(\mathcal{V}, \chi) := \{\phi \in \mathrm{Aut}\,\mathcal{V} \ : \ \phi|_{\partial\mathcal{V}} = \mathrm{id}_{\partial\mathcal{V}}\}$$

be the group of automorphisms of $\mathcal{V}$ which restrict to the identity map on the boundary $\partial\mathcal{V}$. We shall call $\mathrm{Aut}(\mathcal{V}, \chi)$ the group of *signed automorphisms* of $(\mathcal{V}, \chi)$.

Now let us consider the following condition, that will play an important role in Section 1.4.4.

S.1 $\mathrm{Aut}(\mathcal{V}, \chi) = \{\mathrm{id}\}$.

In other words, S.1 states that any non-trivial automorphism of $\mathcal{V}$ has to move some vertex of its boundary $\partial\mathcal{V}$.

For a given vertex $v \in VC$, a signed graph $\mathcal{V}$ can be inserted in place of v by connecting each edge $e \in \mathrm{Adj}\ v$ to the vertex $\chi(\tau_v(e)) \in \partial\mathcal{V}$. Broadly speaking, condition S.1 forbids any automorphism local to $\mathcal{V}$ to appear when v is replaced by $\mathcal{V}$.

A graph $\mathcal{V}$ described above will be called the *vertex-graph* associated with v.

1.4.2 Edge graphs

The case of edge substitution is simpler: if s is a label, an *edge-graph* for s is a connected graph $\mathcal{E}_s$ with distinct distinguished vertices h_s^+ and h_s^-. We substitute an edge e labelled s by first removing e, and then identifying the origin of e with h_s^+ and its terminus with h_s^-.

1.4.3 Vertex and edge substitution

Below we describe the complete substitution procedure. Recall that we start by considering the Cayley graph C of a group Γ with respect to generators S, as well as

- a connected signed vertex graph $(\mathcal{V}, \chi)$,
- a connected edge graph $\mathcal{E}_s$, for each edge label s, with two distinct distinguished vertices h_s^+ and h_s^- and trivial automorphism group $\operatorname{Aut}\mathcal{E}_s = \{\mathrm{id}\}$.

Let C' be the result of replacing each vertex v of C with an instance of $\mathcal{V}$ as explained in Section 1.4.1. More precisely, if v is a vertex with signature $\tau_v : \operatorname{Adj} v \to \Sigma$, we remove v, insert a copy of $\mathcal{V}$ and connect each edge $e \in \operatorname{Adj} v$ to the vertex $\chi(\tau_v(e))$ of $\mathcal{V}$.

Now, let C^* be the result of replacing each "old" edge (i.e. an edge that is not in any of the vertex graphs) labelled s in C' by a copy of $\mathcal{E}_s$, with h_s^+ identified with the origin of e, and h_s^- with its terminus, as described in Section 1.4.2.

Here and below, copies of $\mathcal{V}$ will be always called "vertex-graphs", and copies of $\mathcal{E}_s$ will be "edge-graphs".

For $v \in VC$, write $\mathcal{V}_v$ for the instance of the vertex-graph $\mathcal{V}$ inserted in place of v, and let $\iota_v : \mathcal{V} \hookrightarrow C^*$ be the corresponding graph embedding. Similarly, for $e \in EC$ with label $\mu(e) = s$, we shall write $\mathcal{E}_e$ for the instance of the edge-graph labelled s which is inserted in place of e, and $\iota_e : \mathcal{E}_s \hookrightarrow C^*$ will be the corresponding graph embedding.

If e is an edge with origin u, resp. terminus v, and label s, then we identify $\iota_e(h_s^+)$ with $\iota_u(\chi(s,+))$, resp. $\iota_e(h_s^-)$ with $\iota_v(\chi(s,-))$. The vertex that we obtain after such identification is shared between $\mathcal{E}_e$ and $\mathcal{V}_u$, resp. $\mathcal{V}_v$. We shall call these vertices $h^+(e)$ and $h^-(e)$, respectively, in order to distinguish them in C^*.

Observe that our construction implies the following:

- The vertex-graphs are disjoint from each other, since h_s^+ and h_s^- are distinct in any $\mathcal{E}_s$, and so are the edge-graphs, by injectivity of χ.
- The vertices of $\partial\mathcal{V}_v$ are exactly those of the form $h^\pm(e)$ for an edge e adjacent to v in the initial graph C.
- $h^+(e) \in \mathcal{V}_u$ if and only if u is the origin of e, and $h^-(e) \in \mathcal{V}_v$ if and only if v is the terminus of e.
- $\chi^{-1}\iota_u^{-1}(h^+(e)) = (\mu(e), +)$, resp. $\chi^{-1}\iota_u^{-1}(h^-(e)) = (\mu(e), -)$, if u is the origin, resp. the terminus, of e.

Finally, consider an ordering $s_1 < \cdots < s_n$ on the edge labels. Let, for each $i = 0, \ldots, n$, $C^*_{(i)}$ be the subgraph of C^* that consists of the vertex-graphs and only the edge-graphs corresponding to the labels s_j with $j \le i$. In other words, this is a subgraph obtained by removing the "interiors" (i.e. everything but the vertices $h^\pm(e)$) of all the edge-graphs corresponding to the labels s_j for $j > i$.

1.4.4 Main statement

Below we formulate the main statement regarding our graph substitution procedure.

Proposition 1.4.1. *With the notation above, if S.1 holds, as well as if*

- ***S.2*** *one can order the labels $s_1, \dots, s_n$ in such a way that for any $i \in \{0, \dots, n-1\}$, the graph $C^*_{(i)}$ contains no subgraph isomorphic to $\mathcal{E}_{s_i}$ except for the edge-subgraphs $\mathcal{E}_e$, for $e \in EC$ with label $\mu(e) = s_i$,*

then

$$\operatorname{Aut} C^* \cong \operatorname{Aut} C.$$

Here, condition S.2 ensures that under any automorphism of C^* the instances of $\mathcal{V}$, resp. the instances of any $\mathcal{E}_s$, must be sent to each other. As can be understood from the proof, any other condition ensuring this fact can be used instead of S.2. Then, it is enough to define, given an automorphism of C^*, a corresponding automorphism of C by looking at the correspondence

$$\{\text{vertex / edge of } C\} \leftrightarrow \{\text{vertex-graph / edge-graph in } C^*\}.$$

Condition S.1 ensures that not too much liberty is gained by making the aforementioned substitutions.

Proof of Proposition 1.4.1. For $v \in VC$, recall that $\mathcal{V}_v$ stands for the instance of $\mathcal{V}$ inserted in place of v, and $\iota_v : \mathcal{V} \hookrightarrow C^*$ is the corresponding graph embedding. For $e \in EC$, $\mathcal{E}_e$ stands for the instance of the edge-graph with label $\mu(e)$, inserted in place of e, while $\iota_e : \mathcal{E}_{\mu(e)} \hookrightarrow C^*$ is the corresponding embedding.

Constructing a morphism $\mathfrak{B} : \operatorname{Aut} C^* \to \operatorname{Aut} C$. Fix an automorphism $\phi \in \operatorname{Aut} C^*$. We claim that, for any edge $e \in EC$, there exists a unique $e' \in EC$ such that ϕ restricts to an isomorphism $\phi|_{\mathcal{E}_e} : \mathcal{E}_e \to \mathcal{E}_{e'}$ and $\mu(e) = \mu(e')$. From this follows that $\phi : h^+(e) \mapsto h^+(e')$ and $\phi : h^-(e) \mapsto h^-(e')$, since $\mathcal{E}_{\mu(e)}$ has trivial automorphism group.

Let, without loss of generality, $s_1 < \cdots < s_n$ be the ordering on the edge labels required by condition S.2. We shall verify by reverse induction on $1 \le i \le n$ that

- **(∗)** if e has label $\mu(e) = s_i$, then there exists a unique e' such that $\phi|_{\mathcal{E}_e}$ is an isomorphism from $\mathcal{E}_e$ to $\mathcal{E}_{e'}$, and $\mu(e') = s_i$ (as noted above, this implies that we have $\phi|_{\mathcal{E}_e} : h^\pm(e) \mapsto h^\pm(e')$).

If $i = n$, we know by the hypothesis (since condition S.2 is satisfied) that the only subgraphs of $C^* = C^*_{(n)}$ isomorphic to $\mathcal{E}_{s_n}$ are the graphs $\mathcal{E}_e$ with e an edge labelled s_n. Fix such an edge: then $\phi|_{\mathcal{E}_e}$ defines an isomorphism onto its image, which must therefore be of the form $\mathcal{E}_{e'}$, for some e' of label s_n. This provides the induction base.

Now, fix $i \le n-1$ and assume that property (∗) holds for any $i+1 \le j \le n$. Since, for all $i+1 \le j \le n$, ϕ sends instances of $\mathcal{E}_{s_j}$ to instances of $\mathcal{E}_{s_j}$, and their vertices $h^\pm(e)$ to themselves, it restricts to an automorphism of $C^*_{(i)}$. Then, by applying S.2, we know that ϕ must send an instance of $\mathcal{E}_{s_i}$ to another one. Indeed, $\phi|_{\mathcal{E}_{s_i}}$ has range in $C^*_{(i)}$, while S.2 guarantees that its image must then be an instance of $\mathcal{E}_{s_i}$. This proves the induction step and our claim is thus verified.

Now, let us show that for each vertex $v \in VC$, there exists a unique $v' \in VC$ such that ϕ restricts to an isomorphism $\phi|_{\mathcal{V}_v} : \mathcal{V}_v \to \mathcal{V}_{v'}$. Indeed, it follows from the above that ϕ sends edge-graphs to edge-graphs, and boundary vertices (those of the form $h^\pm(e)$) to boundary vertices, hence it restricts to an automorphism of the subgraph of C^* consisting only of the vertex-graphs employed in the construction. The latter is just a disjoint union of all vertex-graphs. Since $\phi(\mathcal{V}_v)$ is connected, it must lie in some $\mathcal{V}_{v'}$, and thus coincide with it.

Let $\psi := \mathfrak{B}(\phi)$ be defined as follows: $\psi(v)$ is the unique v' such that ϕ sends $\mathcal{V}_v$ to $\mathcal{V}_{v'}$, and $\psi(e)$ is the unique e' such that ϕ sends $\mathcal{E}_e$ to $\mathcal{E}_{e'}$.

Observe that ψ is bijective since applying the above construction to ϕ^{-1} yields an inverse map to ψ, and it remains to verify that ψ is actually a morphism of graphs, i.e. ψ preserves adjacency.

Recall that $h^+(e)\in\mathcal{V}_u$ if and only if u is the origin of e. Fix an edge $e\in EC$ with origin $u\in VC$, such that $h^+(e)\in\mathcal{V}_u$. Let $u' := \psi(u)$ and $e' := \psi(e)$. Then

$$h^+(e') = \phi(h^+(e))\in\phi(\mathcal{V}_u) = \mathcal{V}_{u'},$$

and since $h^+(e')\in\mathcal{V}_{u'}$, u' is the origin of e'. Replacing $+$ by $-$ and "origin" by "terminus", we conclude that ψ is indeed a morphism of graphs.

$\mathfrak{B}$ is a homomorphism. Let $\phi, \phi'\in \mathrm{Aut}(C^*)$ be two automorphisms, and suppose that ϕ sends $\mathcal{V}_v$ to $\mathcal{V}_{v'}$, while ϕ' sends $\mathcal{V}_{v'}$ to $\mathcal{V}_{v''}$. Then $\phi'\phi$ sends $\mathcal{V}_v$ to $\mathcal{V}_{v''}$, and $\mathfrak{B}(\phi'\phi)(v) = v'' = \mathfrak{B}(\phi')(\mathfrak{B}(\phi)(v))$. An analogous statement holds for edges, and thus $\mathfrak{B}$ is a morphism of groups.

Injectivity of $\mathfrak{B}$. Let ϕ be an element of $\mathrm{Aut}\, C^*$. If $e, e' \in EC$ are two edges, then $\phi(\mathcal{E}_e) = \mathcal{E}_{e'}$ if and only if $\mathfrak{B}(\phi)(e) = e'$. If $\mathfrak{B}(\phi) = \mathrm{id}$, then we have that $\phi(\mathcal{E}_e) = \mathcal{E}_e$. However, edge graphs have trivial automorphism group, and thus $\phi|_{\mathcal{E}_e} = \mathrm{id}$. Since the graph C^* is folded, we obtain that ϕ is the identity on C^*.

Surjectivity of $\mathfrak{B}$. Let $\psi\in \mathrm{Aut}\, C$. Our goal is to find some $\phi\in \mathrm{Aut}\, C^*$ such that $\mathfrak{B}(\phi) = \psi$. Define

$$\phi|_{\mathcal{E}_e} := \iota_{\psi(e)}\circ(\iota_e|_{\mathcal{E}_e})^{-1}, \qquad \text{for every edge } e\in EC,$$
$$\phi|_{\mathcal{V}_v} := \iota_{\psi(v)}\circ(\iota_v|_{\mathcal{V}_v})^{-1}, \qquad \text{for every vertex } v\in VC.$$

By construction, we see already that, assuming ϕ to be a well-defined automorphism, its image under $\mathfrak{B}$ is ψ, and it remains to verify that the piecewise definitions above actually agree. The only vertices lying in two different subgraphs (one vertex- and one edge-subgraph) are the boundary vertices. Fix an edge e with terminus v and label s, such that $\iota_e(h_s^-) = h^-(e) = \iota_v(\chi(e,-))$. Then

$$\phi|_{\mathcal{E}_e}(h^-(e)) = \iota_{\psi(e)}\iota_e^{-1}\iota_e(h_s^-) = \iota_{\psi(e)}(h_s^-) = h^-(\psi(e)),$$
$$\phi|_{\mathcal{V}_v}(h^-(e)) = \iota_{\psi(v)}\iota_v^{-1}\iota_v(\chi(e,-)) = \iota_{\psi(v)}(\chi(e,-)) = h^-(\psi(e))$$

and $\phi|_{\mathcal{E}_e}(h^-(e)) = \phi|_{\mathcal{V}_v}(h^-(e))$, as desired. Replacing $-$ by $+$ and "terminus" by "origin", we conclude that $\phi|_{\mathcal{E}_e}(h^+(e)) = \phi|_{\mathcal{V}_v}(h^+(e))$, as well.

Finally, observe that ϕ is an automorphism since the construction above applied to ψ^{-1}, the inverse of ψ, yields an inverse of ϕ. □

1.5 Adding more factors

In this section we shall concentrate mostly on properties of telescopic groups, which do not have direct applications to the combinatorial results from Section 1.7, and rather stay on the purely group-theoretic side of our study.

Proposition 1.5.1. *Let G_1 be a freely telescopic group and let G_2 be a group having a finite index subgroup H such that $N_{G_2}(H) = H$. Then $G_1 * G_2$ is freely telescopic.*

Example 1.5.2. Let us construct a self normalizing finite index subgroup of $F(a, b)$- the free group on two generators. Consider the subgroup $\tilde{H}$ generated by $(1, 2)$ in $Sym(3)$. It is a self-normalizing subgroup of index 3 in $Sym(3)$ and consider H its lift in $F(a, b)$ by the natural homomorphism $\phi : F(a, b) \to Sym(3)$. Then H is a self-normalizing subgroup of index 3 by the lattice subgroup theorem.

Note that, in particular, the free product of two freely telescopic groups is freely telescopic. The proof follows the same kind of argument as Lemmas 1.3.5 – 1.3.6.

Proof. Let Γ be a finite group. Since G_1 is freely telescopic, there exists a finite-index free subgroup H_1 satisfying ${}^{N_{G_1}(H_1)}/_{H_1} \cong \Gamma$. By assumption, there exists a finite-index free subgroup $H_2 < G_2$ such that ${}^{N_{G_2}(H_2)}/_{H_2}$ is trivial.

Consider the inclusions morphisms $\phi_i : G_i \to G_1 \times G_2$, for $i = 1, 2$, and the morphism $\phi : G_1 * G_2 \to G_1 \times G_2$ induced by the universal mapping property of the free product. The latter is easily seen to be surjective.

Observe that $H_1 \times H_2$ is a finite index subgroup of $G_1 \times G_2$, and that for the normaliser we have $N_{G_1 \times G_2}(H_1 \times H_2)/H_1 \times H_2 \cong N_{G_1}(H_1)/H_1 \times N_{G_2}(H_2)/H_2$, while the latter is simply $N_{G_1}(H_1)/H_1$. Then, let us consider the preimage $\widetilde{H}$ of $H_1 \times H_2$ under the morphism ϕ defined above. Since ϕ is an epimorphism, by Lemma 1.3.4, we have immediately that $G_1 * G_2$ is telescopic. It remains to show that $\widetilde{H} = \phi^{-1}(H_1 \times H_2)$ is a free subgroup of $G_1 * G_2$.

By Kurosh's Subgroup Theorem, $\widetilde{H}$ can be written as $\widetilde{H} = F(X) * (*_i u_i K_{1,i} u_i^{-1}) * (*_j w_j K_{2,j} w_j^{-1})$ for a number of subgroups $K_{1,i}$ of G_1, $K_{2,j}$ of G_2, a subset X of $G_1 * G_2$ and words u_i, w_j in $G_1 * G_2$. If $\widetilde{H}$ were not free, then some of the $K_{1,i}, K_{2,j}$ would not be free either. Assume, without loss of generality, that $K = K_{1,i}$ is not free. Then, by construction, $u_1 K u_1^{-1}$ is a non-free subgroup of $\widetilde{H}$ and thus

$$\phi(u_1 K u_1^{-1}) = \phi(u_1)\phi(K)\phi(u_1)^{-1} = \phi(u_1)\phi_1(K)\phi(u_1)^{-1}$$

is a non-free subgroup of H_1, by using the injectivity of ϕ_1. Therefore H_1 cannot be free, and the proposition follows. □

1.6 An asymptotic estimate

Theorem 1.6.1. *Let T be a finite free product of cyclic groups, different from $\mathbb{Z}_2 * \mathbb{Z}_2$. Then for any finite group Γ, there exist constants $A > 1$, $B > 0$ and $M \in \mathbb{N}$ such that for all $d \geq M$ the set $F(T, \Gamma, d) = \{$free subgroups $H < T$ of index $\leq d$ with $N_T(H)/H \cong \Gamma$, up to conjugacy$\}$ has cardinality $\geq A^{Bd \log d}$.*

The proof of the above theorem is conceptually simple. Fix the aforementioned groups T, Γ and index d, and let $N = N(T, \Gamma, d)$ be an integer whose dependence on T, Γ, d will be clarified later. We shall verify that sufficiently many non-isomorphic graphs $\mathcal{H}_\sigma$ on N vertices can be built, with respect to an additional parameter σ, introduced below. Then the previously used edge-links in the construction of the Schreier graph for $H < T$ (based on a Cayley graph of Γ) will be combined with one of many possible choices of $\mathcal{H}_\sigma$. Once we show that there are

$\geq A^{Bd \log d}$ non-isomorphic instances of $\mathcal{H}_\sigma$, the result follows. On the other hand, the cardinality of $F(T, \Gamma, d)$ is $\leq C^{Dd \log d}$, for some constants $C > 1$, $D > 0$, with $d \geq M$ by [Bol82]. Thus, the *growth type* of the cardinality of $F(T, \Gamma, d)$ is d^d, and is independent of T and Γ.

One natural condition on the graphs $\mathcal{H}_\sigma$ is that they do not contain "roots" that have property P from Section 1.3.2, i.e. vertices v for which

- $v \cdot r = v \cdot c$, if $T = \mathbb{Z}_p * \mathbb{Z}_q$ (with $p \geq q$, $p \geq 3$, $q \geq 2$), or
- $v \cdot rgbgrgr = v$ and $v \cdot b = v \cdot g$, if $T = \mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2$.

If an $\mathcal{H}_\sigma$ contained such a root, then the combination of an edge-link with $\mathcal{H}_\sigma$ would not have a *unique* root any more.

The following lemma provides all the necessary details about constructing $\mathcal{H}_\sigma$'s, which is the first step towards the proof of Theorem 1.6.1.

Lemma 1.6.2. *Let T be $\mathbb{Z}_p * \mathbb{Z}_q$, with $p \geq q$, $p \geq 3$, $q \geq 2$, resp. $\mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2$. Then there exist constants $k, K > 0$, such that for any N a multiple of pq, resp. N a multiple of 8, we have at least KN^{kN} non-isomorphic non-degenerate $(T, \{r, c\})$-graphs, resp. $(T, \{r, g, b\})$-graphs, with two split vertices.*

Proof. First, we treat the case $T = \mathbb{Z}_p * \mathbb{Z}_q$, with $p \geq q$, $p \geq 3$ and $q \geq 2$. Let N be a multiple of pq, and let us consider N vertices ordered as $v_0, \ldots, v_{N-1}$. Similar to Section 1.3.2, we start by drawing red p-cycles of the form

$$v_{kp} \to v_{kp+1} \to \ldots \to v_{kp+p-1} \to v_{kp},$$

One sees that since the initial graph is a disjoint union of red p-cycles, then $\leq 2N/p$ vertices will have to be joined by cyan edges in order to get a connected graph, and then at most q other vertices will be joined in order to close up the last cyan q-cycle. Observe that at least $N - 2N/p - q \geq N/3 - q \geq N/4$ (for N large enough) vertices will remain "free" in $\mathcal{H}_\emptyset$. Let $F \geq N/4$ be the exact number of remaining free vertices, and let $D := \lfloor F \rfloor_q$, where $\lfloor x \rfloor_q$ denotes the greatest multiple of q that is smaller than x, for any natural numbers q and x. Then, for N large enough, we also have $D \geq N/4$.

Now, let us consider any ordering σ on the free vertices in $\mathcal{H}_\emptyset$. By drawing cyan edges between the consecutive vertices in σ, and closing up the cycles when necessary, any such choice of σ yields a non-degenerate graph $\mathcal{H}_\sigma$.

First of all, observe that after fixing a basepoint v_0 in $\mathcal{H}_\emptyset$, the graphs $\mathcal{H}_\sigma$ and $\mathcal{H}_\tau$, for any orderings σ and τ, are basepoint-isomorphic if and only if $\mathcal{H}_\sigma = \mathcal{H}_\tau$. Indeed, since $\mathcal{H}_\emptyset \leq \mathcal{H}_\sigma$, any isomorphism $\mathcal{H}_\sigma \to \mathcal{H}_\tau$ restricts to an embedding $\mathcal{H}_\emptyset \hookrightarrow \mathcal{H}_\tau$, only one of which exists by foldedness: namely, the identity map.

Our next step is estimating the number of non-isomorphic graphs $\mathcal{H}_\sigma$, as above, without a basepoint. To this end, observe that two orderings σ and τ of the remaining free vertices yield the same graph if and only if τ can be obtained from σ by a permutation of the respective "q-blocks" (each consisting of the vertices in a cyan cycle created in accordance with σ or τ), as well as by cyclic permutations within each "block". Indeed, permuting the blocks, together with cyclic permutations inside each block, obviously does not change the resulting graph. Also, if two orderings result in the same graph, then the graph isomorphism implies that the orderings are equal up to the aforementioned permutations. Thus, the number of non-basepoint-isomorphic extensions of $\mathcal{H}_\emptyset$ by using orderings is

$$\geq \frac{D!}{(D/q)!\, q^{D/q}}.$$

However, since the resulting graphs will be used later on in the constructions of "edge-links", we do not want them to contain any double edges. This mean that no vertices v with property $P : v \cdot r = v \cdot c$, except for $v_0 \to v_1$ (joined by a double edge by construction) are allowed, since this would contradict the uniqueness of roots (in the sense of having property P from Section 1.3.2).

We will therefore consider only the orderings that will not produce such edges. Therefore, we start with D free vertices and at each step choose any of the remaining vertices, with the restriction that no double edge be created. Observe that there are three possible cases: the to-be-chosen vertex can be the first of a new cyan cycle, it can lie in the middle of such a cycle or be its last one with respect to the chosen ordering. In the first case, no restriction is imposed since choosing the first vertex of a cycle does not create edges at all. In the second case, one edge is created, and thus one vertex is "illegal" in the sense that choosing it would result in a double edge. In the third case, two vertices are "illegal".

We have to determine if any initial choice of vertices in cyan cycles can be extended to a legal ordering as above. Such an extension may fail under certain circumstances: e.g. if two vertices remain to be ordered, then no legal choice may be possible. In order to circumvent this problem, we shall only consider orderings of $D-q$ vertices, since any legal prefix of length $D-q$ *can* be extended to a legal ordering of length D (if $q = 2$, then one has to consider orderings of the first $D-4$ vertices instead).

With the above points in mind, one readily sees that the number of legal orderings is at least $\geq (D-2)(D-3)\dots(q-1) = (D-2)!/(q-2)!$ Once again, taking into account the fact that orderings yield non-basepoint isomorphic graphs if and only if they are obtained one from another by permutations of the blocks and by cyclic permutations within those blocks, we conclude the existence of

$$\geq \frac{1}{(q-2)!} \cdot \frac{(D-2)!}{(D/q)!\, q^{D/q}} \qquad (*)$$

legal choices of σ defining non-isomorphic graphs $\mathcal{H}_\sigma$ without basepoint.

Now, take any $\mathcal{H}_\sigma$, split its vertex v_0, and call the resulting graph $\mathcal{H}'_\sigma$, while v_r, v_c will be its dangling vertices. We claim that:

1. each $\mathcal{H}'_\sigma$ is connected;
2. if σ and τ are distinct (up to the aforementioned permutations), then $\mathcal{H}'_\sigma$ and $\mathcal{H}'_\tau$ are non-isomorphic without basepoint;
3. $\mathcal{H}'_\sigma$ does not have double edges.

The first property holds since we split at the *unique* double edge of $\mathcal{H}_\sigma$. The second property follows from the fact any isomorphism between $\mathcal{H}'_\sigma$ and $\mathcal{H}'_\tau$ must fix the dangling vertices. Then, were the split graphs $\mathcal{H}'_\sigma$ and $\mathcal{H}'_\tau$ isomorphic, then $\mathcal{H}_\sigma$ and $\mathcal{H}_\tau$ would be so. Finally, since we always split the vertex v_0 of $\mathcal{H}_\sigma$ at the end of its unique double edge, none of the latter remains, and the third property is verified.

Thus we conclude the existence of as many as $(*)$ non-isomorphic graphs $\mathcal{H}'_\sigma$, without a root and having two dangling vertices. By invoking Stirling's formula and using the fact that $D - 2 \geq \frac{D}{q}$ for D big enough, the desired asymptotic estimate follows.

The argument for $T = \mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2$ is similar and not spelled out in detail here. First, we take N a big enough multiple of 8 and then assemble a chain of alternating red-green edges connecting N vertices. Finally, we estimate the number of different graphs obtained by adding blue edges in an analogous way to the above argument. □

Proof of Theorem 1.6.1. Observe that it suffices to verify the statement for $T = \mathbb{Z}_p * \mathbb{Z}_q$, with $p \geq q$, $p \geq 3$, $q \geq 2$, and $T = \mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2$, only. Indeed, for any surjection $f : G \to K$ and any subgroups $K_1, K_2 \leq K$, their pre-images $f^{-1}(K_1)$ and $f^{-1}(K_2)$ are conjugate if and only if K_1 and K_2 are so. Therefore, the inductive constructions of Section 1.3.3 preserve the size of the set in the statement of the theorem.

Assume that T is either $\mathbb{Z}_p * \mathbb{Z}_q$, as above, or $\mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2$ and choose N as in Lemma 1.6.2. Then there are at least KN^{kN} non-isomorphic graphs $\mathcal{H}'_\sigma$ with two dangling vertices.

Let us fix such a graph $\mathcal{H}_\sigma$. Then consider the edge-link $\mathcal{L}_e$ constructed in Section 1.3.2, and glue $\mathcal{H}'_\sigma$ to $\mathcal{L}_e$ via the dangling vertices in order to obtain $\mathcal{L}_e + \mathcal{H}'_\sigma$. The graph $\mathcal{L}_e + \mathcal{H}'_\sigma$ still has a unique root, which is the only condition for the arguments of Section 1.3.2 to hold. Following those constructions once again, for a given finite group Γ, results in a finite-index subgroup $H_\sigma < T$ with $N_T(H_\sigma)/H_\sigma \cong \Gamma$.

For any two choices of graphs $\mathcal{H}'_\sigma$ and $\mathcal{H}'_\tau$ as above, the resulting subgroups H_σ and H_τ are not conjugate. Indeed if H_σ and H_τ were conjugate, then $\mathrm{Sch}_{T,X}(H_\sigma)$ and $\mathrm{Sch}_{T,X}(H_\tau)$ would be isomorphic (with base-points not necessarily matched), so that the roots of the former Schreier graph would be sent to the roots of the latter one, and the instances of $\mathcal{H}'_\sigma$ would be mapped isomorphically to the instances of $\mathcal{H}'_\tau$, which would imply that σ and τ coincide up to the above mentioned permutations. Thus, there are $\geq KN^{kN}$ non-conjugate free subgroups H_σ of T with $N_T(H_\sigma)/H_\sigma \cong \Gamma$ having same index.

Finally, let us estimate the index of H. For any fixed Γ with a given generating set S, let e_Γ and v_Γ be the following numbers:

$$\begin{aligned} e_\Gamma &:= \text{the number of edge-links appearing in the original construction} \\ &\quad \text{of Section 1.3.2} = |\Gamma| \cdot |S| \cdot (|S| - 1)/2; \\ v_\Gamma &:= \text{the number of vertices appearing in the original construction.} \end{aligned}$$

Now, with the modified construction involving $\mathcal{H}'_\sigma$, we obtain that the index of the corresponding subgroup H_σ is

$$d = v_\Gamma + e_\Gamma N,$$

which is the number of vertices in the resulting Schreier graph. Thus, for such a d and N sufficiently large, there are

$$\geq K \left(\frac{d - v_\Gamma}{e_\Gamma} \right)^{k \frac{d - v_\Gamma}{e_\Gamma}}$$

conjugacy classes of index d subgroups $H < T$ such that $N_T(H)/H \cong \Gamma$.

In order to conclude for an arbitrary d, as in the statement, just consider index $\lfloor d \rfloor_{v_\Gamma + e_\Gamma N}$ subgroups. Then the above construction of $\mathcal{H}_\sigma$'s yields the desired result. □

1.7 Symmetries of maps, pavings and constellations

As mentioned in Section 1.1, an oriented map M on n darts can be thought of as index n free subgroup H_M of $\Delta^+ = \mathbb{Z}_2 * \mathbb{Z}$. Moreover, the group of orientation-preserving automorphisms of M is $\mathrm{Aut}\, M \cong N(H_M)/H_M$, where $N(H_M)$ is the normaliser of H_M in Δ^+, cf. [BdMN10] and [JS78, Theorem 3.8].

An analogous statement holds for several other classes of combinatorial objects, such as

- oriented hypermaps, with $\Delta^+ = \mathbb{Z} * \mathbb{Z}$ [MN06, MN10];
- oriented (p, q)-hypermaps, with $\Delta^+ = \mathbb{Z}_p * \mathbb{Z}_q$ ($p \geq 3$, $q \geq 2$) [CK17];
- oriented pavings (or three-dimensional maps), with $\Delta^+ = \mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2$ [BCK17],
- length $k \geq 3$ constellations (in the sense of [LZ04, Definition 1.1.1]), with $\Delta^+ = \underbrace{\mathbb{Z} * \dots * \mathbb{Z}}_{k-1}$.

The following theorem generalises the respective results of [CM82, CM92, Fru38, ŠŠ93] to the case of (p, q)-hypermaps.

Theorem 1.7.1. *For any finite group Γ, and n sufficiently large, there exist $\sim n^n$ non-isomorphic oriented (p, q)-hypermaps H on n darts with* $\mathrm{Aut}\, H \cong \Gamma$.

Here by the symbol $f(x) \sim x^x$, for a function $f(x) : \mathbb{N} \to \mathbb{R}$, we mean its *rate of growth*, i.e. that there exist positive constants A_1, B_1, A_2, B_2 such that $A_1 x^{B_1 x} \leq f(x) \leq A_2 x^{B_2 x}$, for x sufficiently large.

The family of (p, q)-hypermaps naturally comprises the cases of maps ($p = \infty, q = 2$) and hypermaps ($p = q = \infty$). Some other interesting classes of maps, such as triangulations ($p = 3, q = 2$) and quadranqulations ($p = 4, q = 2$) of surfaces, or their bi-coloured triangulations ($p = q = 3$) also satisfy the above theorem.

As well, we can now easily estimate the number of three-dimensional pavings, cf. [AK89, BCK17], with a given automorphism group.

Theorem 1.7.2. *For any finite group Γ, and n sufficiently large, there exist $\sim n^n$ non-isomorphic oriented pavings P on n darts with* $\mathrm{Aut}\, P \cong \Gamma$.

An analogous result holds for constellations as defined in [LZ04, Definition 1.1.1].

Theorem 1.7.3. *For any finite group Γ, any $k \geq 3$, and n sufficiently large, there exist $\sim n^n$ non-isomorphic length k constellations C on n darts with* $\mathrm{Aut}\, C \cong \Gamma$.

Since each length $k \geq 3$ constellation defines a branched covering of the sphere $\mathbb{S}^2$ with k branch points, we can reformulate the above theorem in a more geometric language.

Theorem 1.7.4. *For any finite group Γ, any $k \geq 3$, and n sufficiently large, there exist $\sim n^n$ non-isomorphic degree n branched coverings of $\mathbb{S}^2$ with k branch points and deck transformation group Γ.*

All the above theorems are fairly obvious corollaries of the results in Section 1.6. Indeed, the lower bound on the number of non-isomorphic hypermaps, resp. pavings, on n darts with given automorphism group follows from Theorem 1.6.1, and the upper bound of $(n!)^k$, with an appropriate fixed $k \geq 2$, is trivial. Then, invoking Stirling's formula provides the rate of growth $\sim n^n$.

The statements of Theorems 1.7.1–1.7.4 should be contrasted with [DN12, Lemma 1], which implies that the corresponding combinatorial objects "mostly" have only trivial automorphism groups, i.e. are *asymmetric*. However, those with a given non-trivial automorphism group are still numerous, though not as abundant as asymmetric ones.

Apparently, this technique can be applied to many naturally arising classes of oriented maps, hypermaps, and pavings: exactly those describable as free subgroups of a certain "universal group" Δ^+, which is a finite free product of cyclic groups. For any such class we obtain that any finite group is realisable as automorphism group by infinitely many of its members (more precisely, super-exponentially many depending on the number of darts). It is worth mentioning that not all families of maps admit such a wide variety of symmetries, e.g. the maps with underlying graph a tree [Fei69].

2 Subgroup enumeration of Hecke groups

2.1 Introduction

In this chapter, we give an asymptotic estimate of the number of index n subgroups of Hecke groups, that is groups of the form $\Delta_q = \mathbb{Z}_2 * \mathbb{Z}_q$ and our results apply for $q \geq 3$ prime. The main result of this chapter is the following theorem.

Theorem (Corollary 2.7.9)**.** *The number λ_n of index n subgroups of Δ_q for $q \geq 3$ prime satisfies*

$$\lambda_n \sim \frac{\sqrt{2\pi n} \cdot e^{-1/4}}{2\pi n \sqrt{2q}} n^{n-n/q-n/2} \exp\left(\left(\frac{1}{2} + \frac{1}{q} - 1\right) \cdot n + n^{1/2} + \left(1 + \frac{1}{q} - \frac{1}{q^2}\right) n^{1/q}\right)$$

The results presented in this chapter are not new and different proofs are given in [Lub95]. However, our method is different and uses graph enumeration and exponential generating functions.

In [Sta83], Stallings introduced a simple yet very powerful tool now called a *Stallings graph*. To a finitely generated subgroup of a free group, one can associate its Stallings graph which is a finite labelled graph which encodes the essential information about the subgroup. Stallings graphs are similar to Schreier (coset) graphs but have the advantage of being finite for any finitely generated subgroup of a free group, not only for those of finite index. The Stallings graph is the *core* of the Schreier graph. It has the same information but is easier to access since the graph is finite. With this tool, Stallings proved in simpler ways several already known results and a few new ones. Lots of effort have been made to generalise Stallings graphs to other groups than just free groups.

In this chapter, we use the tools developed in [KMW17] to study subgroups of free products of cyclic groups of the form $\Delta_q = \mathbb{Z}_2 * \mathbb{Z}_q$ for q prime ≥ 3 and associate to any finitely generated subgroup of Δ_q its Stallings graph (cf. Section 2.3). We give information about how the combinatorial properties of the Stallings graph and the algebraic properties of the subgroup are related (cf. Section 2.5). Once this is established, we apply the theory of generating functions to the question of enumerating finite index subgroups by deriving the asymptotics of coefficients of the corresponding generating functions (cf. Section 2.7).

The methods used here provide a more combinatorial approach to determining the number of subgroups of finite index. The tools presented in Section 2.3 and the use of generating functions can be applied not only to finitely generated subgroups and these tools present a unifying method to questions regarding subgroup enumeration, whether the subgroups are of finite index or not, free or with other parameters that could be described combinatorially by means of the Stallings graph.

Note also that the method presented here is from [BNW20]. The method is somehow extended to study not only $PSL_2(\mathbb{Z}) = \mathbb{Z}_2 * \mathbb{Z}_3$ but also any group of the form $\Delta_q = \mathbb{Z}_2 * \mathbb{Z}_q$ for $q \geq 3$ prime. The computations done in [BNW20] are more extensive and not just concerned with finite index subgroups.

2.2 More on Schreier graphs

Since one of our motivation is the study of Stallings graphs and Schreier graphs, we will look at graphs in a broad sense. We allow multi-edges and loops as well as, the less usual, degenerated loops (or semi-edges), which are necessary to treat the case of Schreier graphs of groups with generators of order 2. Heuristically, a degenerated loop can either be seen as an edge with only one end, or as a loop that adds only one to the degree.

Let us start now give a formal definition of (di)graphs.

Definition 2.2.1. A *digraph* (or directed graph) is a couple $(V, \vec{E})$ (the sets of *vertices* and of *arcs*) together with two maps $\iota, \tau \colon \vec{E} \to V$ (*initial vertex* and *terminal vertex*).

A *graph* is a digraph $(V, \vec{E}) = (V, \vec{E}, \iota, \tau)$ together with an involution $\mathrm{inv} \colon \vec{E} \to \vec{E}$ (*inverse arc*) such that $\iota \circ \mathrm{inv} = \tau$ and $\tau \circ \mathrm{inv} = \iota$.

If $(V, \vec{E})$ is a graph, there is a natural equivalence relation on arcs defined by $\vec{e} \sim \vec{f}$ if and only if $\vec{e} = \vec{f}$ or $\vec{e} = \mathrm{inv}(\vec{f})$. The quotient set $\vec{E}/\sim$ is denoted by E and is the set of *edges*. When it's clear from context, we will often write (V, E) for the graph $(V, \vec{E}, \iota, \tau, \mathrm{inv})$.

A vertex v and an edge $e = \{\vec{e}, \mathrm{inv}(\vec{e})\}$ are *adjacent* if $v = \iota(\vec{e})$ or $v = \tau(\vec{e})$, in which case we also say the v is an *end* of e. An edge that is adjacent to only one vertex is called a *loop*. A loop $e = \{\vec{e}, \mathrm{inv}(\vec{e})\}$ is *degenerated* (or a *semi-edge*) if $\vec{e} = \mathrm{inv}(\vec{e})$, that is if $\vec{e}$ is a fixed point of inv.

Remark 2.2.2. Graph theorists usually avoid degenerated loops, but this special kind of edges naturally arises in the context of Schreier graphs. It is particularly necessary for the bijections of Lemma 2.2.5.

In a graph, the *degree* of a vertex is the number of arcs $\vec{e}$ such that $v = \iota(\vec{e})$. In particular, a non-degenerated loop adds 2 to the degree of its adjacent vertex, while a degenerated loop adds only 1. All graphs under consideration will be *locally finite*, that is the degree of any vertex is finite.

If (V, E) is a graph, classical notions like *paths* and *connected components* are defined in a natural way. In particular, a *morphism* of graphs from (V, E) to (W, F) is a map $\varphi \colon V \sqcup \vec{E} \to W \sqcup \vec{F}$ which is compatible with the graph structure. Finally, for a vertex v we define $\mathrm{Star}(v) := \{\vec{e} \in \vec{E} \mid \iota(\vec{e}) = v\}$.

Definition 2.2.3. Let $\mathcal{A}$ be a set. A *labelling* of a digraph $(V, \vec{E})$ by $\mathcal{A}$ is a map from $\vec{E}$ to $\mathcal{A}$. If $\mathcal{A}^{\pm}$ is a set endowed with an involution $^{-1}$, a *labelling* of a graph $(V, \vec{E})$ by $\mathcal{A}^{\pm}$ is a map $l \colon \vec{E} \to \mathcal{A}^{\pm}$ such that $l(\mathrm{inv}(\vec{e})) = l(\vec{e})^{-1}$.

If $\{\vec{e}\}$ is a degenerated loop, then $l(\vec{e})$ is a fixed point of $^{-1}$. The set $\{l(\vec{e}), l(\vec{e})^{-1}\}$ is the *color* of the edge $\{\vec{e}, \mathrm{inv}(\vec{e})\}$. An $\mathcal{A}^{\pm}$-labelled graph is *freely reduced* if no two arcs out of (respectively into) the same vertex have the same label. A *morphism* of $\mathcal{A}^{\pm}$ labelled graphs from (V, E, l) to (W, F, l') is a graph morphism φ such that $l' \circ \varphi = l$.

An *a-edge* of an $\mathcal{A}^{\pm}$-labelled graph X will denote any edge of X with color $\{a, a^{-1}\}$. Similarly, an *a-Y subgraph* of X is a subgraph of X consisting of edges coloured by a and which is isomorphic to the (unlabelled) graph Y.

Finally, let us recall the definition of Schreier graphs from Section 1.2 too .

Definition 2.2.4. Let Γ be a group with a symmetric generating set S, and let Λ be a subgroup of Γ. The corresponding *(right) Schreier graph* $\mathrm{Sch}(\Gamma, \Lambda; S)$ is the labelled graph with vertex set the right cosets $\{\Lambda g \mid g \in \Gamma\}$ and with an arc labelled by s from Λg to Λh for every s in S such that $\Lambda gs = \Lambda h$. If $\vec{e}$ is an arc from Λg to Λh labelled by s, its inverse arc is the only arc from Λh to Λg labelled by s^{-1}.

Any Schreier graph $\mathrm{Sch}(\Gamma, \Lambda; S)$ comes with a canonical root: the vertex Λ. Schreier graphs are generalizations of the well-known *Cayley graphs*, where $\mathrm{Cay}(\Gamma; S) = \mathrm{Sch}(\Gamma, \{1\}; S)$.

In the graph $\mathrm{Sch}(\Gamma, \Lambda; S)$ there is a loop at the vertex Hg if and only if there exists s in $S \cap g^{-1}\Lambda g$. This loop is degenerated if and only if s is of order 2. This is the main reason why we allow degenerated loops in our definition of a graph.

The following fact about Schreier graphs will be fundamental for us. Its proof is elementary and details can be found in [Lee16].

Lemma 2.2.5. *Let Γ be a group and S be a symmetric generating set. Then the following three sets are in bijection:*

1. *Subgroups of Γ,*
2. *Rooted Schreier graphs of Γ with respect to S,*
3. *Rooted S-labelled graphs that are covered by* $\mathrm{Cay}(\Gamma; S)$ *(where the coverings preserve the root and the labels).*

These bijections induce bijections between the following three sets

1. *Conjugacy classes of subgroups of Γ,*
2. *Schreier graphs of Γ with respect to S,*
3. *S-labelled graphs that are covered by* $\mathrm{Cay}(\Gamma; S)$ *(where the coverings preserve the labels).*

2.3 Stallings graphs

In [Sta83], Stallings introduced a remarkable tool to associate with any finitely generated subgroup H of a free group F a finite, directed, labelled graph which is easily computable and roughly contains all the information about the subgroup. In [KMW17], this combinatorial tool is generalized to quasi-convex subgroups of hyperbolic groups (among other generalizations). In this section, we recall some definitions and properties from [KMW17] and from [Kap97] that will be useful later in this chapter.

Let us fix $\mathcal{A}$ an alphabet, G a group and an epimorphism $\mu : \mathcal{A}^* \to G$ which respects inverses where $\mathcal{A}^*$ is the free monoid over $\mathcal{A}^\pm = \mathcal{A} \sqcup \mathcal{A}^{-1}$. In other words, we are choosing a set of generators of G.

Definition 2.3.1 ([KMW17]). A rational language L over $\mathcal{A}^\pm$ is a subset $L \subseteq \mathcal{A}^*$ which is accepted by a finite state automaton. The pair (G, L) is a rational structure of G relative to $\mathcal{A}$ if $L \subseteq \mathcal{A}^*$ is a rational language consisting only of reduced words and such that $\mu(L) = G$.

Note that if G is a hyperbolic group and we take the language L_{geod} of geodesic words, this yields a rational structure for G. For more details on this, see [Eps92].

Definition 2.3.2. [Definition 3.1, [KMW17]] Let (G, L) be a rational structure for the group G and let H be a subgroup of G. The Stallings graph of H with respect to L is the fragment of $\mathrm{Sch}(G, H; \mathcal{A})$ rooted at H and spanned by the loops at H labelled by the L-representatives of the elements of H. We denote it $\Gamma_L(H)$.

Definition 2.3.3 (Definition 3.4, [KMW17])**.** A reduced rooted $\mathcal{A}$-graph (Γ, v) is a Stallings-like graph for H with respect to L if every loop at v in Γ is labelled by a word in $\mu^{-1}(H)$ and every L-representative of an element of H labels a loop of Γ at v.

These two definitions generalize the concept of core-graph introduced by Stallings in [Sta83]. However, the finiteness of such graphs is not guaranteed and is related to the notion of quasi-convexity.

A subset $S \subseteq G$ is called L-quasi-convex if there exists an integer k such that, for every word $w \in L \cap \mu^{-1}(S)$, the path in $\mathrm{Cay}(G; S)$ which starts at 1 and is labelled by the word w lies in the k-neighbourhood of S.

The following proposition gives a sufficient and necessary condition for the Stallings-graph to be finite.

Proposition 2.3.4 (Proposition 3.5, [KMW17])**.** *Let (G, L) be a rational structure for G. Then a subgroup $H \leq G$ of G is L-quasi-convex if and only if it has a finite Stallings-like graph with respect to L, if and only if $\Gamma_L(H)$ is finite.*

The following results will allow us to apply Proposition 2.3.4 to Hecke groups.

Definition 2.3.5. A group G has property (Q) if it is word hyperbolic and any of its finitely generated subgroups is quasi-convex.

Theorem 2.3.6 ([Kap97])**.** *Suppose $G = A *_C B$ is word hyperbolic where C is virtually cyclic and A, B have property (Q). Then G has property (Q).*

Any finite group has property (Q) and is hyperbolic. Thus, by applying Theorem 2.3.6, groups of the form $\Delta_q = \mathbb{Z}_2 * \mathbb{Z}_q$ have property (Q) (more generally, any free product of two cyclic groups has property (Q)).

2.4 Core graphs of Hecke groups

We will now look at the group $\Delta_q := C_2 * C_q$ for q prime, with standard generating set $S = \{a, b, b^{-1}\}$ and $q \geq 3$.

The Cayley graph $\mathrm{Cay}(\Delta_q; S)$ consists of b-q-gons connected together by a-edges. While the graph X is not a tree when $q \geq 3$, it is quasi-isometric to a tree and will play a role similar to the role played by trees in the study of free groups. In this context, (reduced) thick paths will play an important role.

Definition 2.4.1. A *thick path of length n* in the graph $\mathrm{Cay}(\Delta_q; S)$ is a sequence $g_0, e_1, g_1, \ldots, e_n, g_n$ such that the g_i are b-q-gons, the e_i are a-edges and such that for every i, the graph g_i shares a vertex with e_i and a vertex with e_{i+1}.

Observe that a thick path $g_0, e_1, g_1, \ldots, e_n, g_n$ is uniquely determined by the sequence $(e_i)_{i=1}^n$ of its a-edges. Moreover, thick paths of length $n \geq 1$ are in bijection with sequences $(e_i)_{i=1}^n$ of a-edges such that e_i and e_{i+1} are adjacent to the same b-q-gon.

A thick path is *reduced* if it does not contain subsequences of the form $e_i, g_i, e_{i+1} = e_i$. We say that a thick path $(e_i)_{i=1}^n$ is from v to w if v is any vertex of g_0 and w is any vertex of g_n. In particular, a given reduced thick path of length $n \geq 1$ has $2q$ endpoints (q on each sides). The main feature of thick paths is the following elementary fact.

Lemma 2.4.2. *Let v and w be two vertices of* $\mathrm{Cay}(\Delta_q; S)$*. Then there exists a unique reduced thick path from v to w.*

We now define the analogue of rooted subtrees of a regular tree.

Definition 2.4.3. Let Y be the graph obtained from $\mathrm{Cay}(\Delta_q; S)$ by removing all edges of one b-q-gon. Then all connected components of Y are isomorphic and their isomorphism type does not depend on the chosen b-q-gon. An S-labelled graph isomorphic to a connected component of Y is called a Δ_q*-tree*.

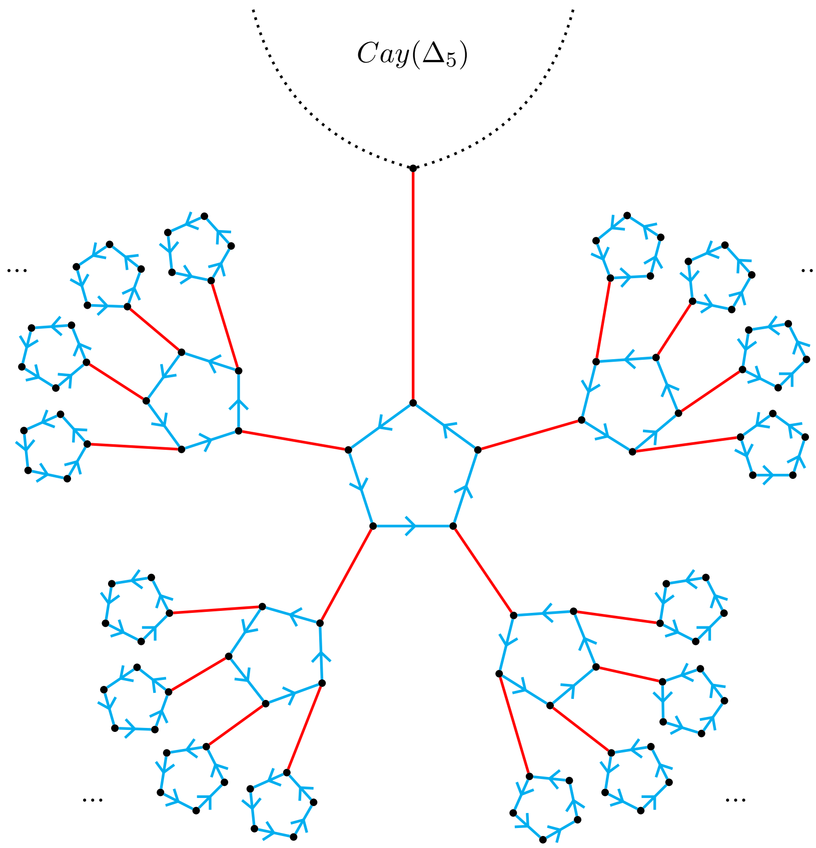


Figure 2.1: Example of a Δ_5-tree. The graph continues repeating the same pattern. The cyan edges are edges labelled b and the red edges are labelled a.

It is an easy observation that if $X = \mathrm{Sch}(\Delta_q, \Lambda; S)$ for some non-trivial Λ, then every Δ_q-tree of X is contained in a unique maximal Δ_q-tree.

We can now define core graphs for subgroups of Δ_q in an analogous way as Stalling defined core graphs for subgroups of free groups.

Definition 2.4.4. Let Λ be a non-trivial subgroup of Δ_q. The corresponding (right) *core graph* $\mathrm{Core}(\Delta_q, \Lambda)$ is the subgraph obtained from $\mathrm{Sch}(\Delta_q, \Lambda; S)$ by cutting all maximal Δ_q-trees. It is naturally labelled by the restriction of the labelling of $\mathrm{Sch}(\Delta_q, \Lambda; S)$. On the other hand, we define $\mathrm{Core}(\Delta_q, \{1\})$ to be a single b-q-gon.

Observe that $\mathrm{Core}(\Delta_q, \Lambda)$ depends only on the conjugacy class of Λ and is always connected. For Λ a non-trivial subgroup of Δ_q and v a vertex of $\mathrm{Sch}(\Delta_q, \Lambda; S)$, there exists a unique shortest reduced thick path (possibly empty) from v to $\mathrm{Core}(\Delta_q, \Lambda)$. This allows us to define a rooted version of the core graph.

Definition 2.4.5. Let Λ be a non-trivial subgroup of Δ_q. The corresponding (right) *pointed core graph* $\mathrm{Core}_p(\Delta_q, \Lambda)$ is the rooted subgraph of $\mathrm{Sch}(\Delta_q, \Lambda; S)$ consisting of the union of $\mathrm{Core}(\Delta_q, \Lambda)$ together with the unique shortest reduced thick path from the vertex Λ to $\mathrm{Core}(\Delta_q, \Lambda)$, it is rooted at Λ.

On the other hand, we put $\mathrm{Core}_p(\Delta_q, \{1\})$ to be a single rooted b-q-gon.

It follows from the definition that core graphs are in bijection with Schreier graphs, and hence with conjugacy classes of subgroups of Δ_q. Similarly, pointed core graphs are in bijection with rooted Schreier graphs and hence with subgroups of Δ_q.

Following Stallings, we will now give an intrinsic definition of a finite core graph.

Definition 2.4.6. A finite S-labelled graph X is a Δ_q-core graph if and only if it satisfies

1. It is connected and freely reduced;
2. Every vertex belongs either to a b-q-gon or to a b-loop;
3. If Y is a b-q-gon of X, then at least one of the following three possibilities happens: Y is not incident to any a-edge, there is an a-loop incident to Y or there are at least two vertices of Y incident to an a-edge.

In other words, condition 3 forbids only the situation where Y has exactly one vertex incident to an a-edge and that this edge is not a loop.

Definition 2.4.4 and Definition 2.4.6 agree in the case of finite graphs. More precisely, we have:

Proposition 2.4.7. *For any subgroup Λ of Δ_q, if the graph $\mathrm{Core}(\Delta_q, \Lambda)$ is finite, it is a Δ_q-core graph in the sense of Definition 2.4.6. On the other hand, if X is a finite Δ_q-core graph in the sense of 2.4.6, then there exists a subgroup Λ of Δ_q such that $X \cong \mathrm{Core}(\Delta_q, \Lambda)$. Moreover, the map $\mathrm{Core}(\Delta_q, \Lambda) \mapsto X$ and $X \mapsto \mathrm{Core}(\Delta_q, \Lambda)$ are inverse one of the other when restricted to finite graphs.*

Proof. Let X be a Δ_q-core graph in the sense of Definition 2.4.6. Then by gluing a Δ_q-tree to every vertex not adjacent to an a-edge or a-loop we retrieve a graph that is covered by $\mathrm{Cay}(\Delta_q; S)$ and thus is the Schreier graph of some subgroup Λ of Δ_q. Conditions 2 and 3 are maximal.

Let now X be a core-graph in the sense of Definition 2.4.4. The graph is freely reduced, connected and every vertex belongs to a b-q-gon or to a b-loop because removing maximal Δ_q-trees does not affect those properties. For condition 3, remark that in a Schreier graph, every vertex is adjacent to an a-edge or a-loop. Thus, if we remove all maximal Δ_q-trees, every b-q-gon is still adjacent to at least two edges labelled a except if it is the Schreier graph.

Finally, the fact that the two maps are inverse of each other from the arguments presented here above. □

The following result explains why we work with core graphs instead of working with Schreier-coset graphs.

Lemma 2.4.8. *Let Λ be a subgroup of Δ_q. Then*

1. *The index of Λ is the number of vertices of* $\mathrm{Sch}(\Delta_q, \Lambda; S)$. *In particular, Λ has finite index if and only if* $\mathrm{Sch}(\Delta_q, \Lambda; S)$ *is finite;*

2. *Λ is finitely generated if and only if* $\mathrm{Core}(\Delta_q, \Lambda)$ *is finite;*

3. *If Λ is finitely generated, then Λ has finite index in Δ_q if and only if* $\mathrm{Sch}(\Delta_q, \Lambda; S) = \mathrm{Core}(\Delta_q, \Lambda)$.

Proof. The first assertion holds in any group, with any generating set, and directly follows from the definition.

For the second assertion, let $v_0$be a vertex of $\mathrm{Core}(\Delta_q, \Lambda)$. The language $L(v_0, \mathrm{Core}(\Delta_q, \Lambda) = \{$labels of loops at v_0 evaluated in $\Delta_q\}$ is equal to Λ by definition of the core-graph. Thus, if the core-graph is finite, the subgroup Λ is finitely generated. On the other hand, the construction of the core-graph coincides with Definition 2.3.2 and thus we can apply Proposition 2.3.4.

We now prove the last assertion. On one hand, if Λ is finitely generated and the equality $\mathrm{Sch}(\Delta_q, \Lambda; S) = \mathrm{Core}(\Delta_q, \Lambda)$ holds, then $\mathrm{Core}(\Delta_q, \Lambda)$ is finite and so is $\mathrm{Sch}(\Delta_q, \Lambda; S)$. This implies that Λ has finite index. On the other hand, the core graph of Λ is equal to its Schreier graph if and only if $\mathrm{Sch}(\Delta_q, \Lambda; S)$ contains no Δ_q-trees. But if Λ has finite index, then $\mathrm{Sch}(\Delta_q, \Lambda; S)$ is finite and we are done. □

Remark 2.4.9. All the results of this subsection generalize easily to $\Delta_F := C_2 * F$ where F is any finite subgroup. In this case, $S = \{a\} \cup T$ where T is some symmetric generating set of F. The role of b-q-gons is then played by copies of $\mathrm{Cay}(F; T)$, while the statement "Every vertex belongs either to a b-q-gon or to a b loop" should be replaced by "Every vertex belongs to a labelled quotient of $\mathrm{Cay}(F; T)$".

The problem when working with a general finite group F, even for F cyclic, is that F might have many subgroups, or equivalently that $\mathrm{Cay}(F; T)$ might have many quotients, which implies that the combinatorial analysis of Δ_F-core graphs become more complicated. This is the main reason why we restrict ourself to the groups Δ_q for q prime. Finally, we note that similar ideas may be used to study groups of the form $G * F$ where both G and F are finite groups but this will not be discussed further in this manuscript.

2.5 Reading properties of a subgroup off its core graph

Core graphs as we defined them are exactly the part of the Schreier graph that encode the information about the subgroup. The aim of this section is to give an easy combinatorial criterion to decide whether a finite S-labelled graph is a core graph and give some properties about the subgroup it corresponds to.

We define the *combinatorial type* of a finite S-labelled graph to be the tuple $(n, l_a, i_a, w_a, l_b, m)$ where:

n is the number of vertices in the graph,

l_a is the number of a-loops,

i_a is the number of a-edges which are not loops, we will call them straight a-edges,

w_a is the number of vertices with a-degree 0,

l_b is the number of b-loops,

m is the number of b-q-gons.

Proposition 2.5.1. *Let $\gamma = (n, l_a, i_a, w_a, l_b, m)$ be a tuple of positive integers with $n \geq 1$. Then γ is the combinatorial type of a Δ_q-core graph if and only if it satisfies the two conditions:*

1. $n = 2i_a + l_a + w_a = q \cdot m + l_b$
2. $(q-2) \cdot m - l_a - l_b - w_a = -2$

Remark 2.5.2. Note that from condition 1, we obtain $n = i_a + \frac{1}{2}(q \cdot m + l_a + l_b + w_a)$ and $i_a = \frac{1}{2}(q \cdot m + l_b - l_a - w_a)$.

Proof. Let γ be the combinatorial type of a Δ_q-core graph X. Then every vertex of X is either on a b-q-gon or on a b-loop and is either on a straight a-edge, on an a-loop or is not adjacent to any edge labelled a. This implies condition 1.

For condition 2, consider a spanning tree of X containing $q-1$ edges of every q-gon. Because X is a Δ_q-core graph, every straight a-edge must be contained in the spanning tree. Thus we can conclude by a simple calculation.

$$
\begin{aligned}
& (q-1) \cdot m + i_a = n - 1 \\
\Leftrightarrow \quad & (q-1) \cdot m + i_a = i_a + \frac{1}{2}(q \cdot m + l_a + l_b + w_a) - 1 \\
\Leftrightarrow \quad & (q-2) \cdot m - l_a - l_b - w_a = -2
\end{aligned}
$$

Now let γ be a tuple of integers satisfying conditions 1 and 2. We will construct Δ_q-core graphs with combinatorial type the given tuple.

First, we will consider the case when $m = 0$. If $m = 0$, we can rewrite the two conditions as

a) $n = 2i_a + l_a + w_a = l_b$

b) $l_a + l_b + w_a = 2$

Since the graph is non-empty, $l_b > 0$. Then by combining conditions a) and b) we obtain $l_b = i_a + 1$.

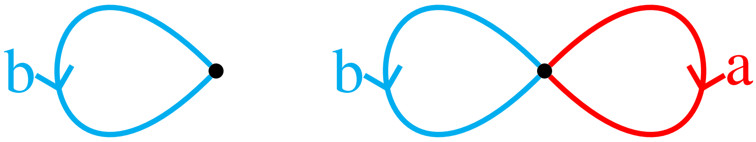
If $l_b = 1$, then $i_a = 0$ and $w_a + l_a = 1$. Depending on the values of w_a and l_a, we can construct two different Δ_q-core graphs as seen in Figure 2.2.

If $l_b = 2$ then $i_a = 1$ and $w_a = l_a = 0$ and we can construct the Δ_q-core graph seen in Figure 2.3.

Definition 2.4.6 guarantees that the graphs constructed above are Δ_q-core graphs with the desired combinatorial type γ.

Now suppose that $m \geq 1$ and consider $Q_1, ..., Q_m$ to be m distinct b-q-gons. Condition 1, implies that $i_a \geq m-1$. Indeed, we can rewrite it as $(q-2)m+2-l_b = l_a+w_a$ thus $m-1+l_b = i_a$. We can thus connect the q-gons using straight a-edges. For an example see Figure 2.4.

Now attach each b-loop to a b-q-gon using a straight a-edge. We obtain a graph with $qm+l_b = n$ vertices from which exactly $l_a + w_a$ have no a-edge adjacent to them. Moreover, the q-gons



(a) The case $w_a = 1$ and $l_a = 0$

(b) The case $w_a = 0$ and $l_a = 1$.

Figure 2.2: The two Δ_q-core graphs in the case $l_b = 1$.

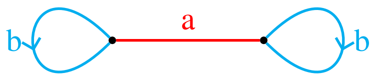


Figure 2.3: The two Δ_q-core graph obtained in the case $l_b = 2, i_a = 1, w_a = l_a = 0$

$Q_2, ..., Q_{m-1}$ are each adjacent to two straight a-edges and we just need to verify it is also the case for Q_1 and Q_m. If $l_a \geq 2$, we can simply attach an a-loop to Q_1 and Q_m and thus we obtain a Δ_q-core graph.

If $l_a = 1$, it suffices to prove that $l_b \geq 1$ and will imply that $i_a \geq m$ and thus that we can close the chain of q-gons. If $l_a = 1$, we have that $2i_a + 1 + w_a = i_a + \frac{1}{2}(qm + 1 + l_b + w_a)$ and thus $l_b = (q-2)m + 3 + w_a > 1$.

If $l_a = 0$, we can do a similar computation and we obtain $l_b = (q-2)m + w_a + 2 > 1$ which concludes the proof.

□

As a direct corollary of Lemma 2.4.8, we obtain Corollary 2.5.3.

Corollary 2.5.3. *Let H be a finitely generated subgroup of Δ_q, whose Δ_q-core graph has combinatorial type $(n, l_a, i_a, w_a, l_b, m)$. Then H has finite index if and only if $w_a = 0$. In that case, H has index n.*

Let us recall Kurosh's subgroup theorem, which describes the subgroups of free products.

Theorem 2.5.4 (Kurosh's Subgroup Theorem)**.** *Let $G_1, \ldots, G_n$ be groups, and $G := *_{i=1}^{n} G_i$ be their free product. Then, any subgroup $H \leq G$ of G has the form:*

$$(*_{i=1}^{n} *_{j=1}^{m_i} w_{ij} H_{ij} w_{ij}^{-1}) * F(X)$$

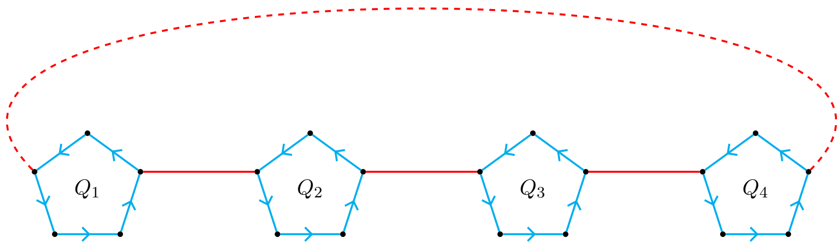


Figure 2.4: Example in the case $q = 5$ and $m = 4$. The dashed arrow is present in the case $l_b = 1$ and is not present in the case $l_b = 0$ and is replaced by a-loops at Q_1 and Q_4.

where each H_{ij} is a subgroup of G_i, $F(X)$ is a free subgroup generated by a subset X of G, and w_{ij} is an element of G, for $1 \leq i \leq n$ and $1 \leq j \leq m_i$.

Thus a subgroup Λ of Δ_q is free if and only if it is torsion free. This implies Lemma 2.5.5.

Lemma 2.5.5. *Let Λ be a finitely generated subgroup of Δ_q whose Δ_q-core graph has combinatorial type $\gamma = (n, l_a, i_a, w_a, l_b, m)$. Then Λ is free if and only if $l_a = l_b = 0$.*

2.6 Enumeration of graphs

We want to enumerate and compute Q_n, the number of connected $\{a, b, b^{-1}\}$-labelled graphs on n vertices such that:

- every vertex belongs either to a b-loop or to a b-q-gon.
- every vertex belongs either to an a-loop, a straight a-edge or has a-degree equal to 0.

Since every vertex is either on a b-loop or on a b-q-gon, the b-edges determine a partition of the vertex set into 1- and q-element sets. Such a structure, that we will call a τ_b-structure, is characterized by the exponential generating function (EGS) defined by

$$T_b(z, u, y) = \exp\left(zu + \frac{z^q}{q}y\right)$$

where the variable z counts the number of vertices, the variable u the number of loops and the variable y the number of q-gons.

The vertices of the graph are also partitioned with respect to a-edges. The vertices can either be on an a-loop, on an a-edge or have no adjacent edge coloured by a. Such a structure will be called a τ_a-structure. The EGS of such structures is given by :

$$T_a(z, u, s, x) = \exp\left(zu + \frac{z^2}{2}x + zs\right)$$

For T_a, the variable z counts the number of vertices, the variable u the number of loops, the number of vertices without a-edges is counted by the variable s and x counts the number of a-edges.

A graph corresponding to the core of a finitely generated subgroup of Δ_q is obtained by a superposition of those two structures. However, not every graph obtained by superposition of these structures is the core graph of a finitely generated subgroup of Δ_q.

If $A(z) = \sum_{i=0}^{+\infty} a_n z^n$, we denote $[z^n]A(z)$ the coefficient in front of the term z^n. We introduce the multivariate EGS $\widetilde{Q}(z, u, s, x, y)$ that counts the number of graphs obtained by superposing aτ_a and a τ_b structure and so $\widetilde{Q}_{n,l,w,i,m} = [z^n u^l s^w x^i y^m]\widetilde{Q}$ is the number of non-empty, possibly disconnected, labelled graphs with defining tuple $\gamma = (n, l, w, i_a, m)$ where $l = l_a + l_b$.

If $Q(z, u, s, x, y)$ is the EGS of such connected structures, we can retrieve it by the functional equation

$$1 + \widetilde{Q} = \exp(Q)$$

or equivalently

$$Q = \log\left(1 + \widetilde{Q}\right)$$

and $[z^n u^l s^w x^i y^m]Q$ is the number of connected labelled graphs obtained by superposition of a τ_b and a τ_a structure.

In particular, the series $\widetilde{Q}$ can be expressed in terms of T_a and T_b :

$$[z^n u^l s^w x^i y^m]\widetilde{Q} = n! \sum_{l_1+l_2=l} [z^n u^{l_1} s^w x^i]T_a \cdot [z^n u^{l_2} y^m]T_b$$

2.7 Asymptotics of finite index subgroups

In this section we give the asymptotic behaviour of the number of subgroups of Δ_q of fixed index n. A finite index subgroup is determined by the condition $w_a = 0$ as proved in Corollary 2.5.3. We thus need to evaluate the asymptotic behaviour of the coefficients of the exponential generating function $G(z) = Q(z, 1, 0, 1, 1)$. To do so, we will first study the coefficients of $\widetilde{G} = \widetilde{Q}(z, 1, 0, 1, 1)$, $T_2(z) = T_a(z, 1, 0, 1)$ and $T_q(z) = T_b(z, 1, 1)$. The tools we use in this section are developed in [BNW20].

Thus $T_2(z) = T_a(z, 1, 0, 1) = \exp(S_2(z))$ where $S_2(z) = z + \frac{z^2}{2}$ and $T_q(z) = T_b(z, 1, 1) = \exp\left(S_q(z)\right)$ for $S_q(z) = z + \frac{z^q}{q}$. To do so we will use the following propositions.

Proposition 2.7.1. *[Proposition 4.1, [BNW20]] Let $S(z) = \sum_{i=1}^{d} s_i z^i$ be an aperiodic polynomial with non-negative coefficients and let $A(z) = \exp(S(z))$. Then*

$$[z^n]A(z) \sim \frac{e^{S(c_n)}}{\sqrt{2\pi\lambda(c_n)}c_n^n}$$

where $\lambda(z) = z^2 \frac{d^2}{dz^2} S(z) + z\frac{d}{dz} S(z)$ and c_n is the least positive solution of the equation $z\frac{d}{dz} S(z) = n$.

The use of Proposition 2.7.1 requires the asymptotics of c_n and c_n^n. Those can be computed using Lemma 2.7.2 and Corollary 2.7.3.

Lemma 2.7.2. *[Lemma 4.2, [BNW20]] Let $S(z) = \sum_{i=1}^{d} s_i z^i$ be a polynomial of degree $d \geq 1$. Then the least positive solution c_n of the equation $z\frac{d}{dz} S(z) = n$ satisfies*

$$c_n = \left(\frac{n}{d s_d}\right)^{1/d} \left(1 + R\left(n^{-\frac{1}{d}}\right) + o\left(n^{-1}\right)\right)$$

where R is a polynomial of degree at most d, depending on S only and such that $R(0) = 0$.

Corollary 2.7.3. *[Corollary 4.3, [BNW20]] With the same notation as in Lemma 2.7.2, we have*

$$c_n^n \sim \left(\frac{n}{ds_s}\right)^{\frac{n}{d}} \exp\left(Q\left(n^{\frac{1}{d}}\right)\right)$$

where Q is a polynomial of degree at most $d-1$, which depends on S only.

Finally, combining Proposition 2.7.1, Lemma 2.7.2 and Corollary 2.7.3 we have an asymptotic estimate of $[z^n]A(z)$ given by Proposition 2.7.4.

Proposition 2.7.4. *[Proposition 4.4, [BNW20]] Let $S(z) = \sum_{i=1}^{d} s_i z^i$ be an aperiodic polynomial with non-negative coefficients and let $A(z) = \exp(S(z))$. Then*

$$[z^n]A(z) \sim \frac{1}{\sqrt{2\pi dn}} n^{-\frac{n}{d}} \exp\left(\frac{n}{d}\left(1 + \log\left(ds_d\right)\right) + T\left(n^{\frac{1}{d}}\right)\right)$$

for some polynomial $T(z)$ of degree at most $d-1$.

2.7.1 Computations of the polynomials R, Q and T

The computations for $S_2(z) = z + \frac{z^2}{2}$ are done in [BNW20]. In this subsection we compute the polynomials R, Q and T for $S_q(z) = z + \frac{z^q}{q}$.

Let c_n be the smallest root of $z^q + z = n$. Then $c_n^q + c_n = n \Rightarrow c_n^q = n - c_n \le n$. Thus

$$c_n = O\left(n^{1/q}\right)$$

Moreover, $c_n^q = n - c_n = n\left(1 - \frac{c_n}{n}\right)$. Thus

$$c_n = n^{1/q}\left(1 - \frac{c_n}{n}\right)^{1/q}$$

We use the Taylor expansion of $\left(1 - \frac{c_n}{n}\right)^{1/q}$:

$$\left(1 - \frac{c_n}{n}\right)^{1/q} = 1 - \frac{1}{q}\frac{c_n}{n} + \frac{(q-1)}{2q^2}\frac{c_n^2}{n^2} - \frac{(q-1)(2q-1)}{6q^3}\frac{c_n^3}{n^3} + \dots$$

But since $c_n = O\left(n^{1/q}\right)$, for all $k \le q-1$:

$$\left(1 - \frac{c_n}{n}\right)^{1/q} = 1 + O\left(n^{-\frac{q-1}{q}}\right) = 1 + O\left(n^{-\frac{k}{q}}\right)$$

Thus $R_k(z) \equiv 1$ for all $1 \le k \le q-1$.

Now we can compute R. We have

$$c_n = n^{1/q}\left(1 + O\left(n^{-\frac{q-1}{q}}\right)\right) \tag{2.7.1}$$

Thus if we divide both sides by n we obtain

$$\frac{c_n}{n} = n^{\frac{1-q}{q}} + O\left(n^{-1}\right) \tag{2.7.2}$$

Now if we substitute Equation (2.7.2) in Equation (2.7.1) and using the Taylor expansion of $(1-x)^{1/q}$:

$$c_n = n^{1/q}\left(1 - \left(n^{\frac{1-q}{q}} - O\left(n^{-1}\right)\right)\right) = n^{1/q}\left(1 - \frac{1}{q}n^{\frac{1-q}{q}} + O\left(n^{-1}\right)\right)$$

Thus

$$c_n = n^{1/q}\left(1 - \frac{1}{q}\left(n^{-1/q}\right)^{q-1} + O\left(n^{-1}\right)\right)$$

and $R(z) = \frac{-1}{q}z^{q-1}$.

Now following Corollary 2.7.3, we will compute the polynomial $Q(z)$.

We want first to find a polynomial $R_1(z)$ such that $R_1(0) = 0$, the degree of R_1 is smaller or equal to q and

$$\log\left(1 + R\left(n^{-1/q}\right) + o\left(n^{-1}\right)\right) = R_1\left(n^{-1/q}\right) + o\left(n^{-1}\right)$$

This condition is equivalent to:

$$\log\left(1 - \frac{1}{q}n^{\frac{1-q}{q}} + o(n^{-1})\right) = R_1(n^{-1/q}) + o(n^{-1}) \tag{2.7.3}$$

Consider first the substitution $u = n^{-1/q}$. Thus when $n \to +\infty$, $u \to 0$. Equation (2.7.3) becomes now :

$$\log\left(1 - \frac{1}{q}u^{q-1} + o(u^q)\right) = R_1(u) + o\left(u^q\right) \tag{2.7.4}$$

We now use the Taylor expansion of $\log(1 + v)$:

$$\log(1 + v) = \sum_{k=1}^{q}\frac{(-1)^{k+1}}{k}v^k + o\left(v^q\right)$$

So if we make the substitution $v = \frac{-1}{q}u^{q-1} + o\left(u^q\right)$, we have that

$$v^k = \left(\frac{-1}{q}u^{q-1}\right)^k (1 + o(u))^k$$

And since for $k \geq 2$, $v^k = o\left(u^q\right)$:

$$\log(1 + v) = \frac{-1}{q}u^{q-1} + o\left(u^q\right) = \frac{-1}{q}n^{\frac{1-q}{q}} + o\left(n^{-1}\right)$$

and thus $R_1(z) = \frac{-1}{q}z^{q-1}$ giving $Q(z) = \frac{-1}{q}z$.

We focus now on the computation of the polynomials P and T following Proposition 2.7.4.

Recall that $c_n = n^{1/q}\left(1 + R\left(n^{-1/q}\right) + o\left(n^{-1}\right)\right)$ and again make the substitution $u = n^{-1/q}$. Thus :

$$c_n = \frac{1}{u}\left(1 - \frac{1}{q}u^{q-1} + o(u^q)\right)$$

Now

$$S\left(c_n\right) = \frac{1}{q}\frac{1}{u^q}\left(1 - \frac{1}{q}u^{q-1} + o(u^q)\right)^q + \frac{1}{u}\left(1 - \frac{1}{q}u^{q-1} + o(u^q)\right) \tag{2.7.5}$$

and we want to give an expression for it of the form $\frac{n}{q} + P(n^{1/q}) + o(1)$ for the asymptotic equivalence.

Now equation 2.7.5 can be rewritten as :

$$\begin{aligned}
&\frac{1}{u} + o(1) + \frac{1}{q}\frac{1}{u^q}\left(1 - \frac{1}{q}u^{q-1} + o(u^q) + O(u^{2q-2})\right) \\
&= \frac{1}{u} + o(1) + \frac{1}{q}\frac{1}{u^q} - \frac{1}{q^2}\frac{1}{u} + \frac{1}{q}\cdot o(1) + \frac{1}{q}\frac{1}{u^q}O(u^{2q-2}) \\
&= \left(1 - \frac{1}{q^2}\right)\frac{1}{u} + \frac{1}{q}\frac{1}{u^q} + o(1) + \frac{1}{q}O(u^{q-2}) \\
&= \left(1 - \frac{1}{q^2}\right)\frac{1}{u} + \frac{1}{q}\frac{1}{u^q} + o(1) \\
&= \left(1 - \frac{1}{q^2}\right)n^{1/q} + \frac{n}{q} + o(1)
\end{aligned}$$

Thus $P(n^{1/q}) = \left(1 - \frac{1}{q^2}\right)n^{1/q}$ and $P(z) = \left(1 - \frac{1}{q^2}\right)z$.

Now the polynomial $T(z)$ can be computed :

$$T(z) = P(z) - Q(z) = \left(1 + \frac{1}{q} - \frac{1}{q^2}\right)z$$

The previous computations can be summarized in the following table:

	$S_2 = \frac{1}{2}z^2 + z$	$S_q = \frac{z^q}{q} + z$
$R(Z)$	$\frac{1}{8}z^2 - \frac{1}{2}z$	$\frac{-1}{q}z^{q-1}$
$Q(z)$	$\frac{-1}{2}z$	$\frac{-1}{q}z$
$T(z)$	$z - \frac{1}{4}$	$\left(1 + \frac{1}{q} - \frac{1}{q^2}\right)z$

Using the table above, we can apply Proposition 2.7.4 to $T_2(z) = \exp(S_2(z))$ and $T_q(z) = \exp\left(S_q(z)\right)$.

Corollary 2.7.5. *We have the following asymptotic estimates.*

$$\begin{aligned}
[z^n]T_2(z) &\sim \frac{e^{-1/4}}{2\sqrt{\pi n}}n^{-n/2}\exp\left(\frac{1}{2}n + n^{1/2}\right) \\
[z^n]T_q(z) &\sim \frac{1}{\sqrt{2\pi q n}}n^{-n/q}\exp\left(\frac{n}{q} + \left(1 + \frac{1}{q} - \frac{1}{q^2}\right)n^{1/q}\right)
\end{aligned}$$

Proposition 2.7.6. *Let $\widetilde{Q}(z, u, s, x, y)$ be the exponential generating function of the graphs obtained by superposing a τ_a and a τ_b structure as defined in Section 2.6. Then*

$$[z^n]\widetilde{G} = [z^n]\widetilde{Q}(z, 1, 0, 1, 1) \sim$$

$$\sim \frac{\sqrt{2\pi n}\cdot e^{-1/4}}{2\pi n\sqrt{2q}} n^{n-n/q-n/2} \exp\left(\left(\frac{1}{2}+\frac{1}{q}-1\right)\cdot n + n^{1/2} + \left(1+\frac{1}{q}-\frac{1}{q^2}\right)n^{1/q}\right).$$

Proof. Note that $[z^n]\widetilde{Q}(z,1,0,1,1) = n![z^n]T_2(z)[z^n]T_q(z)$ and $n! \sim \sqrt{2\pi n}\, e^{-n}n^n$ and the result follows from Corollary 2.7.5. □

Proposition 2.7.6 gives the asymptotics for the possibly disconnected graphs. We want to estimate the asymptotic behaviour of the coefficients of $G = \log(1+\widetilde{G})$, that is only the connected graphs. Proposition 2.7.7 gives us the necessary tool to do so.

Proposition 2.7.7. *[Proposition 6.1, [BNW20]] Let $A(z) = 1 + \sum_{n\geq 1} A_n z^n$ be a formal power series such that*

$$A_n \sim \alpha \cdot n^{\beta} \cdot e^{F(n)} \cdot n^{\epsilon\cdot n}$$

where $F(n) = \sum_{i=1}^{d} \gamma_i n^{\delta_i}$ for some non-zero constants $d, \alpha, \beta, \gamma_i, \delta_i, \epsilon$ satisfying $d \geq 1$ and $0 < \delta_d < ... < \delta_2 < \delta_1 = 1$. Then

$$\frac{A_{n-1}}{A_n} \sim e^{-\gamma_1-\epsilon} n^{-\epsilon}$$

If in addition $\alpha, \epsilon > 0$ and $B(z)$ are such that $B(z) = \log(A(z))$ then for every $s \geq 1$, we have:

$$[z^n]B(z) = \sum_{k=0}^{s-1} b_k A_{n-k} + O(A_{n-s})$$

where $b_0 = 1$ and for each $k \geq 1$:

$$b_k = \sum_{l=1}^{k} (-1)^l \sum_{j_1+...+j_l=k} A_{j_1}\cdot ... \cdot A_{j_l}$$

In our case, $A(z) = 1 + \widetilde{G}(z)$ and $B(z) = \log(1+\widetilde{G}(z)) = G(z)$ and so $A_n \sim \alpha \cdot n^{\beta} \cdot e^{F(n)} \cdot n^{\epsilon\cdot n}$ where :

$$\alpha = \frac{e^{-1/4}}{2\pi\sqrt{2q}}\sqrt{2\pi}$$
$$\beta = \frac{1}{2} - 1 = \frac{-1}{2}$$
$$\epsilon = \left(1 - \frac{1}{q} - \frac{1}{2}\right)$$
$$F(n) = \left(\frac{1}{2}+\frac{1}{q}-1\right)n + n^{1/2} + \left(1+\frac{1}{q}-\frac{1}{q^2}\right)n^{1/q}$$

Now we obtain a result similar to [BNW20, Theorem 7.1].

Theorem 2.7.8. *The probability p_n that an n-vertex labelled graph whose set of a-edges (resp. b-edges) is determined by a τ_2-structure (resp. τ_q-structure) is the core-graph of a finite index subgroup of Δ_q of index n satisfies:*

$$p_n = \frac{[z^n]G}{[z^n]\widetilde{G}} = 1 - e^{1-\frac{1}{q}} \cdot n^{\frac{1}{2}+\frac{1}{q}-1} + O\left(n^{\frac{2}{q}-1}\right)$$

Proof. Recall that $1 + \widetilde{G} = \exp(G)$ and $\widetilde{G}$ satisfies the hypothesis of Proposition 2.7.7 with $\epsilon = \left(1 - \frac{1}{q} - \frac{1}{2}\right)$ and $b_1 = -1$.

Then by the second part of Proposition 2.7.7 and setting $s = 2$, we deduce:

$$[z^n]G = [z^n]\widetilde{G} - [z^{n-1}]\widetilde{G} + O\left([z^{n-2}]\widetilde{G}\right)$$

Thus $p_n = 1 - \frac{[z^{n-1}]\widetilde{G}}{[z^n]\widetilde{G}} + O\left(\frac{[z^{n-2}]\widetilde{G}}{[z^n]\widetilde{G}}\right)$

We can now again apply Proposition 2.7.7 and we obtain:

$$\frac{[z^{n-1}]\widetilde{Q}}{[z^n]\widetilde{Q}} \sim e^{1-\frac{1}{q}} \cdot n^{\frac{1}{2}+\frac{1}{q}-1}$$

Moreover, $\frac{[z^{n-2}]\widetilde{Q}}{[z^n]\widetilde{Q}} = \frac{[z^{n-2}]\widetilde{Q}}{[z^{n-1}]\widetilde{Q}} \cdot \frac{[z^{n-1}]\widetilde{Q}}{[z^n]\widetilde{Q}} \sim e^{2(1-\frac{1}{q})} \cdot n^{\frac{2}{q}-1}$ which concludes the proof. □

Corollary 2.7.9. *The number λ_n of index n subgroups of Δ_q satisfies*

$$\lambda_n \sim \frac{\sqrt{2\pi n} \cdot e^{-1/4}}{2\pi n \sqrt{2q}} n^{n-n/q-n/2} \exp\left(\left(\frac{1}{2} + \frac{1}{q} - 1\right) \cdot n + n^{1/2} + \left(1 + \frac{1}{q} - \frac{1}{q^2}\right) n^{1/q}\right)$$

Proof. Theorem 2.7.8 implies that $[z^n]\widetilde{G} \sim [z^n]G$ and the rest follows from Proposition 2.7.6. □

2.8 Open questions

As defined in [BNW20], the size of a finitely generated subgroup $\Lambda < \Delta_q$ is the number of vertices in $\mathrm{Core}(\Delta_q, \Lambda)$. The following questions arise naturally, and provide possible directions for future research.

Question 2.8.1. Compute the number of subgroups $\Lambda < \Delta_q$ of size n or, at least, determine its asymptotic behaviour.

Question 2.8.2. Same question as above for free subgroups of given size.

Question 2.8.3. Study the statistical properties of size $n \to \infty$ subgroups in Δ_q (for $q \geq 3$). What is the expected isomorphism type? What is the asymptotic relation between free and non-free subgroups? Which properties are "negligeable" (e.g. almost malnormality, if $q = 3$ [BNW20]) and which are "generic" (e.g. parabolicity, if $q = 3$ [BNW20])?

3 Whitehead moves on graphs and their expansion properties

3.1 Introduction

The present chapter is concerned with the graph of 3-regular, or cubic, graphs, and associated objects, such as triangulations and pants decompositions of surfaces. Such a "graph of graphs" Γ_n, represents what can be informally described as the "deformations space" of cubic graphs on $2n$ vertices under the Whitehead moves. In particular, we shall investigate its expansion property.

Let g be a cubic graph (we allow graphs with multiple edges and loops), with set of vertices $V(g)$, and set of edges $E(g)$. Then, for an edge $\varepsilon \in E(g)$ which is not a loop, there are two possible Whitehead moves $w_\varepsilon^{(1)}$ and $w_\varepsilon^{(2)}$ on ε, depicted in Figure 3.1, which are local transformations of g. If $\varepsilon \in E(g)$ is a loop, we assume that the respective Whitehead moves have no effect on g.

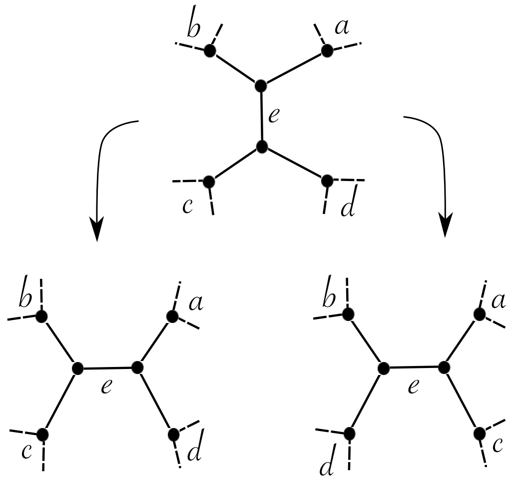


Figure 3.1: Two possible Whitehead moves performed on an edge ε.

Let the vertices of Γ_n be all possible cubic graphs on $2n$ vertices ($n \geq 1$) up to isomorphism, where two vertices $g_1, g_2 \in V(\Gamma)$ are connected by a directed edge for each Whitehead move w_ε, for $\varepsilon \in E(g_1)$, such that $w_\varepsilon(g_1)$ is isomorphic to g_2. It is easy to see that being related by a Whitehead move is an equivalence relation on the set of isomorphism classes of cubic graphs on $2n$ vertices ($n \geq 1$).

Such a graph of graphs is related to moduli spaces of surface triangulations [DP19, DP+18] and pants decomposition of surfaces [RT+13], although not entirely equivalent to said objects. The definition considered here is a directed version of the graph of graphs considered in [RT+13]. Since Whitehead moves are reversible, our graph of graphs is quasi-isometric to the one considered in [RT+13]. The graph of graphs considered in [RT+13] is quasi-isometric to the "wide" component of the thick part of the moduli space of a genus $g \geq 2$ surface. This means that Γ_n describes pants decompositions of the genus $g = n + 1$ surface in which the set of separating curves contains only sufficiently short geodesics. And so is the directed version considered here. Similar local transformations are considered in [Hat99] and two such pants decompositions represent adjacent vertices in Γ_n if they are connected by an A-move, as described in [Hat99], while an S-move (also following [Hat99]) is always performed on a handle that corresponds to a loop in the pants decomposition graph, and thus is not taken into account when passing to Γ_n.

Thus, knowing the combinatorial properties of Γ_n could be useful in the study of the geometric and combinatorial properties of the former.

Here, we introduce a variation $\phi_o ut$ of the usual notion of conductance and show that this quantity tends to 0 for the family Γ_n of cubic graphs on $2n$ vertices.

In the present chapter, we show that the following statements hold.

Theorem (Theorem 3.2.2)**.** *Let Γ_n be the graph of cubic graphs on $2n$ ($n \geq 1$) vertices connected by Whitehead moves. Then $\phi_{out}(\Gamma_n) \to 0$, as $n \to \infty$.*

3.2 Conductance and expansion

Let G be a connected directed graph with vertex set $V(G)$ and edge set $E(G)$. For a subset $S \subset V(G)$, let $d(S)$ denote the sum of vertex out-degrees in S, i.e. $d(S) = \sum_{v \in S} deg_{out}(v)$, where $d_{out}(v)$ is the number of edge of the form (v, u) where $v, u \in V(G)$. The *boundary* $\partial(S)$ of a vertex subset $S \subset V(G)$ is defined as the number of directed edges in $E(G)$ joining a vertex in S with a vertex in $V(G) \setminus S$, that is $\partial(S) = \{e = (u, v) \in E(G) \mid u \in S \ , v \in S^c\}$.

The *outer-conductance* of $S \subset V(G)$ is defined as follows:

$$\phi_{out}(S) = \frac{|\partial(S)|}{\min\{d(S), d(V(G) - S)\}}$$

and the outer-conductance of G is

$$\phi_{out}(G) = \min_{S \subset V(G)} \phi_{out}(S).$$

This is a generalization of the notion of conductance [CG97]. Since we work with directed graphs, the definition is adapted so that the volume is measured with respect to the number of out-going edges. By doing so the outer-conductance measures how hard it is to escape a subset of the graph.

As a generalisation of the notion of expander families to the case of directed graphs with vertex degrees (both out-degree and in-degree) growing with the number of vertices, we introduce Definition 3.2.1.

Definition 3.2.1. We say that a sequence of directed graphs G_n with out-degree and in-degree growing with the number of vertices is an outer-expander family if the conductance $\phi_{out}(G_n)$ tends to 0 as n tends to infinity.

The main theorem in this chapter is the following.

Theorem 3.2.2. *Let Γ_n be the directed graph of cubic graphs on $2n$ ($n \geq 1$) vertices connected by Whitehead moves. Then Γ_n is connected and $\phi(\Gamma_n) \to 0$, as $n \to \infty$.*

3.3 Proof of Theorem 3.2.2

Let Γ_n be the directed graph of (isomorphism classes of) cubic graphs on $2n$ vertices ($n \geq 1$). The following claim will imply Theorem 3.2.2:

There exists a subset $B_n \subset V(\Gamma_n)$ such that its conductance $\phi_n = \phi_{out}(B_n)$ satisfies $\lim_{n\to\infty} \phi_n = 0$.

Γ_n is connected.

This follows from [RT+13]. Even simpler, by means of Whitehead moves we can bring every cubic graph to a certain uniquely determined cubic tree.

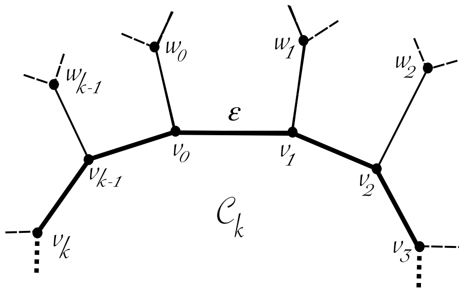


Figure 3.2: Performing a Whitehead move on the edge $\varepsilon = (v_0, v_1)$ that belongs to a k-cycle C_k in $g \in V(\Gamma_n)$.

Let g be a cubic graph, and let C_k be an indecomposable k-cycle (with $k \geq 2$) in g, i.e. an induced cycle subgraph where no edge of g connects any two vertices of C_k, except the edges of C_k itself. If we perform a Whitehead move on the edge ε that belongs to a k-cycle C_k, a part of which is depicted in Figure 3.2, we obtain a modified graph (partly) depicted in Figure 3.3.

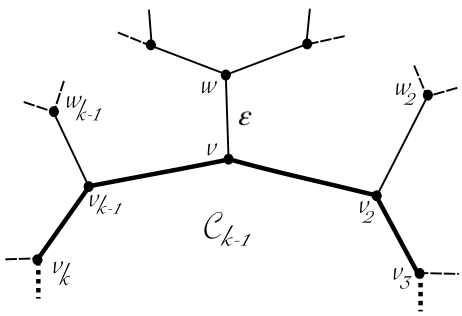


Figure 3.3: A $(k-1)$-cycle C_{k-1} in $\tilde{g} = w_\varepsilon(g)$: the edge $\varepsilon = (v, w)$ does not belong to C_{k-1}.

The overall changes in the structure of $\tilde{g} = w_\varepsilon(g)$ as compared to g are local and amount to the following:

- The k-cycle C_k has been transformed into a $(k-1)$-cycle C_{k-1} ;
- If v_1, w_1 were part of an l-cycle C_l then C_l was transformed into an $(l+1)$-cycle C_{l+1};
- Same applies to any cycle that previously contained w_0 and/or v_0.

By performing a total of $k-1$ Whitehead moves on the edges in C_k, we reduce it to a loop, as depicted in Figure 3.4. Here we reduced the total amount of k-cycles with $k \geq 2$ by one, although the lengths of some other cycles could have been augmented.

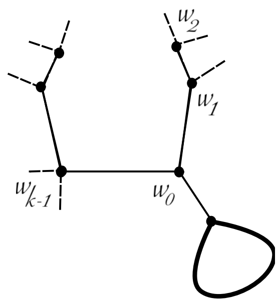


Figure 3.4: A k-cycle C_k ($k \geq 2$) is reduced to a loop by $k-1$ consecutive Whitehead moves.

By repeating the above procedure on the remaining k-cycles (with $k \geq 2$) of the resulting graph, we shall reduce it to what we call a (n, b)-cubic tree, an example of which is depicted in Figure 3.5. Such a tree has $2n$ vertices and $b = n + 1$ loops (note that b is the 1st Betty number of the initial graph g, since Whitehead moves preserve the number of vertices and edges of g, as well as the rank of $\pi_1(g)$).

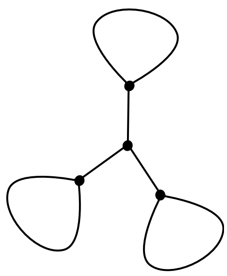


Figure 3.5: A $(2, 3)$-cubic tree

There exists a natural isomorphisms between our (n, b)-cubic trees with a chosen loop (i.e. rooted (n, b)-cubic trees with a chosen loop ε^* as a root), and binary rooted trees. Namely, let us take a rooted (n, b)-cubic tree (T, ε^*) and remove all the loops as well as the edge (v^*, v) adjacent to ε^* (with ε^* being a loop at vertex v). The result, $(\tilde{T}, v^*)$ is a binary rooted tree on $2n-1$ vertices with n leaves.

As shown in Figure 3.6, a tree rotation can be achieved by a Whitehead move. Thus, $(\tilde{T}, v^*)$ can be brought to the unique complete rooted binary tree T_n on n vertices, cf. [CLRS09, §12 – §13] for more information on binary trees and tree rebalancing. This proves that Γ_n is connected.

The outer-conductance of Γ_n tends to 0.

Let $B_n \subset V(\Gamma_n)$ be the set of (isomorphism classes of) connected cubic graphs on $2n$ vertices that have at least one bridge. We shall estimate the probability $\mathbb{P}(g \in B_n)$ asymptotically. For this matter, notice that if $U_n = V(\Gamma_n)$ is the set of unlabelled cubic graphs (i.e. their isomorphism classes), and L_n is the set of labelled ones, then $\mathbb{P}(g \in U_n \text{ has a bridge}) = \mathbb{P}(g \in L_n \text{ has a bridge}) + o(1)$, since asymptotically virtually all graphs in L_n have *no non-trivial* automorphisms [Bol82], and thus $|L_n| \sim |U_n| \cdot (2n)!$, as $n \to \infty$.

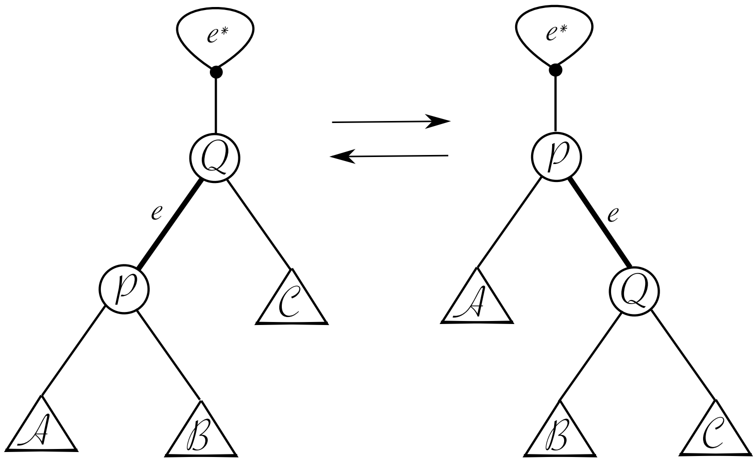


Figure 3.6: A left and right rotation of a rooted binary tree (corresponding to an (n, b)-cubic tree with root e^*) realised by a Whitehead move on edge e. Here P, Q are the vertices of e, and A, B, and C are the respective sub-trees

A *bridge* in a connected graph is an edge such that its removal produces several connected components. A graph with a bridge is called *bridged*, while a graph that has no bridges at all is called *bridgeless*.

A *Hamiltonian circuit* in a (connected) graph is a closed edge path without self-intersection that visits each vertex once. A graph is called *Hamiltonian* if it admits a Hamiltonian circuit.

It is clear that a graph that has a bridge cannot be Hamiltonian, and that every Hamiltonian cubic graph is bridgeless.

Let $Bridged = \{g \in L_n \mid g \text{ has a bridge}\}$ and $Looped = \{g \in L_n \mid g \text{ has a loop}\}$, and let nH be the set of graphs in L_n that do not have a Hamiltonian circuit. Then

$$\mathbb{P}(g \in Looped) \leq \mathbb{P}(g \in Bridged) \leq \mathbb{P}(g \in Looped) + \mathbb{P}(g \notin Looped \,\&\, g \in nH),$$

since $Looped \subseteq Bridged$, as well as $Bridged \subseteq nH$.

By applying [JLR11, Theorem 9.5], we have that $\mathbb{P}(g \in Looped) = 1 - e^{-1} + o(1)$, while [JLR11, Theorem 9.23] implies that $\mathbb{P}(g \notin Looped$ and $g \in nH) = o(1)$, as $n \to \infty$. Thus, $\mathbb{P}(g \in Bridged) = 1 - e^{-1} + o(1)$ asymptotically, as $n \to \infty$.

Together with the remarks above about the probabilities for labelled and unlabelled graphs, we obtain that $\mathbb{P}(g \in B_n)$, for $g \in V(\Gamma_n)$, satisfies $\mathbb{P}(g \in B_n) = 1 - e^{-1} + o(1)$, as $n \to \infty$. Thus, $|B_n| \sim (1 - e^{-1})|V(\Gamma_n)|$, and $|V(\Gamma_n) \setminus B_n| \sim e^{-1}|V(\Gamma_n)|$.

Note that the only vertices in B_n connected with vertices in $V(\Gamma_n) \setminus B_n$ are those $g \in B_n$ having a single bridge, and the corresponding Whitehead moves have to be performed exactly on the bridge edge of g.

Indeed, let $g \in V(\Gamma_n)$ have at least one bridge and let $\varepsilon^* = (u, v)$ denote one of its bridges. Then we can check case-by-case what the result of a Whitehead move on ε^* could be.

For this purpose, let us introduce the following equivalence relation on the vertices of g: for $v, w \in V(g)$, we write $v \sim_g w$ if and only if there are two edge-disjoint paths connecting v to w. Then we write $[v]_g = \{w \in V(g) \text{ s.t. } v \sim_g w\}$ for the equivalence class of v.

Let $\{a, b, c, d\}$ be the vertices adjacent to the bridge $\varepsilon^* \in E(g)$, and let $\tilde{g} = w_{\varepsilon^*}(g)$ be the graph resulting from a Whitehead move performed on ε^*. Then we consider the following five possible cases.

1.$[a]_g = [b]_g = [c]_g = [d]_g$: in this case ε^* is not a bridge neither in g, nor in $\tilde{g}$, cf. Figures 3.7 and 3.8.

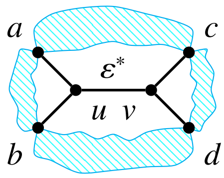


Figure 3.7: Local picture inside g when $[a]_g = [b]_g = [c]_g = [d]_g$

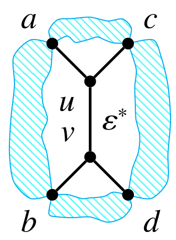


Figure 3.8: Local picture inside $\tilde{g}$ when $[a]_g = [b]_g = [c]_g = [d]_g$

2.$[a]_g = [b]_g = [c]_g \neq [d]_g$: same as above, ε^* is not a bridge neither in g, nor in $\tilde{g}$, cf. Figures 3.9 and 3.10.

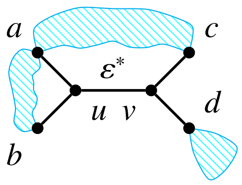


Figure 3.9: Local picture inside g when $[a]_g = [b]_g = [c]_g \neq [d]_g$

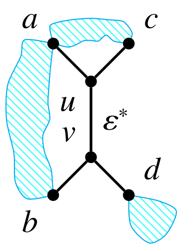


Figure 3.10: Local picture inside $\tilde{g}$ when $[a]_g = [b]_g = [c]_g \neq [d]_g$

3.$[a]_g = [b]_g \neq [c]_g = [d]_g$: in this case, ε^* is a bridge in g, but not in $\tilde{g}$, cf. Figures 3.11 and 3.12.

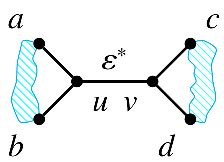


Figure 3.11: Local picture inside g when $[a]_g = [b]_g \neq [c]_g = [d]_g$

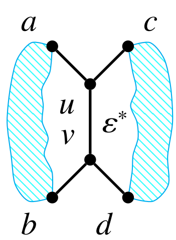


Figure 3.12: Local picture inside $\tilde{g}$ when $[a]_g = [b]_g \neq [c]_g = [d]_g$

4.$[a]_g = [b]_g \neq [c]_g$, $[a]_g = [b]_g \neq [d]_g$, and $[c]_g \neq [d]_g$: again, ε^* is a bridge in g, but not in $\tilde{g}$, cf. Figures 3.13 and 3.14.

5.$[a]_g, [b]_g, [c]_g, [d]_g$ are all distinct: obviously, ε^* is a bridge in both g and $\tilde{g}$, cf. Figures 3.15 and 3.16.

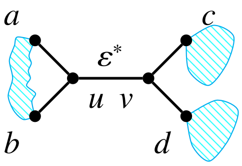


Figure 3.13: Local picture inside g when $[a]_g = [b]_g \neq [c]_g$, $[a]_g = [b]_g \neq [d]_g$, and $[c]_g \neq [d]_g$

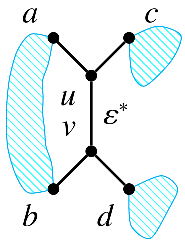


Figure 3.14: Local picture inside $\tilde{g}$ when $[a]_g = [b]_g \neq [c]_g$, $[a]_g = [b]_g \neq [d]_g$, and $[c]_g \neq [d]_g$

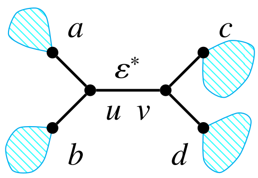


Figure 3.15: Local picture inside g when $[a]_g, [b]_g, [c]_g, [d]_g$ are all distinct

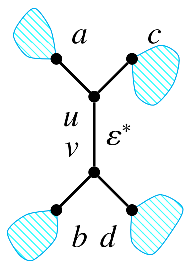


Figure 3.16: Local picture inside $\tilde{g}$ when $[a]_g, [b]_g, [c]_g, [d]_g$ are all distinct

According to the observations above, if a bridged graph $g \in B_n$ can be transformed into a bridgeless one $\tilde{g} = w_\varepsilon(g)$, then ε is the only bridge of g. Every time, both Whitehead moves either succeed or fail to bring $\tilde{g}$ outside of B_n. This implies that $|\partial(B_n)| \leq 2|B_n|$. Thus the outer-conductance of the set B_n satisfies

$$\phi_{out}(B_n) = \frac{|\partial(B_n)|}{\min\{d(B_n), d(V(\Gamma_n) \setminus B_n)\}} \leq \frac{2|B_n|}{6n|V(\Gamma_n) \setminus B_n|} \sim$$
$$\sim \frac{2|V(\Gamma_n)|(1 - e^{-1})}{6n|V(\Gamma_n)|e^{-1}} = \frac{e-1}{3n} \to 0,$$

as $n \to \infty$.

3.4 Open questions

By analogy to the above, we can define the graph of graphs $\Gamma_n^{(k)}$ for k-regular graphs on n vertices, with $k \geq 4$. However, in this case, most of the k-regular graphs will be Hamiltonian, even if they have loops. Thus, our approach by creating a bottleneck on the set of bridged graphs will not work.

Question 3.4.1. Is $\{\Gamma_n^{(k)}\}_{n\in\mathbb{N}}$ an expander family (w.r.t. conductance) of graphs, for $k \geq 4$?

Another question is related to edge switching in trivalent graphs, cf. [FLNU18]. It is easy to see that a Whitehead move can be realised as an edge switch. Thus, the switching graph Σ_n defined in [FLNU18] (with one modification: its vertices should be considered as isomorphism types of graphs), contains the graph of graphs. However, edge switching is a more "global" operation than a Whitehead move, and thus we cannot imply anything about the expansion properties of the family of switching graphs.

Question 3.4.2. Is $\{\Sigma_n\}_{n\in\mathbb{N}}$ an expander family (w.r.t. conductance) of graphs?

4 Vulnerability detection in trophic networks

This chapter is based on an ongoing collaboration with M. Mulot.

4.1 Introduction

An ecosystem is a community of living organisms living in a bounded area characterized by a set of environmental properties (temperature, pH, altitude, etc.). Ecosystems' composition and equilibrium are due to the interactions between the biotic component (species community) and the abiotic components (environmental properties). These interactions define niches that are ecological spaces whose conditions permit a given species to live in. These interactions are of 3 types:

1. abiotic $\leftrightsquigarrow$ abiotic
2. abiotic $\leftrightsquigarrow$ biotic
3. biotic $\leftrightsquigarrow$ biotic

Abiotic–abiotic interactions are essentially physics. For instance, high altitude combined with water input will lead to the formation of ice, and then add *ice* as a property of the ecosystem. Abiotic–biotic interactions are for instance the degradation of mineral particles by microorganisms, or a beaver creating a dam that will raise the water level and flood an area. Finally, the biotic–biotic interactions are mainly the trophic relationships between the species composing the communities.

These trophic interactions (simply put, "who eats whom") shape the structure of the community, which in turn impacts the magnitude of biotic–abiotic interactions. For instance, removing a predator from a community will lead to more grazing, preventing the growth of shrubs or forest, therefore exposing the soil to direct sunlight because of the lack of canopy. Finally, the set of all trophic interactions within an ecosystem drives the biogeochemical (carbon and nitrogen, for instance) cycles leading to ecosystem's equilibrium.

In ecology, it is now common to model these interactions as a network. The trophic interactions are hence crucial to maintain ecosystem stability. However, not all interactions are equally important to maintain equilibrium. Notably, processes of the biogeochemical cycles that are redundant in the trophic network will resist to the loss of a species involved in these processes. Some species are crucial, and their removal will break the biogeochemical cycles and disturb the whole ecosystem. Such species are called keystone species.

Identification of keystone species is essential in conservation biology, that aim at preserving ecosystems. Preserving ecosystems serves other purposes than simply protecting exotic flowers. The primary objective is to preserve ecosystem services, *i.e.,* functions that are provided by an ecosystem, that would be costly to replace by man-made solutions. For instance, mountain forests prevent water from flushing directly down the valley when it rains, some ecosystems act as reservoir of species whose role is crucial in other ecosystems, etc.

However, existing methods to infer topological keystone species do not take into account the prior knowledge we have on organisms about their trophic habits. In this chapter, we propose a new method to identify important species by taking into account both their trophic position and their place in the network.

Chapter 4 is organized as follows. In Section 4.2, we introduce some ecological concepts and give some basic properties of trophic networks. We also explain some challenges when working with trophic networks. In Section 4.3, we explain an existing method for randomly generating correlation/co-occurrence networks. In Section 4.4 we present the method we develop. We give an algorithm to identify nodes that have a pivotal role in the ecosystem and thus can be considered keystone in some sense. This method is new and provides a new approach for detecting essential species in an ecosystem. Finally, in Section 4.5 we provide an example and illustrate it with a randomly generated trophic network.

4.2 Definitions and preliminaries

In Section 4.2, we define precisely the biological concepts we will be working with, and we give the necessary insights on certain objects to allow a clear presentation of our methods in the subsequent sections.

An ecosystem is a community of living organisms living in a bounded area characterized by a set of environmental properties (temperature, pH, altitude, etc.). In an ecosystem, every species has a role and the removal of a given species can have diverse and unpredictable effects. A main interest of ecologists, is to look for Keystone species, see Definition 4.2.1.

Definition 4.2.1 (Keystone species)**.** Species whose presence is crucial for the stability of an ecosystem. Removal or loss of a keystone species will lead to the ecosystem to shift from its actual state.

The problem with such a definition is that there are no precise ways to define what "crucial" means and thus to identify such species. One starting point is to represent the ecosystem as a graph, where the vertices represent species. The edges of the graph roughly represent the "who eats whom" relations, see Definition 4.2.2.

Definition 4.2.2 (Trophic network)**.** A trophic network is a directed graph (V, E) where the vertices are biological species and edges connect pairs of species who exchange substances such as carbon or nitrogen. In certain cases, it simply describes "who eats whom". An edge from A to B indicates that species A preys on species B. Sometimes they are also called "food webs" or "food chains".

The vertices of a trophic network are labelled by the name of the species. The graph can be stratified by the *trophic position*. Usually, we represent on the top part the highest predators of the ecosystem and at the bottom are the primary producers.

There is a natural direction for the flow of substances in the graph, which is the opposite direction that the edges point to. As in rooted trees, we have a natural concept of *parents* and of *descendants*. For a vertex v, its parents are the species that feed on it, that is the vertices w such that there exists an edge $(w, v) \in E$ and its descendants are the species upon which v feeds, that is vertices u such that there exists an edge (v, u). We can thus naturally speak about the *ascending induced graph* and the *descending induced graph* for a vertex v. The ascending induced graph of v is the subgraph of all the "ancestors" of v and the descending induced graph of v is the subgraph of the trophic network consisting of all the "descendants" of v.

Identifying keystone species is a key to protect ecosystems, but the tools for such identifications are lacking. While for large organisms (larger than 1 cm) it is quite trivial to understand their role in ecosystems, it is more complex to understand the structure of the trophic network at the microbial scale, while biogeochemical processes are driven mostly by microorganisms (virus, bacteria, protists, and fungi). One gram of soil can contain several thousands of microbial species interacting with each others.

The amount of information required to completely and accurately represent the trophic network is often not accessible. For instance, when considering microbial communities and ecosystems, not only are ecologists challenged by the number of species but also by the impossibility of observing the trophic interactions. However, methods exists to approximate the structure of such trophic interactions. Briefly, the protocol is to scan for all microbial species in several samples from the same ecosystem using DNA barcoding methods. Then, based on the distribution of the identified species across samples, we can compute co-occurence or correlation networks. Correlation networks are computed from (Spearman or Pearson) correlations or similarity distribution measures.

The trophic network is at least partly embedded in the correlation/co-occurence network, as it represents the interactions between species (c.f. [FR12]). From this, it has been proposed to identify "topological keystone species" (c.f. [Est07], [Jor09]), which are species that are keystone from a network topology perspective by calculating several centrality measures, from local to global and using factor analysis. However, those methods do not take into account the trophic levels of the species considered. The method presented in Section 4.4 offers a new approach to the problem of identifying keystone species while taking into account the trophic position in the network. As it is difficult in ecology to obtain real accurately annotated foodweb from natural settings, we test the method *in silico* by generating realistic artificial foodbwebs following a protocol described in Section 4.3.

4.3 Random generation of trophic networks

In Section 4.3, we present the model we used to generate the mock graphs we use throughout this chapter. Our purpose is to generate graphs that are similar to the ones obtained from real world Next Generation Sequencing samples, while adding trophic relationship information within this network.

This is a 3-step process:

1. Generate a food web with trophic coherence, based on the Preferential Preying Model (c.f. [KJ17]). This leads to the construction of an unweighted directed graph $G = (V, E)$ represented by an adjacency matrix T. This graph represents the *potential* trophic relations, meaning that a given relation (represented by an edge) will happen only if the involved species inhabit the same niche (see Definition 4.3.1).

2. Generate species distribution across space under randomly attributed ecological preferences, and using trophic relations defined in the graph G to constrain the niches. The procedure is described in [EO19].

3. Randomly sample the space populated with species to infer co-occurrence and build a final graph Γ representing the co-occurrences. This is the graph we are interested in.

Several models exists to generate foodwebs, each with their own shortcomings.

One of the oldest algorithms to generate trophic network is the cascade model described in [CBN12]. The cascade model assigns each species a random value from the interval [0,1], corresponding to its trophic position. Each species has a probability of consuming only species with values less than its own, this probability decreasing as the difference between the trophic positions of the two species increases. This underestimates inter-specific trophic similarity and overestimates food-chain length and number in larger webs.

Another popular model is the niche model (see [WM00]), that allows to build trophic networks based both on the trophic position of each species and their ecological niches. Although this creates trophic networks similar to what we can observe in nature, the niche is only in one dimension, which limits the possible overlaps between niches that can occur in space, time, and accross environmental gradients. As we seek to reproduce data gathered from environmental studies, we chose to add the niche constraint further down the network generation pipeline.

Finally, the nested hierarchy model (see [CBBR+04]) brings phylogenetic information to the trophic network, assuming that phylogenetically close taxa share some prey preferences. This performs well. However, assigning a phylogenetic constraint at this stage would be somehow redundant with the simulation of n-dimensional niches afterward. Indeed, the phylogenetic relationships between species and their actual realized niches are the results of biogeographical processes leading to evolutionary processes followed by speciations and/or disjunction events, etc. We thus chose to generate a foodweb based on simple topological properties and weak ecological assumptions, and to add ecological constraints later on.

4.3.1 Generating a foodweb

The protocol described in this section is from [KJ17].

Let B, N and L be the number of basal nodes, total nodes and links respectively in the network, all parameters to be fitted using the empirical network data.

We assign trophic levels $s = 1$ to all basal nodes. We then add $N - B$ new nodes to the network sequentially according to the following rule:

- For each new node j, pick exactly one prey i at random from among all the existing nodes in the network, thus creating an edge from i to j. We temporarily define the trophic level of node j as $\hat{s}_j = \hat{s}_i + 1$.
- We have a network of N nodes and $N-B$ edges, and each node i has a (temporary) trophic level $\hat{s}_i$.

Remaining edges are created to reach L edges, with an edge probability P_{ij} that decays with the (temporary) trophic distance $\hat{x}_{ij} = \hat{s}_j - \hat{s}_i$ between them. We set

$$P_{ij} \propto \exp\left(-\frac{(\hat{r}_{ij} - 1)^2}{2T^2}\right)$$

where T is a free tunable parameter representing the trophic coherence. The probabilities P_{ij} are scaled to get the number of desired edges.

Finally, trophic levels s are recomputed such that for each species j

$$s_j = \frac{1}{k_j^{\text{in}}} \sum_i a_{ij} s_i$$

where $k_j^{in} = \sum_i a_{ij}$ is the number of prey of species j (also known as the in-degree) and the coefficients a_{ij} are the entries of the adjacency matrix T of the food web. Here, the convention is that the directed trophic links point from prey i to predator j.

As this equation is recursive, it implies that two conditions must hold. First, there must be at least one node of zero in-degree – such nodes are called basal; and second, every node in the network must be reachable by a path from at least one basal node. We assign $s_j = 1$ for all basal nodes.

4.3.2 Generating species distribution across niches

For two species to interact, they need to co-exist. The abundance and existence of a species in a given place is described by its ecological niche, defined in Definition 4.3.1.

Definition 4.3.1 (Ecological niche)**.** The thriving of a species can be affected by several environmental parameters (e.g., temperature, pressure, pH). For each species, the number, and the impact of each of the parameters is different. The ecological niche of a given species is the match of the thriving of it to the environmental conditions. Formally, it is a multidimensional space where each dimension corresponds to an environmental parameter affecting the species abundance and a last dimension represents the possible abundance of the species under such conditions.

We present in Section 4.3.2 the method for generating species distribution across niches as described in [EO19].

For n -dimensional niches, we consider an $m \times n$ matrix

$$\begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,3} & \dots & x_{1,n} \\ x_{2,1} & x_{2,2} & x_{2,3} & \dots & x_{2,n} \\ x_{3,1} & x_{3,2} & x_{3,3} & \dots & x_{2,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{m,1} & x_{m,2} & x_{m,3} & \dots & x_{m,n} \end{bmatrix}$$

that defines m niche coordinates in n dimensions of environmental space for which virtual niche values are required. The fundamental niche of a virtual species is defined by three parameters:

- $\lambda_{\max}$, the species' maximum finite rate of increase at the fundamental niche optimum,
- μ, an $n \times 1$ column vector of means that gives the optimum location of the niche in each dimension, and
- Σ, an $n \times n$ variance–covariance matrix that gives the size and orientation of the niche in each dimension. That defines m niche coordinates in n dimensions of environmental space for which virtual niche values are required.

Each species is represented as a set of these variables, so $S = \{\lambda_{\max}, \mu, \Sigma\}$.

For a coordinate column vector $\mathbf{x}_m = \left[x_{m,1}, x_{m,2}, x_{m,3}, \dots, x_{m,n}\right]^T$, the fundamental niche finite rate of increase $\lambda_{F(\mathbf{x})}$ of a virtual species for that coordinate is then calculated as:

$$\lambda_{F(\mathbf{x}_m)} = \lambda_{\max} \times e^{-\frac{1}{2}(\mathbf{x}_m-\mu)^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\mathbf{x}_m-\mu)}$$

The parameter λ_F varies by location, size, orientation, and magnitude within the niche space. The fundamental niche is static for a species. The realized niche is a dynamic realization of niche space as it will fluctuate with changes in species populations, which in some systems will be cyclical or even chaotic.

Realized niches can be viewed as a static point in time, by defining a square matrix I that quantifies the pairwise interspecific species interactions in the virtual community as if all interacting + and − interactions have been balanced out. The diagonal of the matrix is zero, so species do not affect themselves, and the remainder of the matrix has values that represent positive and negative effects, where $\mathbf{I}_{ij}$ is the effect of species j on species i. For example, the interactions' matrix

$$\mathbf{I} = \begin{bmatrix} 0.0 & 0.2 & -0.4 \\ -1.0 & 0.0 & -1.0 \\ -0.6 & 0.3 & 0.0 \end{bmatrix}$$

represents a community of three species in which:

- $\mathbf{I}_{1,2} = 0.2$ and $\mathbf{I}_{2,1} = -1.0$ so S_1 is a predator of S_2,
- $\mathbf{I}_{1,3} = -0.4$ and $\mathbf{I}_{3,1} = -0.6$ so S_1 and S_3 are competitors,
- and $\mathbf{I}_{2,3} = -1.0$ and $\mathbf{I}_{3,2} = 0.3$ so S_3 is also a predator of S_2.

Given a vector $\mathbf{f}_{(\mathbf{x})}$ of fundamental niche values λ_F for n species in a community, that is $\mathbf{f}_{(\mathbf{x})} = \left[\lambda_{F1(\mathbf{x})}, \lambda_{F2(\mathbf{x})}, \ldots, \lambda_{Fn(\mathbf{x})}\right]$, then a vector of realized niches $\mathbf{r}_{(\mathbf{x})} = \left[\lambda_{R1(\mathbf{x})}, \lambda_{R2(\mathbf{x})}, \ldots, \lambda_{Rn(\mathbf{x})}\right]$ can be calculated as:

$$\mathbf{r}_{(\mathbf{x})} = \mathbf{f}_{(\mathbf{x})} \circ e^{\mathrm{f}_{(\mathrm{x})}\mathrm{I}^{\mathrm{T}}}$$

From this we can retrieve the distribution of each species in the niche space.

4.3.3 Parametrizing the realized niche model with trophic relationship

The idea is for each species to (semi) randomly set: $\lambda_{\max}, \mu$, and Σ. It is possible at this stage to decide which species live in the same places. Then, we fill the matrix I such that : an interaction from a prey to a predator is positive for the predator - an interaction from a predator to a prey is negative for the prey.

Note that I_{ij} is not necessarily equal to $-I_{ji}$, when two species share the same prey, they have a negative effect on each others. These effects are not necessarily equal - otherwise set interaction to zero. Note that this model can be further improved by considering other relationships such as commensalism (+/0), symbiosis (+/+), etc. (see [FR12]).

From the niche space, we can randomly sample samples for which we have our species distribution. We are thus in the classical framework of an $n \times m$ Species matrix S with n-the number of species, and m-the number of samples. From there, we can use all "traditional" network inference tool. We chose to use [Vee13], but any tool can be used. Note that there are a lot of thresholds that we can play around with, such as the lower limit to consider a species absent from a sample.

With this configuration, we end up with a graph computed from realistic

- trophic relationships,
- distributions.

The trophic positions are known, therefore we have an interesting object to work with. We can even test different relationships, niche constraints, etc.

4.4 Top–down approach to detect vulnerabilities

In Section 4.4, we present an approach to detect vulnerabilities in trophic and co-occurrence networks for ecosystems. The main idea is to introduce a new distance between the species. This distance takes into account both the whole trophic network as well as the individual trophic positions of each species in the network. In this sense, it is a more network based approach than centrality indices as used in [Jor09] or [GMJ21]. The rationale for this approach is that it takes into account the dependent nature of the relationships occurring in a trophic network. For instance, the interaction between a prey and a predator implies that the predator depends on the prey for its own survival. This methods allow for identifying clusters of species depending on the same trophic cascade.

As stated in Section 4.2, a trophic network is a vertex-labelled graph. For each vertex v representing a species, there is a natural notion of descendants and of parents. Indeed,

- for a vertex v, its parents are the vertices one trophic level above that are connected to v. These are vertices w such that there exists an edge $(w, v) \in E$
- for a vertex v, its children are the vertices one trophic level below that are connected to v. These are vertices u such that there exists an edge (v, u)

see Section 4.4 for an illustration.

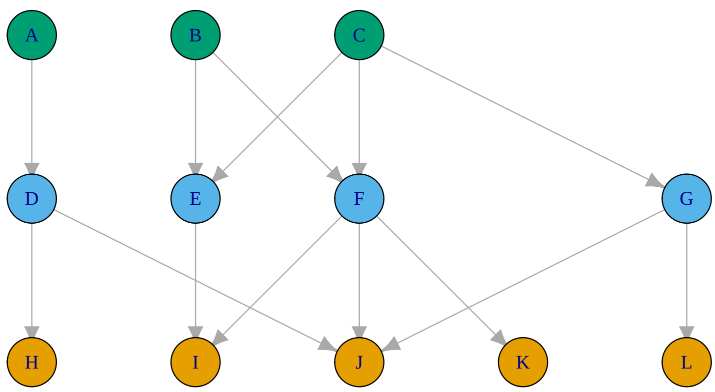


Figure 4.1: The parents of the node F are nodes C and B and the children of the node F are I, J and K. The nodes in green are the "predators", or simply the nodes with the highest trophic level, 3 in this case. The nodes in orange are the basal nodes, usually the primary producers. Their trophic level is 1.

Given this, one can consider the subgraph of all the descendants of a given node: the node and all its children and the children of its children etc. down to the basal nodes. In a trophic network, such basal nodes are called *primary producers* and are organisms capable of photosynthesis, *ie* producing organic matter using sunlight energy. This organic matter is then propagated upward the network due to trophic interactions. If we consider again Section 4.4, the subgraph of descendants induced by C is the subgraph containing the nodes C, E, F, G, I, J, K and L. For a vertex v, we will denote T_v the subgraph of its descendants.

In other terms, T_v is the set of vertices reachable from v by a directed path in the trophic network $G = (V, E)$.

We can now define the main tool of Section 4.4, a new notion of distance between two nodes in a trophic network (see Definition 4.4.1).

Definition 4.4.1 (Trophic distance)**.** Let v, w be two nodes of a trophic network. Then we define the distance between v and w to be :

$$d(v, w) = 1 - \frac{|T_v \cap T_w|}{|T_v \cup T_w|}$$

where T_v and T_w are the graphs of descendants of v and w respectively, and the set operations are considered on the vertices of the subgraphs.

The distance defined in Definition 4.4.1 is a value between 0 and 1.

This is a natural distance to consider when studying trophic networks. Two vertices on different levels are going to be further away than vertices on the same trophic level, unless one is a descendant of the other. Moreover, two species that are close with respect to this distance are species that feed on similar species.

Definition 4.4.1 is the *Jaccard distance* between the two subgraphs induced by the trophic network. Jaccard's index is commonly used in biology and ecology to assess (dis)similarity in communities (e.g., [HCRGE+19]). The difference here is that we consider the Jaccard index of two subgraphs. This distance takes into account the network as a whole and the trophic positions of the species considered.

The first step to identifying keystone species is to determine the distance matrix associated to the trophic network, as described in Algorithm 1.

Algorithm 1 Creation of the trophic distance matrix

Summary: Create a matrix such that an entry of the matrix gives the trophic distance between two nodes.
Input: $G = (V, E)$ - trophic network or correlation network
$v, w \in V$ - two species
Output: trophic distance between v and w
Description:

function $Trophic\_distance(G)$
for each vertex v in V:
  $T_v \leftarrow$ the induced subgraph of v;
create $M = (m_{ij})$ a size $|V| \times |V|$ matrix.
for each $v_i, v_j \in V$:
  compute $d(v_i, v_j) = \frac{|T_{v_i} \cap T_{v_j}|}{|T_{v_i} \cup T_{v_j}|}$
  $m_{ij} \leftarrow d(v_i, v_j)$;
return M;

With this matrix distance, one can now apply some clustering algorithm and obtain $C_1, ..., C_n$, the clusters. We can now look at the subgraphs induced by each of these clusters. That is, we consider $T_{C_i} = \bigcup_{v \in C_i} T_v \subseteq V(G)$. The vertices at the intersection of all/several clusters are the keystone species we identify. This process is described in Algorithm 2.

Now, the vertices of $G = (V, E)$ such that the sum of the coefficients of their corresponding row are maximal or high (with respect to the others) are pivotal: they ensure the connection between two clusters of the graph. In that sense, they are keystone species. Their removal affects the connectivity of the graph.

Algorithm 2 Compute the keystone species of a given trophic network G

Summary: From a trophic network, determine keystone species
Input: $G = (V, E)$ - trophic network or correlation network
Output: Matrix of keystone species
Description:

function $Keystone(G)$
Compute $M = Trophic\_distance(G)$;
From M, determine clusters $C_1, ..., C_n$ (e.g., using a method of hierarchical clustering);
for each C_i:
    $T_{C_i} \leftarrow \bigcup_{v \in C_i} T_v$
Create $K = (k_{v,i})$ a size $|V| \times n$ matrix, where n is the number of clusters
    $k_{v,i} \leftarrow \begin{cases} 1 \text{ if } v \in C_i \\ 0 \text{ otherwise} \end{cases}$;
return K

This method is a top-down approach because of the way we use the network. For a given vertex, we look at its descendants and so we search the graph from top to bottom.

With our approach, we try to solve an issue ecologists face: identifying important species within an ecological network. Most of the existing approaches rely mostly on the sole topology of the network, not taking into account the feeding habits of the studied organisms. For years, this information was lacking, especially for micro-organisms such as protists. Recently, some studies published databases, allowing to inject trophic information into the ecological network(e.g. [BER+20] and [SAM+20]). Therefore, we can use this information to characterize the nature of the relationships unveiled by the co-occurrence computations. Here, we intend to bridge a gap between mathematics and ecology. The major issue we had to face was that most mathematical approaches or theorem on graphs are not easily (if not at all) usable on real word objects, except for the classical centrality indices. Great effort has thus been made to develop a usable mathematical framework that transposes well into ecological data, while being ecologically meaningful. Our ongoing work shows promising results, allowing to identify species that are crucial for the environment. The method is not computationally intensive, and output coherent results on mock communities. Further developments will be:

1. testing on real datasets;

2. optimizing the software implementation.

4.5 Example of application of the Jaccard distance based method

4.5.1 Simple setting

To illustrate the method, let's first consider a very simple trophic network that could be analysed by eye. We will see what vertices does the method point at for being pivotal in maintaining the cohesion of the network. It will then be easy to see why, by looking at the graph.

Figure 4.2 depicts a trophic network G generated using the method described in Section 4.3 with $B = 4$ basal nodes, $N = 23$ nodes and $L = 71$ links, and we denote V for the set of nodes of the graph.

We see that the nodes are partitioned according to their trophic levels as discussed in Section 4.2, and these levels are represented by the different colours of the nodes. In orange are the basal nodes and in red is the top predator.

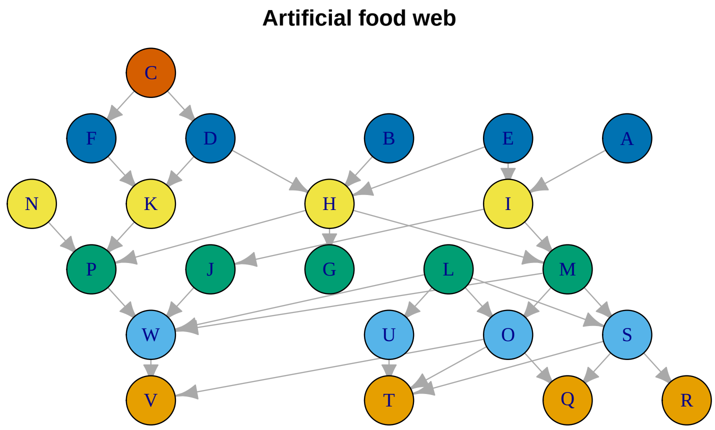


Figure 4.2: An artificial food web G generated using an R implementation of the method described in Section 4.3. The colours of the node represent their trophic level.

We can compute the distance for each pair of nodes. The results of the computations using Definition 4.4.1 can be found in Table 4.1.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
A	0.00	0.47	0.56	0.50	0.33	0.86	1.00	0.43	0.09	0.73	0.85	0.46	0.27	0.85	0.64	0.83	0.91	0.91	0.64	0.91	0.92	0.91	0.82
B	0.47	0.00	0.31	0.21	0.27	0.79	0.92	0.08	0.43	0.85	0.77	0.50	0.33	0.77	0.67	0.75	0.92	0.92	0.67	0.92	0.92	0.92	0.83
C	0.56	0.31	0.00	0.13	0.39	0.67	0.93	0.27	0.53	0.88	0.73	0.59	0.47	0.81	0.73	0.80	0.93	0.93	0.73	0.93	0.94	0.93	0.87
D	0.50	0.21	0.13	0.00	0.31	0.71	0.92	0.15	0.47	0.86	0.69	0.53	0.38	0.79	0.69	0.77	0.92	0.92	0.69	0.92	0.93	0.92	0.85
E	0.33	0.27	0.39	0.31	0.00	0.81	0.93	0.21	0.29	0.79	0.80	0.56	0.43	0.80	0.71	0.79	0.93	0.93	0.71	0.93	0.93	0.93	0.86
F	0.86	0.79	0.67	0.71	0.81	0.00	1.00	0.77	0.85	0.67	0.20	0.83	0.82	0.50	0.88	0.40	1.00	1.00	1.00	1.00	1.00	0.80	0.60
G	1.00	0.92	0.93	0.92	0.93	1.00	0.00	0.91	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
H	0.43	0.08	0.27	0.15	0.21	0.77	0.91	0.00	0.38	0.83	0.75	0.46	0.27	0.75	0.64	0.73	0.91	0.91	0.64	0.91	0.92	0.91	0.82
I	0.09	0.43	0.53	0.47	0.29	0.85	1.00	0.38	0.00	0.70	0.83	0.42	0.20	0.83	0.60	0.82	0.90	0.90	0.60	0.90	0.91	0.90	0.80
J	0.73	0.85	0.88	0.86	0.79	0.67	1.00	0.83	0.70	0.00	0.60	0.80	0.78	0.60	0.83	0.50	1.00	1.00	1.00	1.00	1.00	0.67	0.33
K	0.85	0.77	0.73	0.69	0.80	0.20	1.00	0.75	0.83	0.60	0.00	0.82	0.80	0.40	0.86	0.25	1.00	1.00	1.00	1.00	1.00	0.75	0.50
L	0.46	0.50	0.59	0.53	0.56	0.83	1.00	0.46	0.42	0.80	0.82	0.00	0.30	0.82	0.56	0.80	0.89	0.89	0.56	0.89	0.78	0.89	0.78
M	0.27	0.33	0.47	0.38	0.43	0.82	1.00	0.27	0.20	0.78	0.80	0.30	0.00	0.80	0.50	0.78	0.88	0.88	0.50	0.88	0.89	0.88	0.75
N	0.85	0.77	0.81	0.79	0.80	0.50	1.00	0.75	0.83	0.60	0.40	0.82	0.80	0.00	0.86	0.25	1.00	1.00	1.00	1.00	1.00	0.75	0.50
O	0.64	0.67	0.73	0.69	0.71	0.88	1.00	0.64	0.60	0.83	0.86	0.56	0.50	0.86	0.00	0.83	0.75	1.00	0.67	0.75	0.80	0.75	0.80
P	0.83	0.75	0.80	0.77	0.79	0.40	1.00	0.73	0.82	0.50	0.25	0.80	0.78	0.25	0.83	0.00	1.00	1.00	1.00	1.00	1.00	0.67	0.33
Q	0.91	0.92	0.93	0.92	0.93	1.00	1.00	0.91	0.90	1.00	1.00	0.89	0.88	1.00	0.75	1.00	0.00	1.00	0.75	1.00	1.00	1.00	1.00
R	0.91	0.92	0.93	0.92	0.93	1.00	1.00	0.91	0.90	1.00	1.00	0.89	0.88	1.00	1.00	1.00	1.00	0.00	0.75	1.00	1.00	1.00	1.00
S	0.64	0.67	0.73	0.69	0.71	1.00	1.00	0.64	0.60	1.00	1.00	0.56	0.50	1.00	0.67	1.00	0.75	0.75	0.00	0.75	0.80	1.00	1.00
T	0.91	0.92	0.93	0.92	0.93	1.00	1.00	0.91	0.90	1.00	1.00	0.89	0.88	1.00	0.75	1.00	1.00	1.00	0.75	0.00	0.50	1.00	1.00
U	0.92	0.92	0.94	0.93	0.93	1.00	1.00	0.92	0.91	1.00	1.00	0.78	0.89	1.00	0.80	1.00	1.00	1.00	0.80	0.50	0.00	1.00	1.00
V	0.91	0.92	0.93	0.92	0.93	0.80	1.00	0.91	0.90	0.67	0.75	0.89	0.88	0.75	0.75	0.67	1.00	1.00	1.00	1.00	1.00	0.00	0.50
W	0.82	0.83	0.87	0.85	0.86	0.60	1.00	0.82	0.80	0.33	0.50	0.78	0.75	0.50	0.80	0.33	1.00	1.00	1.00	1.00	1.00	0.50	0.00

Table 4.1: In this table are the distances between each pair of vertices of the graph G. The values are not exact, they have been rounded to make the table readable.

From this distance matrix, one can compute clusters. We computed them using *Ward's minimum variance method*, that is a hierarchical clustering method already implemented in *R*. Any other method could be used, our approach does not depend on the clustering algorithm.

Using this method, we obtain 4 different clusters. They are depicted both in Figure 4.4 as a dendrogram and also in the example graph we have been working with, as can be seen in Figure 4.3. We obtain the clusters :

- $C_1 = \{F, K, N, P, J, W\}$ depicted in blue in Figure 4.3;
- $C_2 = \{C, D, E, B, H, L, M, A, I, O, S\}$ depicted in orange in Figure 4.3;
- $C_3 = \{T, U\}$ depicted in yellow in Figure 4.3;
- $C_4 = \{R, G, Q\}$ depicted in green in Figure 4.3.

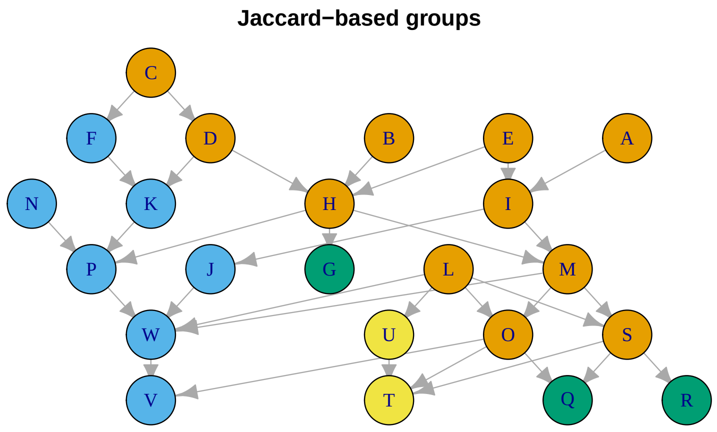


Figure 4.3: We represent the clusters computed using Ward's algorithm for the distance in Definition 4.4.1. The colours represent the different clusters.

From this we can determine the subgraphs induced by each of the clusters C_1, C_2, C_3, C_4. We obtain:

- $T_{C_1} = C_1$
- $T_{C_2} = C_2 \cup \{F, K, P, J, W, V\}$
- $T_{C_3} = C_3$
- $T_{C_4} = C_4$

And thus we can determine the vertices at the intersection of several/all clusters. In Figure 4.5 we coloured the edges of the graph G to point at vertices contained in multiple subgraphs induced by the different clusters. The method thus points at multiple vertices for being pivotal and keystone to maintain the cohesion of the ecosystem:

- vertices F, K, P, J, W, V connect the blue (C_1) and orange (C_2) clusters;
- vertices U, T connect the yellow (C_3) and orange (C_2) clusters;

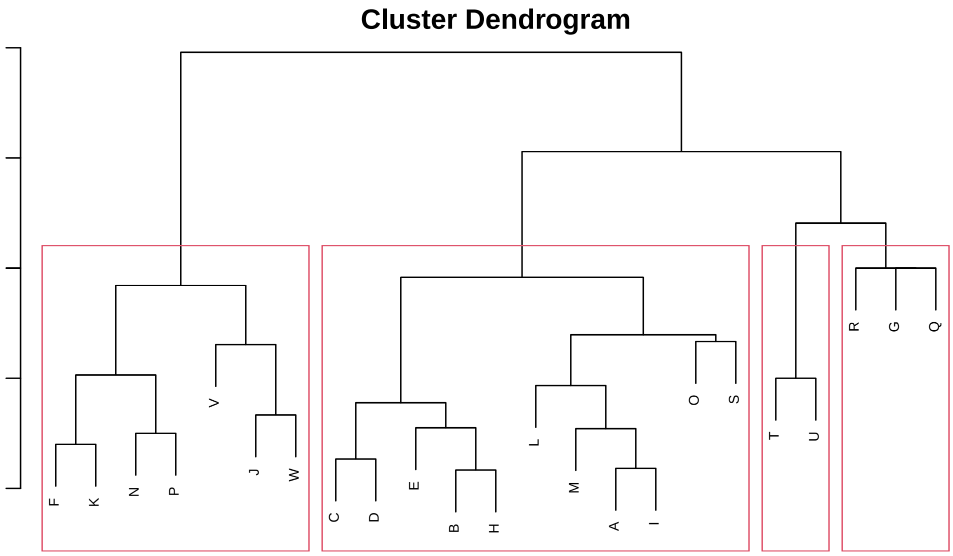


Figure 4.4: The clusters computed for the graph G depicted in Figure 4.2 and matrix distance in Table 4.1.

	W	V	U	T	S	R	Q	P	O	N	M	L	K	J	I	H	G	F	E	D	C	B	A
T_{C_1}	1	1	0	0	0	0	0	1	0	1	0	0	1	1	0	0	0	1	0	0	0	0	0
T_{C_2}	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1
T_{C_3}	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
T_{C_4}	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0

Table 4.2: The rows of this table represent the different subgraphs induced by each cluster and the columns are the vertices of the graph G.

- vertices R, Q, G connect the orange (C_2) and green (C_4) clusters;

These results are coherent with what we would expect by "eyeballing" the graphs. Lots of arrows point to vertices F, K, P, J, W, V in particular, so we know that a lot of the flow of biomass comes from this part of the graph.

4.5.2 Realistic setting

Now let's generate a more realistic graph, close to what can be obtained from Next Generation Sequencing, following Section 4.3.

First, a trophic network of potential interactions is generated with $B = 100$, $N = 1000$ and $L = 2 \times 10^5$. For all the nodes (species) of this trophic network, we randomly define an ecological optimum within a $2 - D$ niche space. For this example, the dimensions of our niche space are Temperature (in degree Celsius, with $T \in [-10, 50]$) and rainfall (in $mm/year$, with $R \in [0, 200]$. For each species i, the maximum increase rate is set to $\lambda_{\max} = 3$, the ecological optima from the niche space is randomly drawn from a uniform distribution and the variance–covariance matrix is set to be

$$\Sigma = \begin{bmatrix} 30 & 70 \\ 70 & 625 \end{bmatrix}$$

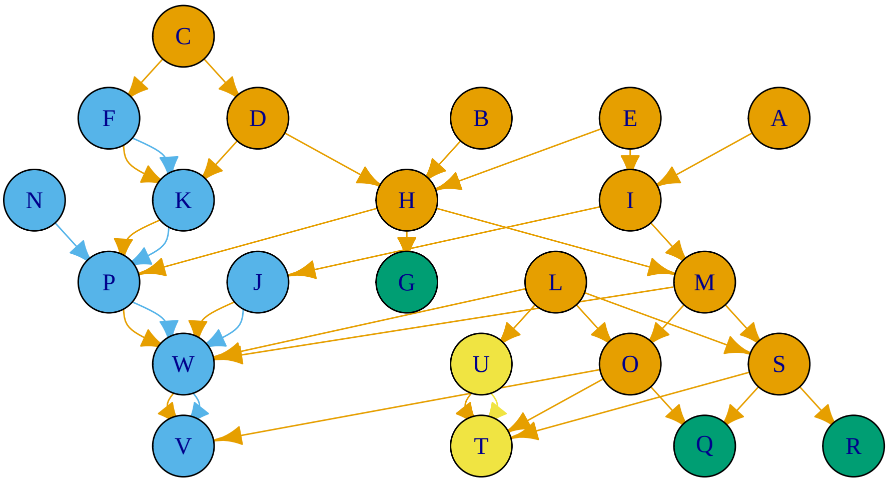


Figure 4.5: Vertices contained in multiple subgraphs induced by different clusters.

Thus, each species i is described as

$$S_i = \left\{ 3, \begin{bmatrix} T_i \sim \mathcal{U}_{[-10,50]} \\ R_i \sim \mathcal{U}_{[0,200]} \end{bmatrix}, \begin{bmatrix} 30 & 70 \\ 70 & 625 \end{bmatrix} \right\}$$

The interactions were defined such that the effect of preying is 0.3 and the effect of being a prey is -0.1, not to give a too strong penalty to the preys. The trophic relationships are given by the previously generated trophic network.

The realized niche (*i.e.*, the niche that a species actually occupies given the trophic constraints) is then computed, and an occurrence matrix is built such that a species is considered at one point in the niche space if its increase rate is above 1. Given this occurrence matrix we compute the co-occurrence between species according to [Vee13]. The output is an adjacency matrix. From this adjacency matrix, we discard the edges between nodes whose trophic level are discontinued. We then discard the nodes with 0 degree. This yields a network with 336 species and 2174 edges.

Ecologically speaking, this network is reasonably small, yet it is already unreadable (see Table 4.3).

We then computed the trophic distance as described in Algorithm 1, and using Ward's hierarchical clustering we retrieved 17 clusters. We then computed the subgraph induced by each cluster and computed, for each node, the number of clusters it belonged to as seen in fig. 4.6.

n clusters	1	2	3	4	5	6	7	8
n nodes	45	140	26	39	43	31	11	1

Table 4.3: Number of Nodes belonging to N clusters

We can thus easily identify which nodes are shared by the most clusters, and hence which nodes are the most heavily relied on. These nodes, in some ways, "carry" the network. For instance, 12 nodes are shared by at least 7 clusters. Removing these nodes induces losing 128

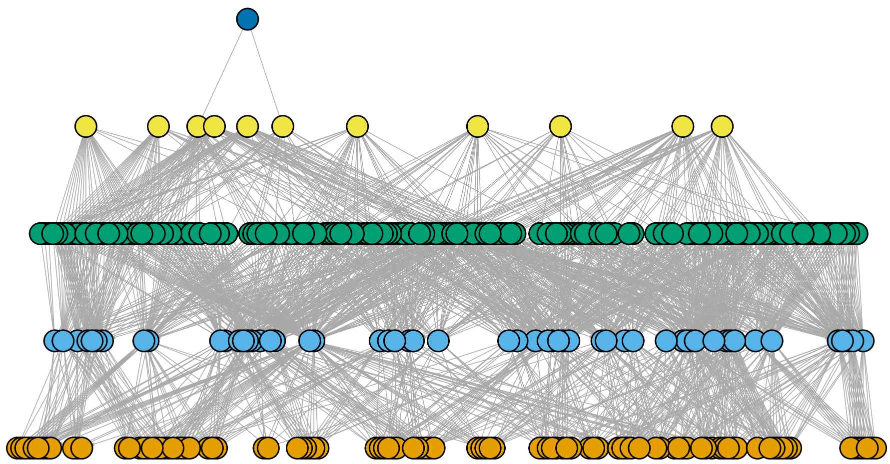


Figure 4.6: Artificial co-occurrence network. Nodes are coloured by trophic level. 336 nodes, 2174 edges

edges. In other words, 2.1% of the nodes are directly responsible for 5.9% of the edges.

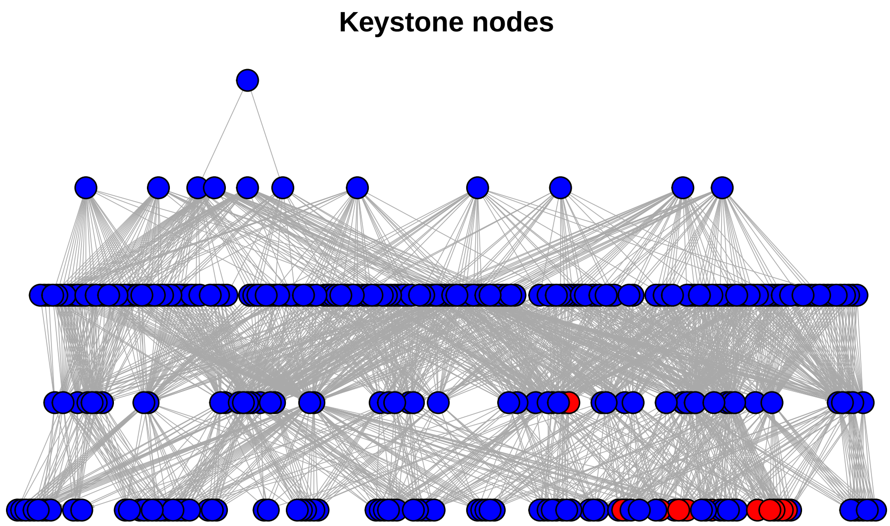


Figure 4.7: Keystone nodes. In red are the nodes that are shared by the induced subgraph of at least 7 clusters (maximum is 8)

List of Figures

Bibliography

[AK89] D. Arquès and P. Koch. Pavages tridimensionels. *Bigre+Globule*, 61-62:5–15, 1989.

[BCK17] Rémi Bottinelli, Laura Ciobanu, and Alexander Kolpakov. Three-dimensional maps and subgroup growth. *arXiv preprint arXiv:1712.01418*, 2017.

[BdMN10] A. Breda-d'Azevedo, A. Mednykh, and R. Nedela. Enumeration of maps regardless of genus: geometric approach. *Discr. Math.*, 310(6):1184–1203, 2010.

[BER+20] Marit F Markussen Bjorbækmo, Andreas Evenstad, Line Lieblein Røsæg, Anders K Krabberød, and Ramiro Logares. The planktonic protist interactome: where do we stand after a century of research? *The ISME journal*, 14(2):544–559, 2020.

[BGdPK20] Rémi Bottinelli, Laura Grave de Peralta, and Alexander Kolpakov. Telescopic groups and symmetries of combinatorial maps. *Algebraic Combinatorics*, 3(2):483–508, 2020.

[BNW20] Frédérique Bassino, Cyril Nicaud, and Pascal Weil. Statistics of subgroups of the modular group. *arXiv preprint arXiv:2004.00437*, 2020.

[Bol82] Béla Bollobás. The asymptotic number of unlabelled regular graphs. *Journal of the London Mathematical Society*, 2(2):201–206, 1982.

[CBBR+04] Marie-France Cattin, Louis-Félix Bersier, Carolin Banašek-Richter, Richard Baltensperger, and Jean-Pierre Gabriel. Phylogenetic constraints and adaptation explain food-web structure. *Nature*, 427(6977):835–839, 2004.

[CBN12] Joel E Cohen, Frédéric Briand, and Charles M Newman. *Community food webs: data and theory*, volume 20. Springer Science & Business Media, 2012.

[CG97] Fan RK Chung and Fan Chung Graham. *Spectral graph theory*. American Mathematical Soc., 1997.

[CK17] L. Ciobanu and A. Kolpakov. Free subgroups of free products and combinatorial hypermaps. *Discr. Math.*, 342(5):1415–1433, 2017.

[CLRS09] Thomas H Cormen, Charles E Leiserson, Ronald L Rivest, and Clifford Stein. *Introduction to algorithms*. MIT press, 2009.

[CM82] R. Cori and A. Machì. Construction of maps with prescribed automorphism group. *Theor. Comp. Sci.*, 21(1):91–98, 1982.

[CM92] R. Cori and A. Machì. Maps, hypermaps and their automorphisms: a survey, i, ii, iii. *Expositiones Mathematicae*, 10:403–427, 429–447, 449–467, 1992.

[Cor75] R Cori. Un code pour les graphes planaires et ses applications. *Astérisque*, 27:169 pp., 1975.

[DN12] M. Drmota and R. Nedela. Asymptotic enumeration of reversible maps regardless of genus. *Ars Math. Contemp.*, 5:77–97, 2012.

[DP+18] Valentina Disarlo, Hugo Parlier, et al. Simultaneous flips on triangulated surfaces. *Michigan Mathematical Journal*, 67(3):451–464, 2018.

[DP19] Valentina Disarlo and Hugo Parlier. The geometry of flip graphs and mapping class groups. *Transactions of the American Mathematical Society*, 372(6):3809–3844, 2019.

[EO19] Thomas R Etherington and O Pascal Omondiagbe. “virtualnicher“: generating virtual fundamental and realised niches for use in virtual ecology experiments. *Journal of Open Source Software*, 4(41):1661, 2019.

[Eps92] David BA Epstein. *Word processing in groups*. CRC Press, 1992.

[Est07] Ernesto Estrada. Characterization of topological keystone species: local, global and “meso-scale” centralities in food webs. *Ecological Complexity*, 4(1-2):48–57, 2007.

[Fei69] V. Z. Feinberg. Automorphism groups of trees. *Dokl. Akad. Nauk BSSR*, 13:1065–1067, 1969.

[FLNU18] Bailey K Fosdick, Daniel B Larremore, Joel Nishimura, and Johan Ugander. Configuring random graph models with fixed degree sequences. *Siam Review*, 60(2):315–355, 2018.

[FR12] Karoline Faust and Jeroen Raes. Microbial interactions: from networks to models. *Nature Reviews Microbiology*, 10(8):538–550, 2012.

[Fru38] R. Frucht. Herstellung von graphen mit vorgegebener abstrakter gruppe. *Compositio Math.*, 6:239–250, 1938.

[GMJ21] Catarina Gouveia, Ágnes Móréh, and Ferenc Jordán. Combining centrality indices: maximizing the predictability of keystone species in food webs. *Ecological Indicators*, 126:107617, 2021.

[Hat99] Allen Hatcher. Pants decompositions of surfaces. *arXiv preprint math/9906084*, 1999.

[HCRGE+19] Minhui Hao, J Javier Corral-Rivas, M Socorro González-Elizondo, K Narayanagowda Ganeshaiah, M Guadalupe Nava-Miranda, Chunyu Zhang, Xiuhai Zhao, and Klaus Von Gadow. Assessing biological dissimilarities between five forest communities. *Forest Ecosystems*, 6(1):1–8, 2019.

[JLR11] Svante Janson, Tomasz Luczak, and Andrzej Rucinski. *Random graphs*, volume 45. John Wiley & Sons, 2011.

[Jon18] G. A. Jones. Realisation of groups as automorphism groups in permutational categories. *arXiv:1807.00547*, 2018.

[Jor09] Ferenc Jordan. Keystone species and food webs. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1524):1733–1741, 2009.

[JS78] G. A. Jones and D. Singerman. Theory of maps on orientable surfaces. *Proc. LMS*, 3(2):273–307, 1978.

[Kap97] Ilya Kapovich. Quasiconvexity and amalgams. *International Journal of Algebra and Computation*, 7(06):771–811, 1997.

[KJ17] Janis Klaise and Samuel Johnson. The origin of motif families in food webs. *Scientific reports*, 7(1):1–11, 2017.

[KM02] I. Kapovich and A. Myasnikov. Stallings foldings and subgroups of free groups. *J. Algebra*, 248(2):608–668, 2002.

[KMW17] Olga Kharlampovich, Alexei Miasnikov, and Pascal Weil. Stallings graphs for quasi-convex subgroups. *Journal of Algebra*, 488:442–483, 2017.

[Lee16] Paul-Henry Leemann. Schreir graphs: Transitivity and coverings. *Internat. J. Algebra Comput.*, 26(1):69–93, 2016.

[Lub95] Alexander Lubotzky. Counting finite index subgroups. In *Groups*, volume 93, pages 368–404, 1995.

[LZ04] S. K. Lando and A. K. Zvonkin. *Graphs on surfaces and their applications*. Springer, 2004.

[MN06] A. Mednykh and R. Nedela. Enumeration of unrooted maps of a given genus. *J. Combin. Theory, Ser. B*, 96(5):706–729, 2006.

[MN10] A. Mednykh and R. Nedela. Enumeration of unrooted hypermaps of a given genus. *Discr. Math.*, 310(3):518–526, 2010.

[RT+13] Kasra Rafi, Jing Tao, et al. The diameter of the thick part of moduli space and simultaneous whitehead moves. *Duke Mathematical Journal*, 162(10):1833–1876, 2013.

[SAM+20] Lisa K Schneider, Konstantinos Anestis, Joost Mansour, Anna A Anschütz, Nathalie Gypens, Per J Hansen, Uwe John, Kerstin Klemm, Jon Lapeya Martin, Nikola Medic, et al. A dataset on trophic modes of aquatic protists. *Biodiversity data journal*, 8, 2020.

[Ser03] J.-P. Serre. *Trees*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2003.

[Spe91] J.-C. Spehner. Merging in maps and in pavings. *Theoret. Comput. Sci.*, 86:205–232, 1991.

[ŠŠ93] J. Širáň and M. Škoviera. Orientable and non-orientable maps with given automorphism groups. *Australasian J. Comb.*, 7:47–53, 1993.

[Sta83] J. R. Stallings. Topology of finite graphs. *Invent. Math.*, 71(3):551–565, 1983.

[Sto78] W. W. Stothers. Free subgroups of the free product of cyclic groups. *Math. Comp.*, 32(144):1274–1280, 1978.

[Tut73] W. T. Tutte. What is a map? In F. Harary, editor, *New Directions in Graph Theory*, pages 309–325. Academic Press, New York, 1973.

[Vee13] Joseph A Veech. A probabilistic model for analysing species co-occurrence. *Global Ecology and Biogeography*, 22(2):252–260, 2013.

[Wel17] C. Welsch. Subgroup graph methods for presentations of finitely generated groups and the contractibility of associated simplicial complexes. *J. Algebra*, 486:119–156, 2017.

[WM00] Richard J Williams and Neo D Martinez. Simple rules yield complex food webs. *Nature*, 404(6774):180–183, 2000.