WHY ARE BRAIDS ORDERABLE?

Patrick Dehornoy, Ivan Dynnikov
Dale Rolfsen, Bert Wiest

Panoramas et Synthèses

Numéro 14 2002

SOCIÉTÉ MATHÉMATIQUE DE FRANCE
Publié avec le concours du Centre National de la Recherche Scientifique et du Ministère de la culture et de la communication (aide de la délégation générale à la langue française)

Patrick DEHORNOY
Ivan DYNNIKOV
Dale ROLFSEN
Bert WIEST

WHY ARE BRAIDS ORDERABLE?

Patrick DEHORNOY

Laboratoire de Mathématiques Nicolas Oresme, CNRS UMR 6139, Université de Caen, 14032 Caen, France.

E-mail : dehornoy@math.unicaen.fr

Url : //www.math.unicaen.fr/∼dehornoy

Ivan DYNNIKOV

Dept. of Mechanics and Mathematics Moscow State University; Moscow 119992 GSP-2, Russia.

E-mail : dynnikov@mech.math.msu.su

Url : //mech.math.msu.su/∼dynnikov

Dale ROLFSEN

Mathematics Department; Univ. of British Columbia; Vancouver BC, V6T 1Z2 Canada.

E-mail : rolfsen@math.ubc.ca

Url : //www.math.ubc.ca/∼rolfsen

Bert WIEST

IRMAR, Université de Rennes 1, Campus Beaulieu, 35042 Rennes, France.

E-mail : bertw@maths.univ-rennes1.fr

Url : //www.maths.univ-rennes1.fr/∼bertw

2000 *Mathematics Subject Classification.* — 20F36, 06F15, 20C40, 20N02, 57M60, 57M07.

Key words and phrases. — Braid groups; ordered groups; mapping class groups; self-distributive systems; finite trees; well-orderings; free group automorphisms; automatic groups; laminations; hyperbolic geometry.

The work of Ivan Dynnikov was supported by RFBR (grant # 99-01-00090).

WHY ARE BRAIDS ORDERABLE?

Patrick DEHORNOY, Ivan DYNNIKOV, Dale ROLFSEN, Bert WIEST

Abstract. — In the decade since the discovery that Artin's braid groups enjoy a left-invariant linear ordering, several quite different approaches have been applied to understand this phenomenon. This book is an account of those approaches, involving self-distributive algebra, uniform finite trees, combinatorial group theory, mapping class groups, laminations, and hyperbolic geometry.

Environ dix ans ont passé depuis la découverte du caractère ordonnable des groupes de tresses, et des méthodes diverses ont été proposées pour expliquer le phénomène. Le but de ce texte est de présenter ces approches variées, qui mettent en jeu l'algèbre auto-distributive, les arbres finis, la théorie combinatoire des groupes, les groupes de difféomorphismes, la théorie des laminations, et la géométrie hyperbolique.

Ein Jahrzehnt ist vergangen seit der Entdeckung, dass Artins Zopfgruppen eine links-invariante Ordnung besitzen, und verschiedene Methoden wurden in der Zwischenzeit vorgeschlagen, um zu einem tieferen Verständnis dieses Phänomens zu gelangen. Ziel dieses Buches ist es, ein Resümee dieser Techniken zu geben. Selbstdistributive Algebren, endliche Bäume, kombinatorische Gruppentheorie, Abbildungsklassengruppen, Laminationen, und hyperbolische Geometrie kommen dabei zusammen.

За десять лет, прошедшие после открытия, что артиновские группы кос обладают левоинвариантным линейным порядком, возник целый ряд различных подходов для объяснения этого явления. Данная книга посвящена описанию этих подходов, которые основаны на самодистрибутивных операциях, теории однородных конечных деревьев, комбинаторной теории групп, группах классов отображений, ламинациях и гиперболической геометрии.

CONTENTS

Introduction vii
An idea whose time was overdue. vii
The importance of being orderable ix
Organization of the text x
Guidelines to the reader xii

1. A linear ordering of braids 1
1.1. Braid groups 1
1.2. A linear ordering on B_n 7
1.3. Applications 17

2. Self-distributivity 23
2.1. The action of braids on LD-systems 24
2.2. Special braids 34
2.3. The group of left self-distributivity 40
2.4. Normal forms in free LD-systems 49
2.5. Appendix: Iterations of elementary embeddings in set theory 51

3. Handle reduction 55
3.1. Handles 55
3.2. Convergence of handle reduction 57
3.3. Implementations and variants 66

4. Finite trees 71
4.1. Encoding positive braid words in trees 72
4.2. A well-ordering on positive braid words 75
4.3. Reduction of positive braid words 78
4.4. Applications 87

5. Automorphisms of a free group 93
5.1. Keeping track of the letters 93

5.2. From an automorphism back to a braid 99

6. Curve diagrams 103
6.1. Mapping class groups and curve diagrams 103
6.2. A braid ordering using curve diagrams 105
6.3. Generalizations 111

7. Hyperbolic geometry 113
7.1. Uncountably many orderings of the braid group 113
7.2. The classification of orderings induced by the action on $\mathbb{R}$ 123
7.3. A proof of Property **S** for all Thurston-type orderings 130
7.4. A combinatorial approach and an extension to B_∞ 131

8. Triangulations 135
8.1. Singular triangulations 136
8.2. The Mosher normal form 142
8.3. Laminations 150
8.4. Integral laminations 152
8.5. Action of the braid group on laminations 155

9. Bi-ordering the pure braid groups 163
9.1. Descending central series 163
9.2. An effective ordering on P_n 164
9.3. Incompatibility of the orderings: local indicability 170

10. Open questions 173
10.1. Well-ordering on positive braids 173
10.2. Finding σ-positive representatives 175
10.3. Topology of the space of orderings 177
10.4. Generalizations and extensions 178

Bibliography 181

Index 191

Index of Notation 193

INTRODUCTION

An idea whose time was overdue.

This book is about braids and orderings. The braid groups B_n were introduced by Emil Artin [**2**] in 1925 (see also [**3**]) and have been studied intensively ever since [**18, 21**]. Indeed, many of the ideas date back to the 19th century, in the works of Hurwicz, Klein, Poincaré, Riemann, and certainly other authors. One can even find a braid sketched in the notebooks of Gauss [**67**]—see [**124**] for a discussion about Gauss and braids, including a reproduction of the picture he drew in his notebook. The n-strand braid group B_n has the well-known presentation (other definitions will be given later):

$$B_n = \langle \sigma_1, \dots, \sigma_{n-1} \,;\, \sigma_i\sigma_j = \sigma_j\sigma_i \text{ for } |i-j| \geqslant 2,\ \sigma_i\sigma_j\sigma_i = \sigma_j\sigma_i\sigma_j \text{ for } |i-j| = 1\rangle.$$

We use B_n^+ for the monoid with the above presentation, which is called the n-strand braid monoid. To each braid, there is an associated permutation of the set $\{1, \dots, n\}$, with $\sigma_i \mapsto (i, i+1)$, defining a homomorphism $B_n \to \mathfrak{S}_n$, where $\mathfrak{S}_n$ denotes the symmetric group on n objects. The kernel of this mapping is the *pure* braid group P_n.

The theory of ordered groups is also well over a hundred years old. One of the basic theorems of the subject is Hölder's theorem, published in 1902 [**76**], that characterizes the additive reals as the unique maximal Archimedian ordered group. It is remarkable, and somewhat puzzling, that it has taken so long for these two venerable subjects to come together as they now have.

A group or a monoid G is *left-orderable* if there exists a strict linear ordering $<$ of its elements which is left-invariant: $g < h$ implies $fg < fh$ for all f, g, h in G. If, in addition, the ordering is a well-ordering, we say that G is *left-well-orderable*. If there is an ordering of G which is invariant under multiplication on both sides, we say that G is *orderable*, or for emphasis, *bi-orderable*. This book is devoted to explaining the following results, discovered within the last decade:

Theorem I.1. — *The Artin braid group B_n is left-orderable, by an ordering which is a well-ordering when restricted to B_n^+.*

Theorem I.2. — *The pure braid group P_n is bi-orderable, by an ordering which is a well-ordering when restricted to $P_n \cap B_n^+$.*

The braid groups have been an exceptionally active mathematical subject in recent decades. The field exploded in the mid 1980's with the revolutionary discoveries of Vaughan Jones [**79**], providing strong connections with operator theory, statistical mechanics and other notions of mathematical physics. Great strides have been made in recent years in understanding the very rich representation theory of the braid groups. The classical Burau representation [**13**] was shown to be unfaithful [**109**], [**98**]. The long-standing question of whether the braid groups are linear (isomorphic to finite-dimensional matrix groups) was recently answered in the affirmative, by Daan Krammer [**85, 86**] and Stephen Bigelow [**6**].

Despite the high degree of interest in braid theory, the importance of the left-orderability of the braid groups, announced in 1992 [**34**], was not widely recognized at first. A possible explanation for this is that the methods of proof were rather unfamiliar to most topologists, the people most interested in braid theory. As will be seen in Chapter 2, that proof involves rather delicate combinatorial and algebraic constructions, which were partly motivated by (while being logically independent of) questions in set theory — see [**82**] for a good introduction. Subsequent combinatorial work brought new results and proposed new approaches: David Larue established in [**89, 88**] results anticipating those of [**60**], Richard Laver proved in [**94**] that the restriction of the braid ordering to B_n^+ is a well-ordering (presumably the deepest result known so far about the braid order), Serge Burckel gave an effective version of the latter result in [**14, 15**]. However, these results were also not widely known for several years.

The challenge of finding a topological proof of left-orderability of B_n led to the five-author paper [**60**], giving a completely different construction of an ordering of B_n as a mapping class group. Remarkably, it leads to exactly the same ordering as [**34**]. Soon after, a new technique [**131**] was applied to yield yet another proof of orderability of the braid groups (and many other mapping class groups), using ideas of hyperbolic geometry, and moreover giving rise to many possible orderings of the braid groups. This argument, pointed out by William Thurston, uses ideas of Nielsen [**112**] from the 1920's. It is interesting to speculate whether Nielsen himself might have solved the problem, if asked whether braid groups are left-orderable in the following language: Does the mapping class group of an n-punctured disk act effectively on the real line by order-preserving homeomorphisms? Nielsen had laid all the groundwork for an affirmative answer.

More recently, a new topological approach using laminations was proposed, one that is also connected with the Mosher normal form based on triangulations [**110**]. Also, a combinatorial interpretation of the results of [**131**] was proposed by Jonathon Funk in [**65**], including a connection with the theory of topoi.

The braid groups are known to be automatic [**138**]. Without burdening the reader with technical details, it should be mentioned that the ordering of B_n and certain

other surface mapping class groups (nonempty boundary) can be considered automatic as well, meaning roughly that it may be determined by some finite-state automaton [**129**].

Theorem I.2 appeared in [**83**], and it relies on a completely different approach, namely using the Magnus representation of a free group. Subsequent work [**128**] has shown how different general braid groups B_n and the pure braid groups P_n are from the point of view of orderability: in particular, for $n \geqslant 5$, the group B_n is not locally indicable, which implies that it is not bi-orderable in a strong sense, namely that no left-ordering of B_n can bi-order a subgroup of finite index, such as P_n [**126**].

The importance of being orderable

As will be recalled in Chapter 1, the orderability of a group implies various structural consequences about that group and derived objects. The fact that B_n is left-orderable implies that it is torsion-free, which had been well known. However, it also implies that the group ring $\mathbb{Z}B_n$ has no zero-divisors, which was a natural open question. Biorderability of P_n shows that $\mathbb{Z}P_n$ embeds in a skew field. In addition, it easily implies that the group P_n has unique roots, a result proved in [**4**] by complicated combinatorial arguments, and definitely not true for B_n.

One may argue that such general results did not dramatically change our understanding of braid groups. The main point of interest, however, is not—or not only—the mere existence of orderings on braid groups, but the particular nature and variety of the constructions we shall present. Witness the beautiful way the order on P_n is deduced from the Magnus expansion in Chapter 9, the fascinating connection between the uncountable family of orderings on B_n constructed in Chapter 7 and the Nielsen–Thurston theory, and, chiefly, the specific properties of one particular ordering on B_n. Here we refer to the ordering of B_n sometimes called the Dehornoy ordering in literature, which will be called the σ-ordering in this text.

Typically, it is the specific form of the braids greater than 1 in the σ-ordering that led to the new, efficient algorithm for the classical braid isotopy problem described in Chapter 3, and motivated the further study of the algorithms described in Chapters 6 and 8. But what appears to be of the greatest interest here is the remarkable convergence of many approaches to one and the same object: at least six different points of view end up today with the σ-ordering of braids, and this, in our opinion, is the main hint that this object has an intrinsic interest. Just to let the reader feel the flavour of some of the results, we state below various characterizations of the σ-ordering—the terms will be defined in the appropriate place. So, the braid β_1 is smaller than the braid β_2 in the σ-ordering if and only if

- (*in terms of braid words*) the braid $\beta_1^{-1}\beta_2$ has a braid word representative where the generator σ_i with smallest index i appears only positively (no σ_i^{-1});
- (*in terms of action on self-distributive systems*) for some/any ordered LD-system $(S, *, <)$, and for some/any sequence $\vec{x}$ in S, we have $\vec{x} \cdot \beta_1 <^{\mathrm{Lex}} \vec{x} \cdot \beta_2$;
- (*in terms of braid words combinatorics*) any sequence of handle reductions from any braid word representing $\beta_1^{-1}\beta_2$ ends up with a σ-positive word;

– (*in terms of trees, assuming* β_1 *and* β_2 *to be positive braids*) the irreducible uniform tree associated with β_1 is smaller than the one associated with β_2;
– (*in terms of automorphisms of a free group*) for some i, the automorphism associated with $\beta_1^{-1}\beta_2$ maps x_j to x_j for $j < i$, and it maps x_i to a word that ends with x_i^{-1};
– (*in terms of mapping class groups*) the standardized curve diagram associated with β_1 first diverges from the one associated with β_2 towards the left;
– (*in terms of hyperbolic geometry*) the endpoint of the lifting of $\beta_1(\gamma_a)$ is larger (as a real number) than the endpoint of the lifting of $\beta_2(\gamma_a)$;
– (*in terms of free group ordering*) we have $\beta_1 \cdot z \lhd \beta_2 \cdot z$ in $\widehat{F_\infty}$;
– (*in terms of Mosher's normal form*) the last flip in the normal form of $\beta_2^{-1}\beta_1$ occurs in the upper half-plane.
– (*in terms of laminations*) the first nonzero coefficient of odd index in the sequence $\beta_1^{-1}\beta_2 \cdot (0, 1, \ldots, 0, 1)$ is positive.

Even if the various constructions of the σ-ordering depend on choosing a particular family of generators for the braid groups, namely the Artin generators σ_i, this convergence might suggest to call this ordering canonical or, at least, standard. This convergence is the very subject of this text: our aim here is not to give a complete study of any of the different approaches—so, in particular, our point of view is quite different from that of [**40**] which more or less exhausts the combinatorial approaches—but to try to let the reader feel the flavour of these different approaches. More precisely—and with the exceptions of Chapter 7 which deals with more general orderings, and of Chapter 9 which deals with ordering pure braids—our aim will be to describe the σ-ordering of braids in the various possible frameworks: algebraic, combinatorial, topological, geometric, and to see which properties can be established by each technique. As explained in Chapter 1, exactly three properties of braids, called **A**, **C**, and **S** here, are crucial to prove that the σ-ordering exists and to establish its main properties. Roughly speaking, each chapter of the subsequent text (except Chapters 7 and 9) will describe one possible approach to the question of ordering the braids, and, in each case, explain which of the properties **A**, **C**, and **S** can be proved: some approaches are relevant for establishing all three properties, while others enable us only to prove one or two of them, possibly assuming some other one already proved.

Organization of the text

Various equivalent definitions of the braid groups are described in Chapter 1, which also includes a general discussion of orderable groups and their rather special algebraic properties. The well-ordering of B_n^+ is also introduced in this chapter.

The remaining chapters contain various approaches to the orderability phenomenon. The combinatorial approaches are gathered in Chapters 2 to 5, while the topological approaches are presented in Chapters 6 to 8.

Chapter 2 introduces left self-distributive algebraic systems (LD-systems) and the action of braids upon such systems. This is the technique whereby the orderability of braids was first demonstrated and the σ-ordering introduced. The chapter sketches a

self contained proof of left-orderability of B_n, by establishing Properties **A**, **C** (actually details are given only for its weak variant $\mathbf{C}_\infty$) and **S** with arguments utilizing LD-systems. Here we consider colourings of the strands of the braids, and observe that the braid relations dictate the self-distributive law among the colours. Then we can order braids by choosing orderable LD-systems as colours, a simple idea yet the existence of an orderable LD-system requires an indirect argument. The chapter concludes with a discussion of the historical origins of orderable LD-systems, which arise in the study of elementary embeddings in the foundations of set theory.

A combinatorial algorithm called handle reduction is the subject of Chapter 3. This procedure, which extends the idea of word reduction in a free group, is a very efficient procedure in practice for determining whether a braid word represents a braid larger than 1, and incidentally gives a rapid solution to the word problem in the braid groups. Handle reduction gives an alternative proof of Property **C**, under the assumption that Property **A** holds.

Another combinatorial technique, due to Serge Burckel, is to encode positive braid words by finite trees. This is the subject of Chapter 4, in which one proves that the restriction of the σ-ordering to B_n^+ is a well-ordering by considering a natural ordering of the associated trees and using a tricky transfinite induction. This approach provides arguments for Properties **C** and **S**, however assuming (as with handle reduction) Property **A**. The advantage of the method is that it assigns a well-defined ordinal to each braid in B_n^+. By using a variant of the σ-ordering, one obtains a well-ordering of B_∞^+.

Chapter 5 contains an approach to the σ-ordering using a very classical fact, that the braid groups can be realized as a certain group of automorphisms of a free group. As observed by David Larue, this method yields a quick proof of Property **A**, an alternative proof of Property **C** (hence an independent proof of left-orderability of B_∞) and a simple criterion for recognizing whether a braid is σ-positive, in terms of its action on the free group.

We begin the topological description of the σ-ordering in Chapter 6. Here we realize B_n as the mapping class group of a disk with n punctures. The braid action can be visualized by use of curve diagrams which provide a canonical form for the image of the real line, if the disk is regarded as the unit complex disk. This was the first geometric argument for the left-orderability of the braid groups, and it is remarkable that the ordering described in this way is identical with the original, *i.e.*, with the σ-ordering. An advantage of this approach is that it also applies to more general mapping class groups. We emphasize that Chapters 5 and 6 are based on very similar ideas, except that the first one is algebraic while the second is more geometric.

The discussion in Chapter 7 interprets braid orderings in terms of Nielsen–Thurston theory. The key observation is that the universal cover of the punctured disk has a natural embedding in the hyperbolic plane. Thereby, braids act on a family of hyperbolic geodesics, which have a natural ordering. This point of view provides an infinitude of inequivalent orderings of braid groups and many other mapping class groups. The σ-ordering on B_n corresponds to choosing a particular geodesic in $\mathbb{H}^2$. We also outline in this chapter the interpretation developed by Jonathon Funk, in

which a certain linear ordering of words in the free group is preserved under the braid automorphisms as considered in Chapter 5.

Chapter 8 continues the discussion of the σ-ordering in terms of mapping classes. However, here, the geometric approach is rephrased in combinatorial terms by use of two somewhat different devices involving triangulations. The first was inspired by the technique employed by Lee Mosher to establish that mapping class groups are automatic. It develops a new canonical form for braids and a method for determining σ-ordering by means of a finite state automaton. The second approach, developed in [**54**], uses integral laminations. One encodes the action of a braid on the disk by counting intersections of the image of a certain triangulation with a lamination. This leads to an independent proof of Property **A** and yet another characterization of braids larger than 1 in the σ-ordering.

The final chapter is an account of an ordering of the *pure* braid groups. Unlike the full braid groups, the groups P_n of pure braids can be given an ordering which is invariant under multiplication on both sides. This ordering is defined algebraically, using the Artin combing technique, together with a specific ordering of free groups using the Magnus expansion. By appropriate choice of conventions, this ordering has the property that braids in $P_n \cap B_n^+$ are larger than 1 and well-ordered. The chapter ends with a discussion showing that any two-sided ordering of P_n is necessarily incompatible with every left-ordering of B_n for $n \geqslant 5$.

Guidelines to the reader

There are many sorts of readers who may be interested in this text, with various styles of mathematical understanding. Thus different approaches are bound to appeal to different readers. The reader with a mostly algebraic or combinatorial culture may feel uncomfortable with informal definitions in the geometric constructions of Chapters 6, 7, or 8, while another reader coming from the world of topology or geometry may find Chapter 2 and, even more, Chapter 4 quite mysterious, and lacking conceptual understanding in the case of the latter. It is impossible to claim that one approach is definitely better than another, as every one of them brings some specific result or intuition that is so far inaccessible to the others. An attempt has been made to keep the chapters relatively self-contained; so, apart from Chapter 1, all chapters are parallel one to the other rather than logically interdependent, and, therefore, from Chapter 2, the reader can take the chapters essentially in whatever order he or she likes.

We mentioned that three properties of braids called **A**, **C**, and **S** play a crucial role, and that our main task in this text will be to prove these properties using various possible approaches. In spite of the above general remarks, it might be useful that we propose answers to the question: which of these approaches offers the quickest, or the most elementary, proof of Properties **A**, **C**, and **S**? The answer depends of course on the mathematical preferences of the reader. As for Property **A**, the shortest proofs are the one using the automorphisms of a free group in Chapter 5, and—even shorter once the formulas (8.5.14) have been guessed—the one using laminations in Chapter 8. The argument involving self-distributivity by contrast is more conceptual

and naturally connected with orderings, but it is technically quite involved. As for Property **C**, the shortest argument is probably the one involving self-distributivity as outlined in Chapter 2, but one may prefer the approach through handle reduction, which uses nothing exotic and gives an efficient algorithm in addition, or the curve diagram approach of Chapter 6, which gives a less efficient method and requires considerable effort to be made rigorous, but appeals to a natural geometric intuition. Finally, for Property **S**, the hyperbolic geometry argument of Chapter 7 is probably the more interesting one, as it gives the result not only for the σ-ordering, but also for a whole family of different orderings. On the other hand, even if it may appear conceptually less satisfactory in its present exposition, the combinatorial approach of Chapter 4 gives the most precise and effective version of Property **S**.

Although they are conceptually simple, the braid groups are very subtle nonabelian groups which have given up their secrets only reluctantly over the years. They will undoubtedly continue to supply us with surprises and fascination, and so will in particular their orderings: despite the many approaches and results mentioned in the text below, a lot of questions about braid orderings remain open today, and further developments can be expected. For the moment, we hope that this small text, which involves techniques of algebra, combinatorics, hyperbolic geometry, topology, and has even a loose connection with set theory, can illuminate some facets of the question, "Why are braids orderable?"

The preparation of this text was coordinated by P.D.; Chapters 2 to 4 have been mostly written by P.D., Chapters 6 and 7 by B.W., Chapter 8 by I.D., and Chapter 9 by D.R.; the other chapters are common work of two or more authors.

We thank all colleagues and friends who have suggested corrections or improvements, specially Roger Fenn and Christian Kassel. We also thank Hervé Sibert for his help in practical implementations, and the referees for their valuable suggestions.

CHAPTER 1

A LINEAR ORDERING OF BRAIDS

In this chapter, we briefly recall some of the standard definitions of the braid groups. Then we introduce three basic properties of braids, called **A**, **C**, and **S** in the sequel, and we show how they allow us to define a linear ordering of braids that is compatible with multiplication on one side. Finally, we describe some general properties of this ordering.

1.1. Braid groups

As mentioned before, the Artin braid group on n strands, denoted by B_n, is defined by the presentation

$$B_n = \langle \sigma_1, \dots, \sigma_{n-1} \,;\, \sigma_i\sigma_j = \sigma_j\sigma_i \text{ for } |i-j| \geqslant 2,\ \sigma_i\sigma_j\sigma_i = \sigma_j\sigma_i\sigma_j \text{ for } |i-j| = 1 \rangle.$$

The braid group on infinitely many strands, denoted B_∞, is defined by a presentation with infinitely many generators $\sigma_1, \sigma_2, \dots$ subject to the same relations.

The aim of this section is to show how the groups B_n arise in several different ways as special cases of some natural mathematical objects in geometry and algebra.

1.1.1. Isotopy classes of braid diagrams. — Let D^2 be the unit disk with centre 0 in the complex plane $\mathbb{C}$, and let D_n be the disk D^2 with n regularly spaced points in the real axis as distinguished points; we call these points the puncture points of D^2.

Definition 1.1.1. — We define an *n-strand geometric braid* to be the image of an embedding b of the disjoint union $\coprod_{j=1}^{n}[0_j, 1_j]$ of n copies of the interval $[0,1]$ into the cylinder $[0,1] \times D^2$ satisfying the following properties: (i) for t in $[0_j, 1_j]$, the point $b(t)$ lies in $\{t\} \times D^2$; (ii) the set $\{b(0_1), \dots, b(0_n)\}$ is the set of punctures of $\{0\} \times D^2$, and similarly the set $\{b(1_1), \dots, b(1_n)\}$ is the set of punctures of $\{1\} \times D^2$.

The image of each interval is called a strand of the braid; the idea is that we have n strands running continuously from left to right (visualizing the unit interval as being horizontal), intertwining, but not meeting each other. The set of starting points of strands on the left face of the cylinder is the same as the set of endpoints on the right face, and the n strands induce in an obvious way a permutation of these puncture points of D^2. If this permutation happens to be the trivial permutation, we say that b is a *pure* geometric braid.

We define two geometric braids b, b' to be *isotopic* if there is a continuous $[0,1]$-family of geometric braids b_t with $b_0 = b$ and $b_1 = b'$. The geometric idea is that we can deform one braid into the other while holding their endpoints fixed. Note that isotopic braids induce the same permutation of the set of puncture points.

Define a *braid* to be an isotopy class of geometric braids. The set of braids has a group structure, given by concatenation: given two braids β_1 and β_2, we squeeze the image of β_1 into the cylinder $[0, \frac{1}{2}] \times D^2$, the image of β_2 into $[\frac{1}{2}, 1] \times D^2$, and obtain a new, well-defined braid $\beta_1 \cdot \beta_2$. The trivial element of the group is represented by the trivial braid (whose strands are just straight line segments, not intertwining each other), and the inverse of a braid is its own reflection in the disk $\{\frac{1}{2}\} \times D^2$. We shall see in a moment that this group is isomorphic to the braid group B_n defined above. The function which to every braid associates the permutation it induces on the set of punctures yields then a homomorphism of B_n to the symmetric group $\mathfrak{S}_n$ (in order to obtain a homomorphism, and not an anti-homomorphism, we must define the permutation associated with a braid so that it specifies the initial positions of the punctures in terms of their final positions). The set of pure braids (the kernel of this homomorphism), forms a normal subgroup, called the *pure braid group*, here denoted P_n.

Let us outline a proof that the group of n-strand braids in the cylinder is indeed isomorphic to the group B_n. To that end, we define a homomorphism from B_n to our group of braids by sending the generator σ_i to the clockwise half-twist braid involving the ith and $(i+1)$st strand indicated in Figure 1.1 (in this picture, the cylinder has not been drawn, for simplicity, and our picture represents a side-view of the braid).

The figure also illustrates the fact that this homomorphism is well-defined. Indeed, in our group of braids we have that crossings which are far apart commute (so that we have $\sigma_i\sigma_j = \sigma_j\sigma_i$ for $|i-j| \geqslant 2$), and the Reidemeister III-type relation $\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}$ holds.

We leave it to the reader to verify that this homomorphism is surjective (any geometric braid can be deformed into one in which a side-view offers only finitely many crossings, all of which are transverse) and injective (our two types of relations suffice to transform into each other any two braid diagrams representing isotopic geometric braids). The proofs, which we shall not discuss here, are similar to the

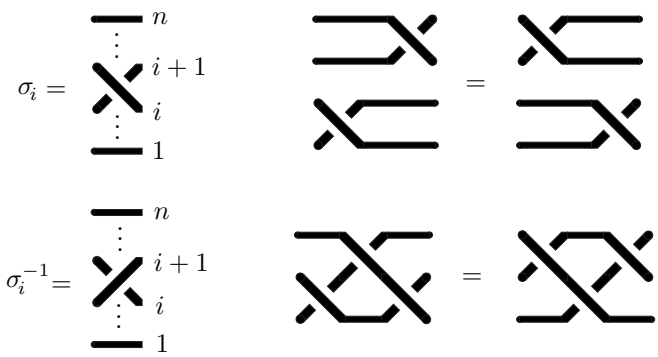


FIGURE 1.1. The generators and relations of the braid group

proof that isotopy classes of knots are the same as knot diagrams up to Reidemeister-equivalence [**125**].

This completes the definition of B_n in terms of geometric n-strand braids. We remark that this definition has a very natural generalization: we can replace the disk D^2 by any compact surface $\mathcal{S}$, possibly with boundary. Choosing n puncture points in $\mathcal{S}$, we can define the *n-strand braid group of the surface* $\mathcal{S}$, denoted $B_n(\mathcal{S})$, to be the group of n-strand braids in $\mathcal{S} \times [0,1]$, in exact analogy with the previous construction. Presentations for these braid groups are known [**7, 130, 69**].

1.1.2. Mapping class groups. — The aim of this section is to identify the braid group B_n with the group of homotopy classes of self-homeomorphisms of an n-punctured disk. The idea is simply to look at braids from one end rather than from the side.

Let $\mathcal{S}$ be an oriented compact surface, possibly with boundary, and $\mathcal{P}$ be a finite set of distinguished interior points of $\mathcal{S}$. The most important example in this section will be the n-punctured disk D_n.

Definition 1.1.2. — The *mapping class group* $\mathcal{MCG}(\mathcal{S},\mathcal{P})$ of the surface $\mathcal{S}$ relative to $\mathcal{P}$ is the group of all isotopy classes of orientation-preserving homeomorphisms $\varphi\colon \mathcal{S} \to \mathcal{S}$ satisfying $\varphi|_{\partial\mathcal{S}} = \mathrm{id}$ and $\varphi(\mathcal{P}) = \mathcal{P}$.

This means that any homeomorphism φ from $\mathcal{S}$ to itself and taking punctures to punctures represents an element of the mapping class group, provided it acts as the identity on the boundary of $\mathcal{S}$. Note that the punctures may be permuted by φ. Two homeomorphisms φ, ψ represent the same element if and only if they are isotopic through a family of boundary-fixing homeomorphisms which also fix $\mathcal{P}$. They will then induce the same permutation of the punctures. The group structure on $\mathcal{MCG}(\mathcal{S},\mathcal{P})$ is given by composition; thus $\varphi\psi$ simply means "first apply ψ, then φ".

We remark that in the previous paragraph the word "isotopic" could have been replaced by the word "homotopic": by a theorem of Epstein [**56**], two homeomorphisms of a compact surface are homotopic if and only if they are isotopic.

Mapping class groups, also known as modular groups, play a prominent role in the study of the topology and geometry of surfaces, as well as in 3-dimensional topology. To illustrate the difficulty of understanding them, we note that simply proving that they admit finite (and in fact quite elegant) presentations already requires deep arguments [**72, 141**].

Our aim now is to sketch a proof (for details see for instance [**8**]) of the following result:

Proposition 1.1.3. — *There is an isomorphism $B_n \cong \mathcal{MCG}(D_n)$.*

Proof. — We shall sketch a proof that $\mathcal{MCG}(D_n)$ is naturally isomorphic to the group of isotopy classes of geometric braids defined above. Let β be a geometric n-braid, sitting in the cylinder $[0,1]\times D^2$, whose n strands are starting at the puncture points of $\{0\}\times D_n$ and ending at the puncture points of $\{1\}\times D_n$. The braid may be considered as the graph of the motion, as time goes from 1 to 0, of n points moving in the disk, starting and ending at the puncture points (letting time go from 0 to 1 would lead to an anti-isomorphism). It can be proved that this motion extends to a continuous family of homeomorphisms of the disk, starting with the identity and fixed on the boundary at all times. The end map of this isotopy is the corresponding homeomorphism $\varphi\colon D_n \to D_n$, which is well-defined up to isotopy fixed on the punctures and the boundary.

Conversely, given a homeomorphism $\varphi\colon D_n \to D_n$, representing some element of the mapping class group, we want to get a geometric n-braid. By a well-known trick of Alexander, any homeomorphism of a disk which fixes the boundary is isotopic to the identity, through homeomorphisms fixing the boundary. The corresponding braid is then just the graph of the restriction of such an isotopy to the puncture points. □

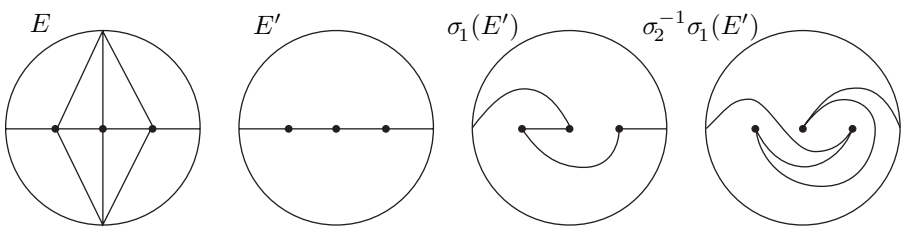


FIGURE 1.2. Two possible diagrams E, and the image of one of them under the homeomorphisms σ_1 and $\sigma_2^{-1}\sigma_1$

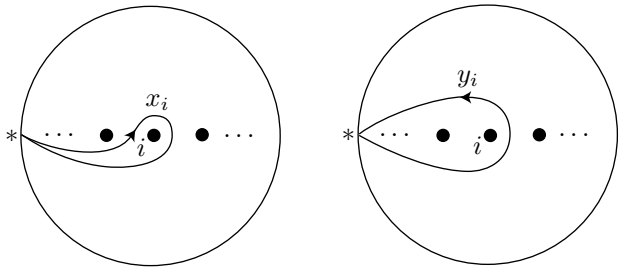


FIGURE 1.3. Generators for the fundamental group of D_n

In order to visualize a mapping of a surface, it is useful to consider images of certain subsets of the surface. Let E be a diagram on the surface $\mathcal{S}$, consisting of a finite number of disjoint, properly embedded arcs (meaning the arcs terminate either on $\partial\mathcal{S}$ or in a puncture of $\mathcal{S}$). Suppose in addition that E fills $\mathcal{S}$, in the sense that the interiors of all components of the surface obtained by cutting $\mathcal{S}$ along the arcs of E are homeomorphic to open disks. Typical examples in the case $\mathcal{S} = D_n$ are the standard triangulation, as well as the collection of $n+1$ horizontal line segments indicated in Figure 1.2. Then the isotopy class of a homeomorphism $\varphi\colon \mathcal{S} \to \mathcal{S}$ is uniquely determined by the isotopy class of the diagram $\varphi(E)$. This fact is also illustrated in Figure 1.2.

It is also well-known that a homeomorphism of D_n can be recovered up to homotopy from the induced isomorphism of the fundamental group $\pi_1(D_n, *)$, where $*$ is a fixed point of the boundary ∂D_n. The group $\pi_1(D_n, *)$ is a free group on n generators, say F_n. So we obtain an embedding $B_n \cong \mathcal{MCG}(D_n) \to \mathrm{Aut}(F_n)$, which can be written explicitly if we choose a base point $*$ and generators of $\pi_1(D_n, *)$. Two choices will play an important role in this text, namely when the base point $*$ is the leftmost point of the disk, and the generators are the loops $x_1, \ldots, x_n$ in the first case, and $y_1, \ldots, y_n$ in the second case, as shown in Figure 1.3.

1.1.3. Braid monoids. — A useful perspective is to restrict attention to a special class of braids, namely *positive* braids: we can think of our presentation of B_n not as a group presentation but as a monoid presentation, so that elements of the monoid are words in the letters σ_i, but not σ_i^{-1}, subject to our two types of relations, which involve only positive words. The resulting monoid is called the *positive braid monoid* and denoted B_n^+. Geometrically, we might think of this as the monoid of geometric braids with only positive crossings in their diagram, up to isotopy, where the isotopies are deformations through a family of braids that are again positive in the same sense.

The reason why this point of view is so fruitful is the following lemma due to Garside [**66**]:

Lemma 1.1.4. — *The canonical mapping of B_n^+ to B_n is injective.*

In other words, the braid monoid embeds in the braid group; thus we could have defined B_n^+ as the subset of B_n consisting of braids which are representable by words in the letters σ_i (not using any σ_i^{-1}).

The positive braid monoid is often easier to handle than the full braid group. On the other hand, the study of B_n^+ can be fruitful for the understanding of B_n, because of the following result:

Proposition 1.1.5. — *For every braid β in B_n there exist positive braids β_1, β_2 in B_n^+ satisfying $\beta = \beta_1^{-1}\beta_2$.*

Proof. — In order to give a simple-minded proof of the proposition, we start by identifying the centre of the group B_n. We define the *Garside fundamental braid* Δ_n to be the positive half-twist braid

$$\Delta_n = (\sigma_1 \dots \sigma_{n-1})(\sigma_1 \dots \sigma_{n-2}) \dots (\sigma_1\sigma_2)(\sigma_1). \tag{1.1.1}$$

It is well-known [**25**] that the centre of B_n is just the infinite cyclic subgroup generated by the full twist Δ_n^2.

Next we observe that Δ_n^2 has $n(n-1)$ crossings, and for every generator σ_i the braid $\sigma_i^{-1}\Delta_n^2$ admits a positive representative with $n(n-1)-1$ crossings.

Thus, given any representative w of a braid β, we can find an equivalent fraction $w_1^{-1}w_2$ in the following way: we replace each letter σ_i^{-1} (with negative exponent) in w by a positive word with $n(n-1)-1$ letters representing $\sigma_i^{-1}\Delta_n^2$ (this yields the word w_2), and define w_1 to be any positive braid word representing Δ_n^{2k}, where k is the number of letters with negative exponent in w. □

The decomposition of the braid β as a fraction of two positive braids provided by Proposition 1.1.5 is obviously not unique. Moreover, the decomposition constructed in the proof of Proposition 1.1.5 is algorithmically wasteful and unhelpful for solving the word problem in braid groups. Garside's deep insight (which was later generalized to many other groups [**12, 42**]) was that there exists a canonical such decomposition, in which the denominator and numerator have no nontrivial common divisor in the monoid B_∞^+. We shall return to this topic in Chapter 2.

1.1.4. Artin groups and generalizations. — Starting from the presentation of B_n, rather than from any geometric description, we can situate braid groups in a larger framework of a completely different nature: they are special cases of Artin groups, and more specifically spherical Artin groups, as introduced in [**47, 12**].

An Artin group—or, better, an *Artin–Tits* group—is, by definition, a group which admits a presentation with a finite number of generators $s_1, \dots, s_n$, such that all relations are of the form $s_is_js_is_j \dots = s_js_is_js_i \dots$, where the words on both sides of the equality sign have the same length (finite and at least 2), and i, j lie in $\{1, \dots, n\}$.

For instance, finitely generated free groups (no relations) and free abelian groups (commutation relations between all pairs of generators) are Artin–Tits groups. An Artin–Tits group is said to be *spherical* if the associated Coxeter group, *i.e.*, the group obtained by adding the relations $s_i^2 = 1$ for $i = 1, \dots n$, is finite [**77**]. Spherical Artin–Tits groups are much better understood than general ones. The archetypical example of a spherical Artin–Tits group with n generators is the braid group B_{n+1}, where the quotient obtained by turning all generators into involutions is the symmetric group $\mathfrak{S}_{n+1}$.

The algebraic, algorithmic and geometric properties of Artin–Tits groups are still a central focus of current research. See for instance [**47, 12, 137, 23, 108**] for a proof that spherical Artin–Tits groups have some very good algorithmic properties (efficient solutions to the word problem), [**48**] for a proof that they are linear, *i.e.*, isomorphic to groups of finite-dimensional matrices, [**116**] for a proof that the natural map of an Artin–Tits monoid into the corresponding group is an embedding, [**28**] for a survey on the geometry of metric spaces on which Artin–Tits groups can act, and [**136, 24, 45**] for the determination of their homology.

It is an open problem to determine which Artin–Tits groups are (left)-orderable. Indeed, we only know that three families, those of so-called type A_n (the braid groups) and B_n, and those of type D_n (which embed in mapping class groups of surfaces with boundary) are left-orderable, but no further orderability or non-orderability results are currently known.

Let us finally mention that many algebraic properties of Artin–Tits groups extend to a larger class of groups called *Garside groups* [**42, 44, 46, 123, 122, 134**].

1.2. A linear ordering on B_n

We turn now to the linear (strict total) ordering of the braid group that is the main subject of this text, namely the one sometimes called the Dehornoy ordering of braids. In this section we shall give a first definition of the ordering. Most of the structures on the braid groups afforded by this ordering can be condensed into three key properties called **A**, **C**, and **S**, which we shall state here without proof. Our task in subsequent chapters will be to study the ordering from various perspectives, always with a view towards proving these three properties.

1.2.1. The basic result. — According to our definition of B_n by a presentation, a braid is an equivalence class of braid words; if the braid β is the equivalence class of the braid word w, we say that w is a *representative* braid word (or an *expression*) of β. As, for instance, the braid word $\sigma_1^k \sigma_1^{-k}$ is a representative of the unit braid for every k, each braid admits infinitely many representative braid words.

Definition 1.2.1. — Assume $2 \leqslant n \leqslant \infty$. For β_1, β_2 in B_n, we say that $\beta_1 < \beta_2$ is true if, for some i, the braid $\beta_1^{-1}\beta_2$ admits at least one representative braid word where σ_i appears, but neither σ_i^{-1} nor any $\sigma_j^{\pm 1}$ with $j < i$ does.

The central property is the following result of [**34**] (first part of Theorem I.1):

Theorem 1.2.2. — *The relation $<$ is a linear ordering on B_n which is compatible with multiplication on the left.*

This statement can be discussed more easily if we introduce the notion of a σ-positive braid word. Here, and in the sequel, we denote by sh the *shift* endomorphism of B_∞ that maps σ_i to σ_{i+1} for every i; we also use sh for braid words in the obvious way.

Definition 1.2.3. — For w a braid word, we say that w is *σ_1-positive* (*resp. σ_1-negative*, *resp. σ_1-free*) if w contains at least one letter σ_1 and no letter σ_1^{-1} (*resp.* at least one σ_1^{-1} and no σ_1, *resp.* no σ_1 and no σ_1^{-1}); we say that w is *σ-positive* (*resp. σ-negative*) if it is the image of a σ_1-positive (*resp.* σ_1-negative) word under sh^{i-1} for some i with $i \geqslant 1$, *i.e.*, w contains at least one letter σ_i, but no letter σ_i^{-1} and no letter $\sigma_j^{\pm 1}$ with $j < i$.

Thus Definition 1.2.1 asserts that $\beta_1 < \beta_2$ is true if and only if the braid $\beta_1^{-1}\beta_2$ admits at least one σ-positive representative, making the name "σ-ordering" natural.

Example 1.2.4. — Let β be the braid $\sigma_1\sigma_2\sigma_1^{-1}$. The braid word $\sigma_1\sigma_2\sigma_1^{-1}$ is neither σ-positive, nor σ-negative. But this word happens to be equivalent to $\sigma_2^{-1}\sigma_1\sigma_2$, *i.e.*, $\sigma_2^{-1}\sigma_1\sigma_2$ is another representative braid word of β. As the latter word is σ_1-positive, hence σ-positive, we conclude that $\beta > 1$ is true: among all representatives of β (*i.e.*, of $1^{-1}\beta$), at least one is σ-positive.

What are the essential properties of the relation $<$ that are needed in order to prove Theorem 1.2.2? We observe that, for any group G, a left-invariant linear ordering exists on G if and only if there exists a subset Π of G satisfting $\Pi \cdot \Pi \subseteq \Pi$ and such that $\{1\}$, Π, and Π^{-1} is a partition of G. Given such Π, define $f < g$ to mean $f^{-1}g \in \Pi$ to obtain a left-invariant ordinering. Conversely, given a left-invariant order, take Π to be its positive cone, *i.e.*, $\Pi = \{g \in G; 1 < g\}$. Note that the formula $gf^{-1} \in \Pi$ would define a right-invariant ordering. In our current context, proving Theorem 1.2.2 amounts to proving that the set $B_n^{\sigma\text{-pos}}$ of all n-strand braids that admit at least one σ-positive representative qualifies as such a Π; the latter result decomposes into two statements:

Property A (Acyclicity). — *A braid that admits at least one σ_1-positive representative braid word is not trivial, i.e., a σ_1-positive braid word never represents the unit braid.*

Property C (Comparison). — *Every braid in B_n admits an n-strand representative braid word that is σ_1-positive, σ_1-negative, or σ_1-free, i.e., every n-strand braid word is equivalent to some n-strand braid word that is σ_1-positive, σ_1-negative, or σ_1-free.*

Let us immediately introduce another form of the latter property:

Property C (Comparison, second form**).** — *Every braid in B_n admits an n-strand representative braid word that is σ-positive, σ-negative, or empty, i.e., every n-strand braid word is equivalent to some n-strand braid word that is σ-positive, σ-negative, or empty.*

By definition, a σ-positive braid word is either σ_1-positive or σ_1-free, so it is clear that the second form of Property **C** implies the first one. The converse implication follows from a straightforward induction on n. For $n = 2$, the two statements coincide. Otherwise, we start with an n-strand braid word w; by the first form of Property **C**, we find w' equivalent to w that is σ_1-positive, σ_1-negative, or σ_1-free. In the first two cases, we are done; in the last one, the word w' is $\mathrm{sh}(w_1)$ for some $(n-1)$-strand braid word w_1, and applying the induction hypothesis to w_1 gives an equivalent word w_1' that is σ-positive, σ-negative, or empty: now w is equivalent to $\mathrm{sh}(w_1')$, which is also σ-positive, σ-negative, or empty.

Proof of Theorem 1.2.2 from Properties **A** *and* **C**. — Our aim is to prove that $\{1\}$, $B_n^{\sigma\text{-pos}}$, $(B_n^{\sigma\text{-pos}})^{-1}$ partition B_n. The fact that $\{1\}$, $B_n^{\sigma\text{-pos}}$, and $(B_n^{\sigma\text{-pos}})^{-1}$ cover B_n is exactly the second form of Property **C**. Now, assume $\beta \in B_n^{\sigma\text{-pos}} \cap (B_n^{\sigma\text{-pos}})^{-1}$. Then both β and β^{-1} admit σ-positive representatives, and so does 1, which is $\beta\beta^{-1}$. As the unit braid 1 is invariant under sh, this would imply that 1 admits a σ_1-positive representative, contradicting Property **A**. □

Property **A** has four different proofs in this text: they can be found in Sections 2.3, 5.1, 6.2, 8.5. As for Property **C**, no less than five proofs are given, in Sections 2.2, 3.2, 4.3, 5.2, 6.2.

For the moment, we shall mention some variations of these properties. First we observe that a slighty weaker form of Property **C** is sufficient.

Property $\mathbf{C}^+$. — *Assume that β_1, β_2 are positive n-strand braids, i.e., belong to B_n^+. Then the braid $\beta_1^{-1}\beta_2$ admits a representative braid word that is σ-positive, σ-negative, or empty.*

Lemma 1.2.5. — *Property $\mathbf{C}^+$ is equivalent to Property* **C**.

Proof. — Clearly, Property $\mathbf{C}^+$ is a special case of Property **C**. On the other hand, by Proposition 1.1.5, every braid in B_n can be expressed as $\beta_1^{-1}\beta_2$ with β_1, β_2 in B_n^+. □

Remark 1.2.6. — For $n \leqslant \infty$, let us call Property $\mathbf{C}_n$ the restriction of Property $\mathbf{C}$ to B_n, *i.e.*, the statement that every braid in B_n can be represented by an n-strand braid word that is σ-positive, a σ-negative, or empty. As B_∞ is the union of all groups B_n, Property $\mathbf{C}_\infty$ is a consequence of the conjunction of Properties $\mathbf{C}_n$ for $n < \infty$. But, conversely, it is not clear that Property $\mathbf{C}_\infty$ implies any of Properties $\mathbf{C}_n$: indeed, it only asserts that every braid in B_n admits a representative that is σ-positive, σ-negative, or empty, but the latter may be an n'-strand braid word, with n' possibly larger than n.

Finally, we state a strengthening of Property $\mathbf{A}$:

Property $\mathbf{A}_i$. — *A braid that admits a representative braid word containing at least one letter σ_i but no letter σ_i^{-1} is not trivial, i.e., a braid word where σ_i occurs but σ_i^{-1} never represents the unit braid.*

It is not obvious that Property $\mathbf{A}$, which is exactly Property $\mathbf{A}_1$, implies Property $\mathbf{A}_2$, or, more generally, $\mathbf{A}_i$ for any i with $i \geqslant 2$: indeed, using the shift endomorphism and Property $\mathbf{A}$, we deduce that a braid which admits a representative braid word with no σ_1 or σ_1^{-1}, at least one σ_2, and no σ_2^{-1} is not trivial. But Property $\mathbf{A}_2$ claims more, as it involves all braid words with at least one σ_2 and no σ_2^{-1}, regardless of whether they also contain σ_1 and σ_1^{-1}.

It is in fact true that Property $\mathbf{A}$ implies Property $\mathbf{A}_i$ for every i, but the most natural proof of this result involves the framework of Chapter 2, and we postpone it.

Property $\mathbf{A}_i$ will be used in Chapter 3 to prove the convergence of handle reduction. It can also be used to define new orderings on B_n: for instance, using Property $\mathbf{A}_2$, we can construct a partial ordering on B_4 satisfying $\sigma_2 > \sigma_1 > \sigma_3$. However, this ordering is not a linear ordering, nor is any analogous ordering except the ones giving $\sigma_1 > \sigma_2 > \dots$ and $\sigma_1 < \sigma_2 < \dots$. The reason is that the counterpart of Property $\mathbf{C}$ is then false: for instance, every braid word representative of the braid $\sigma_2\sigma_1\sigma_3^{-1}\sigma_2^{-1}$ contains both σ_2 and σ_2^{-1}.

1.2.2. Compatibility properties of the braid ordering. — By construction, the linear order $<$ on B_∞ is compatible with multiplication on the left, and with the shift endomorphism: indeed, by definition, any shifted image of a σ-positive braid word is a σ-positive braid word. These properties nearly characterize the σ-ordering:

Proposition 1.2.7. — *Assume that $\prec$ is a partial order on B_∞ that is compatible with multiplication on the left and with shift, and that $1 \prec \mathrm{sh}(\beta)\sigma_1\mathrm{sh}(\beta')$ holds for all β, β'. Then $\prec$ coincides with $<$.*

Proof. — The hypotheses imply

$$1 \prec \mathrm{sh}(\beta_0)\,\sigma_1\,\mathrm{sh}(\beta_1) \prec \mathrm{sh}(\beta_0)\,\sigma_1\,\mathrm{sh}(\beta_1)\,\sigma_1\,\mathrm{sh}(\beta_2) \prec \cdots$$

for all $\beta_0, \beta_1, \dots$ Hence $1 \prec \beta$ holds for every braid β that admits a σ_1-positive representative. The compatibility of $\prec$ with sh implies $1 \prec \beta$ for every braid β that admits a σ-positive representative. Hence $\beta < \beta'$ implies $\beta \prec \beta'$. Finally $<$ being a linear ordering and $\prec$ being an ordering gives the equivalence. □

A word of warning: the order $<$ enjoys no other general compatibility properties than invariance under multiplication on the left. For instance, let $\beta_1 = \sigma_1\sigma_2^{-1}$, and $\beta_2 = \sigma_1\sigma_2\sigma_1$. The word $\sigma_1\sigma_2^{-1}$ contains one occurrence of σ_1 and no occurrence of σ_1^{-1}, so the braid β_1 is σ-positive, and $\beta_1 > 1$ is true. On the other hand, the braid $\beta_2^{-1}\beta_1\beta_2$ is represented by the word $\sigma_1^{-1}\sigma_2^{-1}\sigma_1^{-1}\sigma_1\sigma_2^{-1}\sigma_1\sigma_2\sigma_1$, hence also by the equivalent word $\sigma_2\sigma_1^{-1}$, which contains one letter σ_1^{-1} and no letter σ_1. So, by definition, $\beta_2 < \beta_1\beta_2$ is true. This shows that the order $<$ is *not* compatible with multiplication on the right: $\beta_1 > 1$ does not imply $\beta_1\beta_2 > \beta_2$. As $<$ is compatible with multiplication on the left, it follows that it cannot be compatible with the inverse operation: $\beta_1 < \beta_2$ does not imply $\beta_1^{-1} > \beta_2^{-1}$ in general.

The order $<$ on B_n is not Archimedian for $n \geqslant 3$: for every positive integer p, we have $1 < \sigma_2^p < \sigma_1$, so σ_2 is infinitely small with respect to σ_1. However, we mentioned in the proof of Proposition 1.1.5 that, for Δ_n defined as in (1.1.1), Δ_n^2 is central in B_n. As noted by A. Malyutin, this implies that Δ_n^2 is Archimedian in the following sense:

Proposition 1.2.8. — *For every braid β in B_n, there exists a unique integer r satisfying $\Delta_n^{2r} \leqslant \beta < \Delta_n^{2r+2}$.*

Proof. — Let w be an arbitrary representative word of β. Let us write w as

$$w_0\sigma_1^{-1}w_1\sigma_1^{-1}\dots\sigma_1^{-1}w_p,$$

where, for each i, the word w_i is σ_1-positive or σ_1-free. Fix a braid word $\sigma_1 u$ representing Δ_n^2 and beginning with σ_1, which exists by (1.1.1). As Δ_n^2 is central, an expression of $\Delta_n^{2p+2}\beta$ is $\sigma_1 u w_0 u w_1 u \dots u w_p$, a σ_1-positive word, so $\Delta_n^{-2p-2} < \beta$ is true.

Similarly, if β admits a representative containing q letters σ_1, then $\beta < \Delta_n^{2q+2}$ is true. We conclude that the intervals $[\Delta_n^{2r}, \Delta_n^{2r+2})$ cover B_n. As these intervals are disjoint by construction, the result follows. □

We obtain in this way a decomposition of $(B_n, <)$ into a sequence of disjoint intervals of size Δ_n^2, as suggested in Figure 1.4 below.

As the powers of Δ_n^{2r} are central in B_n, we obtain a weak compatibility result for the product with the σ-ordering:

Lemma 1.2.9. — *Let $\prec$ be $<$, $\leqslant$, $>$, or $\geqslant$—or, more generally, any left-invariant ordering on B_n. Then $\Delta_n^{2q} \prec \beta$ implies $\beta^{-1} \prec \Delta_n^{-2q}$, and the conjunction of $\Delta_n^{2p} \prec \alpha$ and $\Delta_n^{2q} \prec \beta$ implies $\Delta_n^{2p+2q} \prec \alpha\beta$.*

Proof. — Assume $\Delta_n^{2q} \prec \beta$. By multiplying by β^{-1} on the left, we obtain $\beta^{-1}\Delta_n^{2q} \prec 1$, which is also $\Delta_n^{2q}\beta^{-1} \prec 1$. Multiplying by Δ_n^{-2q} on the left, we deduce $\beta^{-1} \prec \Delta_n^{-2q}$.

Assume $\Delta_n^{2p} \prec \alpha$ in addition. By multiplying by Δ_n^{2q} on the left, we obtain $\Delta_n^{2p+2q} \prec \Delta_n^{2q}\alpha = \alpha\Delta_n^{2q}$, while $\Delta_n^{2q} \prec \beta$ implies $\alpha\Delta_n^{2q} \prec \alpha\beta$ by multiplying by α on the left. We deduce $\Delta_n^{2p+2q} \prec \alpha\beta$. □

We deduce that the action of conjugacy cannot move a braid too far:

Proposition 1.2.10. — [**104**] *Assume that β and β' are conjugate in B_n. Then $\Delta_n^{2q} \leqslant \beta < \Delta_n^{2q+2}$ implies $\Delta_n^{2q-2} \leqslant \beta' < \Delta_n^{2q+4}$. So, in particular, $\beta\Delta_n^{-4} < \beta' < \beta\Delta_n^4$ is always true.*

Proof. — Assume $\beta' = \alpha\beta\alpha^{-1}$ and $\Delta_n^{2q} \leqslant \beta < \Delta_n^{2q+2}$. By Proposition 1.2.8, we have $\Delta_n^{2p} \leqslant \alpha < \Delta_n^{2p+2}$ for some p. Lemma 1.2.9 first implies $\Delta_n^{-2p-2} < \alpha^{-1} \leqslant \Delta_n^{-2p}$, and, then,

$$\Delta_n^{2p+2q-2p-2} < \alpha\beta\alpha^{-1} < \Delta_n^{2p+2+2q+2-2p},$$

which gives $\Delta_n^{2q-2} < \beta' < \Delta_n^{2q+4}$. □

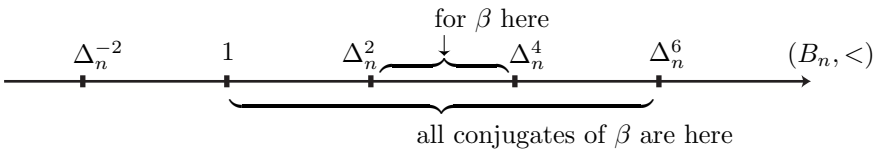


FIGURE 1.4. Powers of Δ_n^2 and action of conjugacy

The previous results are optimal in the sense that we cannot replace intervals of length Δ_n^2 with intervals of length Δ_n in Lemma 1.2.9: for instance, we have $1 < \sigma_1^2\sigma_2 < \Delta_3$ and $\Delta_3^2 < \Delta_3\sigma_1^2\sigma_2 < \Delta_3^3$.

1.2.3. Order and normal form. — As was mentioned after Proposition 1.1.5, each braid β admits a distinguished decomposition $\beta = d(\beta)^{-1}n(\beta)$ where $d(\beta)$ and $n(\beta)$ are positive braids admitting no nontrivial common left divisor in the braid monoid B_∞^+ [**66, 47, 57**] (see also Remark 2.1.13). The denominator $d(\beta)$ can easily be characterized in terms of (any) left-invariant ordering.

Proposition 1.2.11. — *For each braid β, the denominator $d(\beta)$ is the $<$-minimal positive braid α such that $\alpha\beta$ belongs to B_∞^+.*

Proof. — Assume $\alpha\beta \in B_\infty^+$, *i.e.*, $\alpha d(\beta)^{-1}n(\beta) \in B_\infty^+$. Let $\gamma = d(\alpha d(\beta)^{-1})$. We argue in the monoid B_∞^+. By construction, γ is a left divisor of $n(\alpha d(\beta)^{-1})d(\beta)$, and, on the other hand, it is also a left divisor of $\gamma\alpha\beta$, which is $n(\alpha d(\beta)^{-1})n(\beta)$. As $d(\beta)$ and $n(\beta)$ have no nontrivial common left divisor, this implies that γ is a left divisor of $n(\alpha d(\beta)^{-1})$. Now, by construction, γ, which is $d(\alpha d(\beta)^{-1})$, has no nontrivial common left divisor with $n(\alpha d(\beta)^{-1})$, so the only possibility is $\gamma = 1$, which means

that α is a left multiple of $d(\beta)$, *i.e.*, we have $\alpha = \alpha' d(\beta)$ for some α' in B_∞^+. By definition, this implies $\alpha \geqslant d(\beta)$. □

1.2.4. Order type and topology. — We wish to get a better overview over the global properties of the ordering $(B_\infty, <)$. In this section we shall firstly identify its order type, *i.e.*, characterize it up to isomorphism, and secondly describe the topology on B_∞ induced by $<$. In the next section we shall see that a much deeper understanding can be gained by studying the restriction of $<$ to the positive monoid B_∞^+.

Let us begin with the order type of $(B_\infty, <)$, which turns out to be nothing but that of the rationals.

Proposition 1.2.12. — *The ordered set $(B_\infty, <)$ is order-isomorphic to $(\mathbb{Q}, <)$.*

Proof. — A well-known result of Cantor says that any two countable linearly ordered sets that are dense (there always exists an element between any two elements) and unbounded (there is no minimal or maximal element) are isomorphic: assuming that the sets are $\{a_n; n \in \mathbb{N}\}$ and $\{b_n; n \in \mathbb{N}\}$, one alternatively defines $f(a_0)$, $f^{-1}(b_0)$, $f(a_1)$, $f^{-1}(b_1)$, *etc.* so as to keep f order-preserving.

Here the rationals are eligible, and the set B_∞ is countable. So, in order to apply Cantor's criterion, it suffices to prove that $(B_\infty, <)$ is dense and unbounded.

Now let β be an arbitrary braid. Then, by construction, we have $\beta\sigma_1^{-1} < \beta < \beta\sigma_1$, so $(B_\infty, <)$ is unbounded. Assume now $\beta \in B_n$ and $1 < \beta$. Then, by definition, we have $1 < \beta\sigma_n^{-1} < \beta$, which shows that there always exists a braid between 1 and β. Using a left translation by an arbitrary braid β', we deduce that there always exists a braid between β' and $\beta'\beta$, which is enough to conclude that $(B_\infty, <)$ is dense. □

Of course, the order-isomorphism could not be an isomorphism in the algebraic sense, as B_∞ is nonabelian. We also note that the ordering on B_n is actually discrete: σ_{n-1} is the least element of B_n greater than the identity. We conclude with a description of the topology associated with $<$.

Proposition 1.2.13. — *For $\beta_1 \neq \beta_2$ in B_∞, define $d(\beta_1, \beta_2)$ to be 2^{-k} where k is the greatest integer satisfying $\beta_1^{-1}\beta_2 \in \mathrm{sh}^k(B_\infty)$. Then d is a distance on B_∞, and the topology of B_∞ associated with the linear order $<$ is the topology associated with d.*

Proof. — First, it is clear that d is a distance. The open disk of radius 2^{-p} centered at β_0 is the left coset $\beta_0\,\mathrm{sh}^p(B_\infty)$, *i.e.*, the set of all braids $\beta_0\beta$ such that β admits an expression where no generator $\sigma_k^{\pm 1}$ with $k \leqslant p$ occurs.

Assume now that β_1, β, β_2 lie in B_n and β belongs to the open interval (β_1, β_2), *i.e.*, $\beta_1 < \beta < \beta_2$ holds. We wish to show that some open d-disk around β is contained in (β_1, β_2). By hypothesis, $\beta_1^{-1}\beta$ admits an expression of the form $\mathrm{sh}^k(w)$ with w a σ_1-positive word and $k \leqslant n$. Now the same holds for $\beta_1^{-1}\beta\beta'$ for every braid β' in $\mathrm{sh}^{n-1}(B_\infty)$, which means that every braid in the open d-disk of radius 2^{n-1} centered

at β is greater than β_1. A similar argument shows that all braids in this disk are also smaller than β_2.

Conversely, let us start with an arbitrary open d-disk $\beta_0 \operatorname{sh}^p(B_\infty)$. Let β be a braid in this disk; we have to find an open $<$-interval containing β which lies entirely in the disk. By hypothesis, we have $\beta = \beta_0 \operatorname{sh}^p(\beta')$ for some β' in B_∞. Let β_1 and β_2 be any two braids satisfying $\beta_1 < \beta' < \beta_2$. Then the interval $(\beta_0 \operatorname{sh}^p(\beta_1), \beta_0 \operatorname{sh}^p(\beta_2))$ contains $\beta_0 \operatorname{sh}^p(\beta')$ and lies in the disk. This completes the proof that the topologies associated with $<$ and with d coincide. □

1.2.5. The well-ordering property. — We have seen above that the braid ordering $<$ is *not* compatible with multiplication on the right, and, therefore, that a conjugate of a braid larger than 1 need not be larger than 1. This phenomenon cannot, however, occur with conjugates of positive braids, *i.e.*, of braids that can be expressed using the generators σ_i only, and not their inverses. The core of the question is the last of the three fundamental properties we shall mention here:

Property S (Subword). — *Every braid of the form $\beta^{-1}\sigma_i\beta$ admits a σ-positive representative braid word; in other words, we always have $\sigma_i\beta > \beta$.*

Proofs of Property **S** appear in Sections 2.4, 4.4, 6.2, and 7.3.

Using the compatibility of $<$ with multiplication on the left and a straighforward induction, we deduce the following result, which explains our terminology:

Proposition 1.2.14. — *Assume that β_1, β_2 are braids and some representative braid word of β_2 is obtained by inserting positive letters σ_i in a representative of β_1. Then we have $\beta_2 > \beta_1$.*

Another consequence is:

Proposition 1.2.15. — *If β belongs to B_∞^+ and is not 1, then $\beta' > 1$ is true for every conjugate β' of β. More generally, $\beta > 1$ is true for every quasi-positive braid β, the latter being defined as a braid that can be expressed as a product of conjugates of positive braids.*

Proof. — It suffices to establish that $\beta^{-1}\sigma_i\beta > 1$ holds for every braid β: by left multiplication, this inequality is equivalent to $\sigma_i\beta > \beta$, which is Property **S**. □

As was noted by Stepan Orevkov [**114**], the converse implication is not true: the braid $\sigma_2^{-5}\sigma_1\sigma_2^2\sigma_1$ is a non-quasi-positive braid but every conjugate of it is σ-positive.

Remark 1.2.16. — Let us call Property $\mathbf{S}^+$ the assertion that every braid of the form $\beta^{-1}\sigma_i\beta$ with β a positive braid admits a σ-positive representative braid word, *i.e.*, that $\sigma_i\beta > \beta$ is true whenever β is a positive braid. Property $\mathbf{S}^+$ is a particular case of Property **S**. Actually, both properties are equivalent. Indeed, assume that β is an arbitrary braid in B_n. Then, for k large enough, the braid $\Delta_n^{2k}\beta$ is positive (see

the proof of Proposition 1.1.5). If Property $\mathbf{S}^+$ is true, we obtain $\sigma_i \Delta_n^{2k} \beta > \Delta_n^{2k} \beta$ for $i < n$. We deduce $\Delta_n^{2k} \sigma_i \beta > \Delta_n^{2k} \beta$, as Δ_n^2 commutes with σ_i, and further $\sigma_i \beta > \beta$ by multiplying by Δ_n^{-2k} on the left.

The importance of Property **S** stems from the following consequence:

Proposition 1.2.17. — *For every n, the restriction of $<$ to B_n^+ is a well-ordering.*

Proof. — A theorem of Higman [**75**] says: An infinite set of words over a finite alphabet necessarily contains a pair $\{u, v\}$ such that u is a subword of v. Let $\beta_1, \beta_2, \dots$ be an infinite sequence of braids in B_n^+. For each i, choose a positive braid word w_i representing β_i. There are only finitely many n strand braid words of a given length, so, for each i, there exists j such that w_j is at least as long as w_i. So, inductively, we can extract a subsequence $w_{i_1}, w_{i_2}, \dots$ in which the lengths are non-decreasing. If the set $\{w_{i_1}, w_{i_2}, \dots\}$ is finite, we must have $\beta_i = \beta_j$ for some distinct i, j. Otherwise, by Higman's theorem, there exist p, q such that w_{i_p} is a subword of w_{i_q}, and, by construction, we must have $i_p < i_q$. By Property **S**, this implies $\beta_{i_p} \leqslant \beta_{i_q}$ in B_n^+. So, in any case, the sequence $\beta_1, \beta_2, \dots$ is not decreasing. □

More precise results about the well-ordering $(B_n^+, <)$ are known—but they are not known to be consequences of Property **S**: they follow from the specific approach developed in Chapter 4.

Proposition 1.2.18. — *For each n, the ordered set $(B_n^+, <)$ is order-isomorphic to $(\omega^{\omega^{n-2}}, <)$.*

This fundamental result means that there exists a one-to-one order-preserving correspondence between positive n-strand braids equipped with $<$ and the elements of the ordinal $\omega^{\omega^{n-2}}$, which, by construction, are the ordinals smaller than $\omega^{\omega^{n-2}}$—here ω denotes the first infinite ordinal, which coincides with the set $\mathbb{N}$ of natural numbers: we refer to any textbook in set theory, for instance [**95**], for background information about ordinals. In other words, every positive braid in B_n^+ receives a well-defined ordinal index, usually called its rank. In the case of B_2^+, this rank is an ordinal smaller than ω, *i.e.*, a mere natural number, and the bijection is given by $\sigma_1^k \mapsto k$. In the case of B_3^+, Figure 1.5 shows the ranks of some braids in B_3^+, *i.e.*, the numbering of B_3^+ by ordinal numbers between 0 and ω^ω.

1.2.6. A well-ordering on B_∞^+. — The property that the restriction of $<$ to B_n^+ is a well-ordering does not extend to the whole of B_∞^+, as, by construction, we have

$$\sigma_1 > \sigma_2 > \cdots,$$

an infinite strictly descending sequence. However, this sequence is the only obstruction, and, by using a convenient automorphism to reverse the order of the generators, we can easily obtain an ordering $<^\phi$ on the whole of B_∞ whose restriction to B_∞^+ *is*

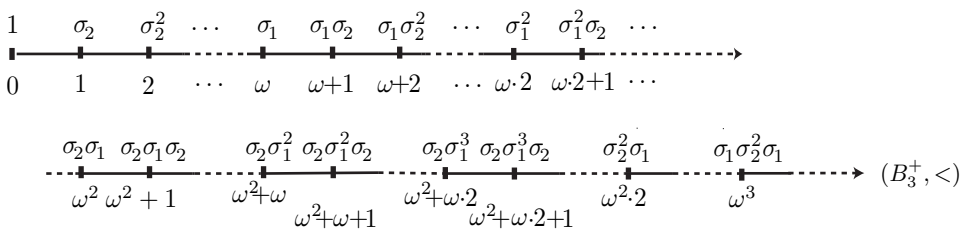


FIGURE 1.5. The well-ordering $(B_3^+, <)$

a well-ordering. Moreover, this well-ordering $<^\phi$, when further restricted to B_n^+, is conjugate to the well-ordering $<$.

Definition 1.2.19. — Assume that β_1, β_2 are braids. We say that $\beta_1 <^\phi \beta_2$ is true if $\phi_n(\beta_1) < \phi_n(\beta_2)$ holds for n large enough, where ϕ_n denotes the flip automorphism of B_n that exchanges σ_i and σ_{n-i} for every i.

In order to see that $<^\phi$ is well-defined, we have to check that if β_1, β_2 lie in B_n, and $n' \geqslant n$ is true, then $\phi_n(\beta_1) < \phi_n(\beta_2)$ is equivalent to $\phi_{n'}(\beta_1) < \phi_{n'}(\beta_2)$. Now, this equivalence is clear, as, for $\beta \in B_n$, we have by definition $\phi_{n'}(\beta) = \mathrm{sh}^{n'-n}(\phi_n(\beta))$ and $<$ is invariant under sh. Observe that we have

$$1 <^\phi \sigma_1 <^\phi \sigma_2 <^\phi \cdots,$$

i.e., the above obstruction to well-ordering is dropped.

Proposition 1.2.20. — (i) *The relation $<^\phi$ is a left invariant linear order on B_∞.* (ii) *The restriction of $<^\phi$ to B_∞^+ is a well ordering of type ω^{ω^ω}. For every n, the set B_n^+ is the initial segment of $(B_\infty^+, <^\phi)$ associated with σ_n, i.e., for β in B_∞^+, the relation $\beta \in B_n^+$ is equivalent to $\beta <^\phi \sigma_n$. For every positive braid β in B_n^+, the rank of β in $(B_\infty^+, <^\phi)$ is equal to the rank of $\phi_n(\beta)$ in $(B_n^+, <)$. In particular, the rank of σ_n in $(B_\infty^+, <^\phi)$ is 1 for $n = 1$, and is $\omega^{\omega^{n-2}}$ for $n \geqslant 2$.*

Proof. — Point (i) is clear. As for (ii), owing to the previous results, it suffices to prove that $\beta \in B_n^+$ is equivalent to $\beta <^\phi \sigma_n$. Assume $\beta \in B_n^+$. Then we have

$$\phi_{n+1}(\beta) = \mathrm{sh}(\phi_n(\beta)), \qquad \phi_{n+1}(\sigma_n) = \sigma_1,$$

and $\mathrm{sh}(\phi_n(\beta)) < \sigma_1$ implies $\beta <^\phi \sigma_n$.

Conversely, assume $\beta <^\phi \sigma_n$, with $\beta \in B_p^+$ and $p \geqslant n$. Then we have $\phi_p(\beta) < \sigma_{p-n}$. By Property **S**, we have $\beta_1\sigma_{p-n}\beta_2 \geqslant \sigma_{p-n}$ for all positive braids β_1, β_2, hence $\phi_p(\beta) < \sigma_{p-n}$ implies that $\phi_p(\beta)$ can be represented by a word involving $\sigma_{p-n+1}, \ldots, \sigma_{p-1}$ only, *i.e.*, we have $\phi_p(\beta) \in \mathrm{sh}^{p-n}(B_n)$, and $\beta \in B_n$. □

The interest of considering the order $<^\phi$ rather than its counterpart $<$ is that the rank of a positive braid β with respect to $<^\phi$ does not depend on a reference set B_n in which β lies. In Figure 1.6 the ordinal rank of a few braids in the well-ordering $(B_\infty^+, <^\phi)$ is indicated.

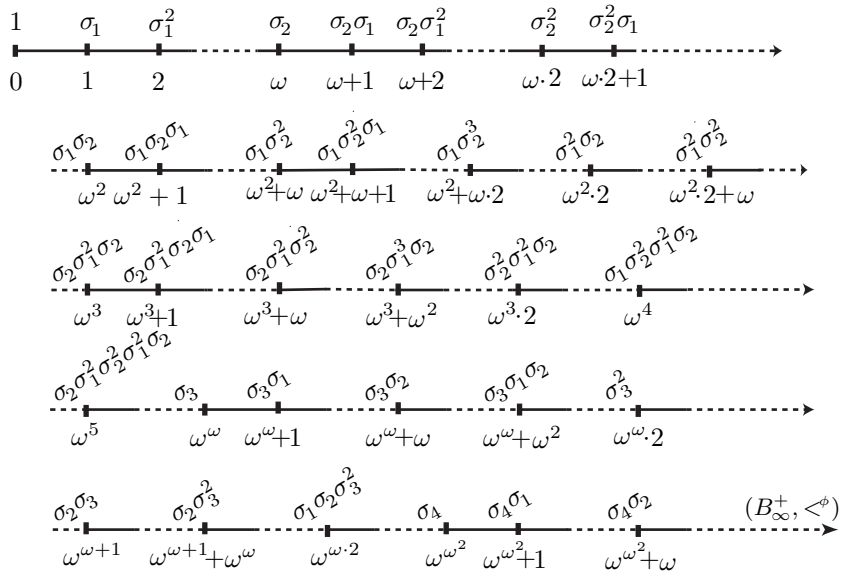


FIGURE 1.6. The well-ordering $(B_\infty^+, <^\phi)$

We can extend the ordinal indexation to arbitrary braids by using a fractionary decomposition. We mentioned in Proposition 1.1.5 that every braid β can be expressed as a fraction $\beta_1^{-1}\beta_2$ where β_1 and β_2 are positive braids. Furthermore, it can be proved [**47, 12, 38**] that the decomposition is unique if we require in addition that β_1 and β_2 have no nontrivial common left divisor in the braid monoid (see Remark 2.1.13). By the previous results, every positive braid is specified by a unique ordinal ξ with $0 \leqslant \xi < \omega^{\omega^\omega}$, and, therefore, every braid β can be specified by a pair of such ordinals, namely those associated with its numerator and its denominator. For instance, consider $\beta = \sigma_2\sigma_1^{-1}$. The distinguished decomposition of β is $(\sigma_2\sigma_1)^{-1}(\sigma_1\sigma_2)$. The rank of $\sigma_2\sigma_1$ in $(B_\infty^+, <^\phi)$ is $\omega + 1$, while the rank of $\sigma_1\sigma_2$ is ω^2. So, we associate with β the pair of ordinals $(\omega + 1, \omega^2)$. Observe that the rank of the numerator, namely ω^2, is bigger than that of the denominator, namely $\omega + 1$: this is natural as $\beta >^\phi 1$ is true.

1.3. Applications

As was mentioned in the introduction, two kinds of applications for the braid ordering can be considered: those that follow from the existence of the ordering, *i.e.*,

from the orderability of the braid groups, and those that follow from the specific characterization of the ordering in terms of σ-positivity.

1.3.1. Consequences of orderability. — The existence of the order $<$ implies that braid groups are left-orderable. We have already seen that $<$ is not compatible with multiplication on the right. Moreover, we have seen that, in B_3, the braid $\sigma_1\sigma_2^{-1}$ is conjugate to its inverse, which shows that no left-invariant ordering on B_3 could be at the same time right-invariant.

In a left orderable group, $1 < g$ implies $g^{-1} < 1$, and also $g < g^2, g^2 < g^3, \ldots$ and we conclude (with a similar argument for $g < 1$) that, if G is left orderable, then G has no elements of finite order. We deduce:

Proposition 1.3.1. — *The braid groups are torsion free.*

The following conjecture, dating from the first half of the 20th century, is still unsolved. Suppose R is a ring (commutative, with unit) and G a group. The group ring RG is the free module generated by the elements of G, endowed with a multiplication in an obvious way. If G has a torsion element, say g in G has order p, then in $\mathbb{Z}G$ there are necessarily zero-divisors. For example one calculates

$$(1-g)(1+g+g^2+\cdots+g^{p-1}) = 1-g^p = 0.$$

The Zero Divisor Conjecture claims that, if G is a torsion-free group, and R has no zero divisors, then the group ring RG also has no zero divisors [**117**]. Frustrating as attempts at this conjecture have been, even for the ring $\mathbb{Z}$, the question is easily settled for left orderable groups.

Lemma 1.3.2. — *The zero divisor conjecture is true if "left-orderable" replaces "torsion-free" in the hypothesis.*

Proof. — Consider a product in RG, say $(\sum_{i=1}^p r_i g_i)(\sum_{j=1}^q s_j h_j) = \sum_{i,j}(r_i s_j)(g_i h_j)$, with $h_1 < \ldots < h_q$. If $g_{i_0}h_{j_0}$ is a minimal term in the right hand side in the given ordering, use left-invariance to deduce $j_0 = 1$, and conclude that $g_{i_0}h_{j_0}$ is the unique minimal term, and, therefore, it cannot cancel with any other term. Similarly, the greatest term cannot cancel with any other term. So the product is nonzero unless all r_i's or all s_j's are zero, and it is equal to 1 if and only if we have $p = q = 1$, $r_1 s_1 = 1$, and $g_1 h_1$ is the identity element of G. □

The previous argument also shows that there are no exotic units in RG, the only invertible elements are the monomials rg, with r a unit of R and g in G.

In studying the braid groups B_n and their representations, the group rings $\mathbb{Z}B_n$ and $\mathbb{C}B_n$ are especially important. It was not until the proof that B_n is left-orderable that we knew the following fact:

Proposition 1.3.3. — *The rings $\mathbb{Z}B_n$ and $\mathbb{C}B_n$ have no zero divisors, and consequently no idempotents.*

For bi-orderable groups we have a stronger conclusion, due independently to Malcev [**101**] and Neumann [**111**], namely that, if G is bi-orderable, then $\mathbb{Z}G$ embeds in a skew field. We shall see in Chapter 9 that the pure braid group P_n is bi-orderable, so we deduce

Proposition 1.3.4. — *For every n, the group ring $\mathbb{Z}P_n$ embeds in a skew field.*

(The corresponding result for $\mathbb{Z}B_n$ has been proved by Linnell and Schicks recently [**97**].)

We have observed that, with left-invariant orderings, we can have $x < y$ and $x' < y'$ but $xx' > yy'$; in a bi-ordered group, one easily establishes that $x < y$ and $x' < y'$ together imply $xx' < yy'$. In particular, $x < y$ implies $x^k < y^k$ for all positive n. So the bi-orderability of P_n gives a new, short proof of the following:

Proposition 1.3.5. — *For every n, the group P_n has unique roots, i.e., if α and β are pure braids and α^k is equal to β^k for some positive k, then α and β are equal.*

The full braid groups B_n, with $n > 2$, certainly do not have unique roots. For instance $(\sigma_1\sigma_2)^3$ and $(\sigma_2\sigma_1)^3$ are equal in B_3 whereas $\sigma_1\sigma_2$ and $\sigma_2\sigma_1$ are distinct; they even determine distinct permutations.

It was recently shown that nonisomorphic groups may have isomorphic integral group rings [**73**]. Another fascinating property of orderable groups is that such a phenomenon is impossible once at least one of the groups is left-orderable [**87**]. Applying this result to the braid groups, we obtain:

Proposition 1.3.6. — *Assume that G is a group and the ring $\mathbb{Z}G$ is isomorphic to $\mathbb{Z}B_n$. Then the group G is isomorphic to B_n.*

We turn briefly to a bit of analysis. Let G be an infinite discrete group and let $L^2(G)$ denote the complex Hilbert space with Hilbert basis $\{g; g \in G\}$. The space $L^2(G)$ may be regarded as the set of formal sums $\sum_{g\in G} a_g g$ with $a_g \in \mathbb{C}$ and $\sum_{g\in G} |a_g|^2 < \infty$. The group ring $\mathbb{C}G$ may be considered as the subset of $L^2(G)$ for which all but finitely many of the a_g are zero. If α, β are two elements of $L^2(G)$, say $\alpha = \sum_{g\in G} a_g g$ and $\beta = \sum_{h\in G} b_h h$, the formal product defined by

$$\alpha\beta = \sum_{g,h\in G} a_g b_h gh$$

may not be in $L^2(G)$, but it will be so, provided α belongs to $\mathbb{C}G$. It is conjectured that, if G is torsion-free, and $\alpha \in \mathbb{C}G$ and $\beta \in L^2(G)$ are both nonzero, then $\alpha\beta$ is also nonzero. This is an extension of the zero divisor conjecture for group rings. Now, if G is left orderable, and α in $\mathbb{C}G$ and f in $L^2(G)$ are both nonzero, then [**96**] we can deduce $\alpha f \neq 0$. In the case of braid groups, we thus obtain:

Proposition 1.3.7. — *Assume $\alpha \in \mathbb{C}B_n$ and $f \in L^2(B_n)$ with α and f nonzero. Then αf is non-zero.*

1.3.2. Applications of σ-positivity. — Besides the previous consequences to the fact that the braid groups are orderable, other properties follow from the specific characterization of the ordering in terms of σ-positive braid words, *i.e.*, from Properties **A**, **C**, or **S**. For instance, Property **C**, *i.e.*, the fact that every nontrivial braid admits a σ-positive or a σ-negative representative, immediately provides the following criterion:

Proposition 1.3.8. — *Assume that f is a homomorphism of B_n into some group G and the image under f of a braid that admits a σ-positive representative braid word is never trivial. Then f is injective.*

As will be shown in Chapter 5, this criterion applies to the well-known homomorphism of B_n to $\mathrm{Aut}(F_n)$, where F_n is a rank n free group, which (re)-proves that this homomorphism is an embedding. The method is applied in [**132**] to show that other homomorphisms of B_n to $\mathrm{Aut}(F_n)$ defined in [**140**] also are embeddings. Whether the criterion applies to any of the classical or recently discovered linear representations of the braid groups is an open question (see [**36**] for a discussion in the case of the Burau representation).

Another consequence, observed by Edward Formanek, is

Proposition 1.3.9. — *For every n, the group B_n is isolated in B_∞, i.e., if β belongs to B_∞ and some power of β belongs to B_n, then β belongs to B_n.*

The result is straightforward using the reversed variant $<^\phi$ of $<$, since we have seen that B_n coincides with the open interval $(\sigma_n^{-1}, \sigma_n)$ in $(B_\infty, <^\phi)$, and $1 \leqslant^\phi \beta$ implies $\beta \leqslant^\phi \beta^k$ for every positive k, so $\sigma_n \leqslant^\phi \beta$ implies $\sigma_n \leqslant^\phi \beta^k$.

Property **C** implies that, if β^k has a σ_1-free expression, *i.e.*, one where neither σ_1 nor σ_1^{-1} occurs, then the same is true for β. We do not know whether a similar property with σ_i instead of σ_1 is true.

1.3.3. Application of well-orderability. — Besides the fact that the elements greater than 1 may have a special form, the fact that inserting one generator always increases the braid, which is Property **S**—a property that the σ-ordering shares with all orderings of Chapter 7—could lead to specific applications.

In a recent (unpublished) paper [**104**], A. Malyutin shows that the closure of every braid that is large enough in such an ordering is a prime non-split link and that such a braid is terminal with respect to the Birman–Menasco operations that are complexity-preserving extensions of Markov moves—see also [**102**].

We have also seen that Property **S** leads to associating with each braid β a canonical pair of ordinals, say $o(\beta)$. Now, each time we are given some equivalence relation $\sim$

on braids, we can define for every braid β a distinguished pair of ordinals

$$o_{\sim}(\beta) = \min\{o(\beta') \ ; \ \beta' \sim \beta\},$$

where the pairs of ordinals are ordered lexicographically. In particular, this approach associates with every conjugacy class of braids, or with every equivalence class of braids under Markov moves, hence with every knot, a unique well defined pair of ordinals. Of course, computing the value of $o(\beta)$ in the general case is not obvious, and computing $o_{\sim}(\beta)$ for various equivalence relations $\sim$ is likely to be still more difficult.

CHAPTER 2

SELF-DISTRIBUTIVITY

This chapter presents an algebraic technique that was developed in the beginning of the 1990's, and led to the first proof of the orderability of the braid groups [**32, 34**]. Subsequently, Richard Laver used a related method to prove that the same ordering is a well-ordering when restricted to positive braids [**94**]. The approach is complete in that it provides proofs of Properties **A**, **C**, and **S**. In the current survey, we shall prove **A**, **S**, and of the weaker form $\mathbf{C}_\infty$ of **C**—which is sufficient for ordering B_∞, and, therefore, B_n for every n. The details are given for the key arguments only, most of the remaining proofs being sketched, or even skipped.

It had been observed for many years that braids are connected with left self-distributive systems (LD-systems for short), an LD-system being defined as a set equipped with a binary operation satisfying the left self-distributivity law, often called (LD),

$$x * (y * z) = (x * y) * (x * z). \tag{2.0.1}$$

In particular, David Joyce [**80**] and Sergei Matveev [**106**] introduced for every knot K a particular LD-system Q_K, the fundamental quandle of K, that characterizes the isotopy type of K up to a mirror image.

Here we use a different approach: instead of associating some particular LD-system S_β with every braid β, we choose one fixed LD-system S, and use it for all braids uniformly by defining an action of B_n on n-tuples in S. It is not surprising that the action leads to an ordering on B_n when S happens to be an ordered LD-system—in a sense that will be made precise below. Most classically considered LD-systems are connected with conjugation in a group, which makes them strongly non-orderable. The point that made the construction described below possible is the discovery of new, orderable LD-systems at the end of the 1980's. It is worth mentioning that the first example of an orderable LD-system came from set theory, and relied on an unprovable large cardinal hypothesis. This example is not needed here, and it was never needed. The construction we shall describe below was precisely made in

order to eliminate any use of this example. However, we insist that the whole story might never have happened without the hint from set theory that such an orderable LD-system could exist.

2.1. The action of braids on LD-systems

Our aim is to define an action of the group B_n on the n-fold product S^n whenever $(S, *)$ is an LD-system. Actually, we shall obtain an action of the monoid B_n^+ only, and see that this action extends into an action of the group B_n only if our LD-system S satisfies some additional requirements that make it a *rack* in the sense of [**61**], or an *automorphic set* in the sence of [**11**]. But—this is the point—even if S is only a left cancellative LD-system and not a rack, we can define a *partial* action of B_n on S^n which enjoys enough properties for the subsequent needs.

Considering a partial action makes it crucial to maintain a clear distinction between braids and braid words in the sequel.

2.1.1. Braid colourings. — The most intuitive approach is to start with the idea of colouring the strands in a braid diagram. Assume that S is a fixed nonempty set, and that w is a braid word. Then we attribute colours from S to the left ends in the diagram encoded by w, we propagate them along the strands, and we compare the sequence of final colours with the sequence of initial colours.

If the colours are just pushed along the strands, the final sequence is a permutation of the initial one, and the only piece of information about the braid we obtain is its projection to the symmetric group.

Things become more interesting when we allow colours to change at crossings. Let us begin with positive braid diagrams. We consider the case when the colour of the back strand is preserved, but the colour of the front strand may change depending on the two colours which have crossed. This amounts to using a function of $S \times S$ into S, *i.e.*, a binary operation $*$ on S, with the rule

Thus, we define a right action of n-strand positive braid words on S^n by

$$\vec{x} \cdot \varepsilon = \vec{x}, \qquad \vec{x} \cdot \sigma_i w = (x_1, \dots, x_i * x_{i+1}, x_i, x_{i+2}, \dots, x_n) \cdot w, \tag{2.1.1}$$

where ε denotes the empty braid word (everywhere in the sequel, when $\vec{x}$ denotes a sequence, we use $x_1, x_2, \dots$ for the successive entries of $\vec{x}$).

Lemma 2.1.1. — *The action defined in* (2.1.1) *is compatible with the braid relations if and only if* $(S, *)$ *is an LD-system.*

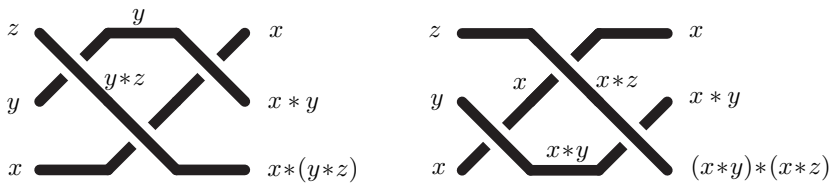


FIGURE 2.1. Compatibility of action with the braid relations

Proof. — Compatibility with the relations $\sigma_i\sigma_j = \sigma_j\sigma_i$ with $|i-j| \geqslant 2$ is obvious. As for the relations $\sigma_i\sigma_j\sigma_i = \sigma_j\sigma_i\sigma_j$ with $|i-j| = 1$, we see on the diagrams of Figure 2.1 that compatibility is guaranteed if and only if the equality $x*(y*z) = (x*y)*(x*z)$ holds in S, *i.e.*, if $(S,*)$ is what we have called an LD-system. □

Let us now consider arbitrary braid words. We have to colour negative crossings. In order to find the most flexible definition, let us first assume that S is equipped with two more binary operations, say $\circ$ and $\overline{*}$, and consider the rule

This amounts to extending the action of braid words on sequences of colours by

$$\vec{x} \cdot \sigma_i^{-1} w = (x_1, \dots, x_i \circ x_{i+1}, x_{i+1} \,\overline{*}\, x_i, x_{i+2}, \dots, x_n) \cdot w. \tag{2.1.2}$$

The reader will easily check that the action defined by (2.1.1) and (2.1.2) is compatible with the relations $\sigma_i\sigma_i^{-1} = \sigma_i^{-1}\sigma_i = 1$ if and only if the following identities are satisfied:

$$y \circ x = x, \qquad x * (x \,\overline{*}\, y) = x \,\overline{*}\, (x * y) = y. \tag{2.1.3}$$

Thus, the operation $\circ$ has to be trivial, while $\overline{*}$ must be chosen so that the left translations associated with $\overline{*}$ are inverses of the left translations associated with $*$. So $\overline{*}$ exists if and only if the left translations associated with $*$ are bijective, in which case we necessarily have

$$x \,\overline{*}\, y = \text{ the unique } z \text{ satisfying } x * z = y. \tag{2.1.4}$$

We can therefore state:

Proposition 2.1.2. — (i) *For every LD-system* $(S,*)$ *and every* n*, the rule* (2.1.1) *defines an action of* B_n^+ *on* S^n*.*

(ii) *Define a rack to be an LD-system* $(S,*)$ *in which all left translations are bijective, i.e., we have* $(\forall x,y)(\exists! z)(x*z=y)$*. Then, for every rack* $(S,*)$ *and every* n*, the rules* (2.1.1)*,* (2.1.2)*, and* (2.1.4) *(with* $x \circ y = y$*) define an action of* B_n *on* S^n*.*

For $\vec{x}$ a sequence of elements of S, and β a braid, we shall write $\vec{x} \cdot \beta$ for the result of applying β to $\vec{x}$, *i.e.*, for $\vec{x} \cdot w$ where w is an arbitrary braid word representing β.

2.1.2. Classical examples. — Considering classical examples of racks—also called LD-quasigroups in [**40**]—leads to not less classical results about braid groups.

***Example 2.1.3* (trivial).** — Let S be an arbitrary set. Defining $x * y = y$ turns S into a rack. For $\vec{x} \in S^n$, and $\beta \in B_n$, we find

$$\vec{x} \cdot \beta = \mathrm{perm}(\beta)^{-1}(\vec{x}),$$

where $\mathrm{perm}(\beta)$ denotes the permutation induced by β. Thus the action gives the surjective homomorphism of B_n onto the symmetric group $\mathfrak{S}_n$.

***Example 2.1.4* (shift).** — Defining $x * y = y + 1$ turns $\mathbb{Z}$ into a rack. For $\vec{x} \in \mathbb{Z}^n$ and $\beta \in B_n$, we find

$$\sum(\vec{x} \cdot \beta) = \sum \vec{x} + e(\beta), \tag{2.1.5}$$

where $\sum \vec{x}$ denotes $x_1 + \cdots + x_n$, and $e(\beta)$ denotes the exponent sum of β, *i.e.*, the difference between the number of positive and negative letters in any representative braid word of β.

***Example 2.1.5* (centre of mass).** — Let E be a $\mathbb{Z}[t, t^{-1}]$-module. Defining $x*y = (1-t)x + ty$ turns E into a rack. For every n-strand braid β, and every sequence $\vec{x}$ in E^n, the sequence $\vec{x} \cdot \beta$ is a linear combination of $\vec{x}$, *i.e.*, we have

$$\vec{x} \cdot \beta = \vec{x} \times \rho(\beta)$$

for some $n \times n$-matrix $\rho(\beta)$. The mapping ρ is a linear representation of B_n, namely the (unreduced) Burau representation.

***Example 2.1.6* (conjugation).** — Let G be a group. Defining $x * y = xyx^{-1}$ turns G into a rack. In particular, let F_n be the free group based on $\{x_1, \ldots, x_n\}$. For $\beta \in B_n$, define elements $y_1, \ldots, y_n$ of F_n by

$$(y_1, \ldots, y_n) = (x_1, \ldots, x_n) \cdot \beta,$$

and let $\varphi(\beta)$ be the endomorphism of F_n that maps x_i to y_i for every i. Then φ is a homomorphism of B_n into $\mathrm{End}(F_n)$, and, as $\varphi(1)$ is the identity by construction, the image of φ is actually included in $\mathrm{Aut}(F_n)$. The action gives the Artin representation

$$\varphi : B_n \to \mathrm{Aut}(F_n),$$

to which we shall return in Chapter 5.

***Example 2.1.7* (free rack).** — For G a group and $A \subseteq G$, defining $(x, a) * (y, b) = (xax^{-1}y, b)$ turns $G \times A$ into a rack. The operation can be called half-conjugacy as the mapping $f : G \times A \to G$ defined by $f((x, a)) = xax^{-1}$ is a surjective homomorphism of $(G \times A, *)$ onto G equipped with conjugation. It is easy to prove that, if F_A denotes

the free group based on A, then $(F_A \times A, *)$ is a free rack based on $\{(1, a); a \in A\}$. So, in some sense, this example is the most general possible one.

2.1.3. Acyclic and orderable LD-systems. — Once we know that braid groups act on the powers of racks, it is natural to think of using ordered racks for possibly deducing an ordering of the braids. Several definitions may be considered, but it turns out that the convenient one is as follows:

Definition 2.1.8. — We say that an LD-system $(S, *)$ is *acyclic* if the left divisibility relation in $(S, *)$ has no cycle, *i.e.*, if no equality of the form

$$x = (\dots((x * x_1) * x_2) * \dots) * x_p. \tag{2.1.6}$$

is true in S. We say that $(S, *, <)$ is an *ordered* LD-system if $(S, *)$ is an LD-system, and $<$ is a strict linear ordering on S such that $x < x * y$ is always true, and $y < z$ implies $x * y < x * z$.

By definition, the order in an ordered LD-system extends the left divisibility relation, and, therefore, the latter admits no cycle: thus an orderable LD-system is necessarily acyclic. Also, an ordered LD-system must be left cancellative, for $y \neq z$ implies $y < z$ or $z < y$, hence $x * y < x * z$ or $x * z < x * y$, and $x * y \neq x * z$ in both cases. Observe that the connection between the order and the operation in an ordered LD-system is the same as the connection between the standard ordering of positive integers and their addition.

An obstruction immediately appears:

Proposition 2.1.9. — *A rack is never acyclic, and, therefore, no rack may be an orderable LD-system.*

Proof. — Assume that $(S, *)$ is a rack. We claim that the identity $(x * x) * y = x * y$ holds in S. Indeed, using the symmetric operation $\overline{*}$ of (2.1.4), we find

$$(x * x) * y = (x * x) * (x * (x \,\overline{*}\, y)) = x * (x * (x \,\overline{*}\, y)) = x * y.$$

(Alternatively, we can check the identity in the free rack $F_A \times A$ using a direct verification.) Thus, in particular, we have $(x * x) * x = x * x$, an equality of the type (2.1.6). □

At this point, it is not clear that orderable LD-systems exist. We shall see below that they do, but then we have a problem for defining a braid group action as, by Proposition 2.1.9, an orderable LD-system cannot be a rack. So, our first task will be to extend the braid action so as to use LD-systems that are not necessarily racks.

2.1.4. Braid word reversing. — The solution relies on a combinatorial tool called braid word reversing. All relations in the standard presentation of B_∞ have the form

$$\dots\sigma_i = \dots\sigma_j,$$

where the dots represent some braid word depending on σ_i and σ_j. Let us denote by Σ the alphabet $\{\sigma_1, \sigma_2, \dots\}$. We use $\equiv$ for the equivalence of braid words, and ε for the empty word. Let us consider the function f of $\Sigma \times \Sigma$ to positive braid words defined by

$$f(\sigma_i, \sigma_j) = \begin{cases} \sigma_i & \text{for } |i-j| \geqslant 2, \\ \sigma_j\sigma_i & \text{for } |i-j| = 1, \\ \varepsilon & \text{for } i = j. \end{cases}$$

Then braid word equivalence is generated by the relations

$$f(\sigma_j, \sigma_i)\ \sigma_i \equiv f(\sigma_i, \sigma_j)\ \sigma_j \tag{2.1.7}$$

for $i \neq j$. Now (2.1.7) implies

$$\sigma_i\ \sigma_j^{-1} \equiv f(\sigma_j, \sigma_i)^{-1}\ f(\sigma_i, \sigma_j), \tag{2.1.8}$$

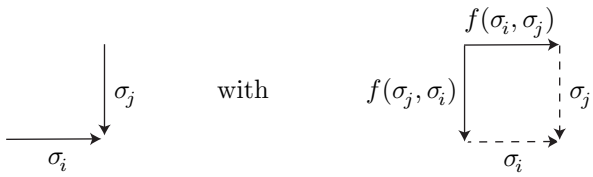
where we denote by w^{-1} the braid word obtained from w by reversing the order of the letters and exchanging σ_i with σ_i^{-1} everywhere. It follows that, if a braid word w' is obtained from another braid word w by repeatedly applying relations of the type (2.1.8), then w and w' are equivalent, *i.e.*, they represent the same braid.

Definition 2.1.10. — If w, w' are braid words, we say that w is *reversible* to w' on the left if w' can be obtained from w by (iteratively) replacing factors of the type $\sigma_i\sigma_j^{-1}$ with the corresponding factors $f(\sigma_j, \sigma_i)^{-1}\ f(\sigma_i, \sigma_j)$.

Example 2.1.11. — (word reversing) Let us consider $w = \sigma_1\sigma_2^{-1}\sigma_2\sigma_3^{-1}$. Then w contains the factor $\sigma_1\sigma_2^{-1}$, so it is reversible to $w_1 = \sigma_2^{-1}\sigma_1^{-1}\sigma_2\sigma_1\sigma_2\sigma_3^{-1}$. (Note that w also contains the factor $\sigma_2\sigma_3^{-1}$, so it is reversible to $\sigma_1\sigma_2^{-1}\sigma_3^{-1}\sigma_2^{-1}\sigma_3\sigma_2$ as well). Then w_1 contains the factor $\sigma_2\sigma_3^{-1}$, so it is reversible to $w_2 = \sigma_2^{-1}\sigma_1^{-1}\sigma_2\sigma_1\sigma_3^{-1}\sigma_2^{-1}\sigma_3\sigma_2$, *etc.* The reader can check that all sequences of word reversing from w end in seven steps with the word $\sigma_2^{-1}\sigma_1^{-1}\sigma_3^{-1}\sigma_2^{-1}\sigma_1^{-1}\sigma_3\sigma_2\sigma_1\sigma_3\sigma_2$. The latter word can no longer be reversed, for it contains no more factor of the form $\sigma_i\sigma_j^{-1}$.

Word reversing can be easily visualized using a planar diagram analogous to a van Kampen diagram. Assume that $(w_0, w_1, \dots)$ is a sequence of braid words such that, for each k, the word w_{k+1} is obtained from w_k by replacing exactly one factor $\sigma_i\sigma_j^{-1}$ with the corresponding factor $f(\sigma_j, \sigma_i)^{-1}\ f(\sigma_i, \sigma_j)$. We associate with w_0 an oriented path in $\mathbb{Q}^2$ by starting from $(0, 0)$, reading the successive letters of w_0, and attaching with every positive letter σ_i an horizontal right-oriented edge labelled σ_i, and with

every negative letter σ_i^{-1} a vertical, down-oriented edge also labelled σ_i. Then we iteratively attach similar paths with the words w_k: the hypothesis that w_{k+1} is obtained from w_k by replacing $\sigma_i\sigma_j^{-1}$ with $f(\sigma_j,\sigma_i)^{-1}\ f(\sigma_i,\sigma_j)$ corresponds to replacing



i.e., by closing the open patterns consisting of two converging edges into a square using the braid relations; in the special case of equal labels $\sigma_i\sigma_i^{-1}$, we add an ε-labeled unoriented arc (or, equivalently, two ε-labeled edges $\varepsilon\varepsilon^{-1}$) contributing nothing when read. Figure 2.2 shows the diagram associated in this way with the reversing sequence of Example 2.1.11.

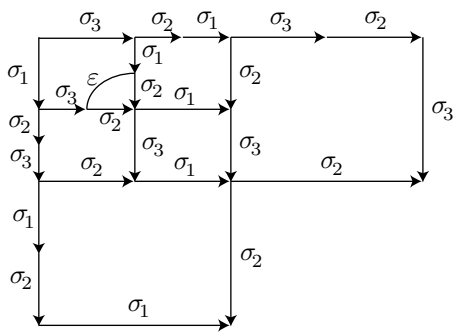


FIGURE 2.2. Example of reversing diagram

Several questions arise: does reversing always terminate, and, if so, does it lead to a unique final word? In the current case of braids, the answers are positive.

Proposition 2.1.12. — *Assume that w is a braid word. Then there exists a unique pair of positive words $N(w), D(w)$ such that every sequence of word reversing from w terminates in a finite number of steps with the word $D(w)^{-1}N(w)$.*

Proof (sketch). — The uniqueness of the reversing diagram implies that word reversing from a given word leads to at most one terminal word of the form $u^{-1}v$ with u, v positive braid words. So the problem is to prove termination. It is sufficient to do it when we start with a word of the form uv^{-1}, with positive u, v. For u, v positive braid words, let us denote by u/v the (unique) positive word u' such that uv^{-1} is reversible to $v'^{-1}u'$ for some positive word v', if such a word exists. Then $(v/u)u \equiv (u/v)v$ holds whenever the involved words exist, and, in this case, $(u/v)v$ represents a common left

multiple of the braids represented by u and v in the monoid B_∞^+. The point is that this common multiple is a *least* common multiple. It is shown in [**38**] that this general result about word reversing is valid provided the condition

$$((x/y)/(z/y))/((x/z)/(y/z)) = \varepsilon$$

is satisfied for every triple of letters (x, y, z), and the latter condition can be checked easily in the case of the braids. Then reversing the word uv^{-1} terminates in a finite number of steps if and only if the braids represented by u and v admit a common left multiple in B_∞^+, which was shown by Garside in [**66**] (we refer the reader to [**44**] for further developments of the word reversing technique). □

For instance, in the case of the word w of Example 2.1.11, the words $D(w)$ and $N(w)$ are $\sigma_1\sigma_2\sigma_3\sigma_1\sigma_2$ and $\sigma_3\sigma_2\sigma_1\sigma_3\sigma_2$, respectively.

It follows from (2.1.8) that, if w is reversible to w', then, in particular, w and w' are equivalent. Thus, Proposition 2.1.12 gives $w \equiv D(w)^{-1}N(w)$, a concrete realization of Garside's result that every braid is a fraction (Proposition 1.1.5). However, Proposition 2.1.12 gives us more than the equivalence, namely a distinguished way for transforming w into $D(w)^{-1}N(w)$, one that avoids introducing any new factor of the type $\sigma_i^{-1}\sigma_i$ or $\sigma_i\sigma_i^{-1}$. Notice that reversing is defined only at the level of braid words, and not of braids. Indeed, the functions N and D do not induce well-defined mappings of B_∞ into B_∞^+: for instance, the words $\sigma_1^{-1}\sigma_1$ and ε are equivalent, as both represent the unit braid, but we find

$$N(\sigma_1^{-1}\sigma_1) = D(\sigma_1^{-1}\sigma_1) = \sigma_1, \quad N(\varepsilon) = D(\varepsilon) = \varepsilon,$$

which shows that $w \equiv w'$ does *not* imply $N(w) \equiv N(w')$ or $D(w) \equiv D(w')$.

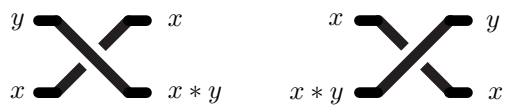
By construction, the braid relations are symmetric, and everything we said so far about reversing braid words on the left can be transposed into similar results about reversing braid words on the right, which means iteratively replacing factors of the form $\sigma_i^{-1}\sigma_j$ with corresponding factors of the form uv^{-1}. In this way, we obtain for every braid word w two positive braid words $\widetilde{N}(w)$ and $\widetilde{D}(w)$ such that w is right reversible to $\widetilde{N}(w)\widetilde{D}(w)^{-1}$.

Remark 2.1.13. — As the functions N and D, the functions $\widetilde{N}$ and $\widetilde{D}$ do not induce well-defined mappings on braids. However, let us mention here that, when a double, left and right, word reversing is used, then the resulting mappings induce well-defined mappings on braids: if, for w a braid word, we define $N_{\mathrm{rl}}(w) = \widetilde{N}(D(w)^{-1}N(w))$ and $D_{\mathrm{rl}}(w) = \widetilde{D}(D(w)^{-1}N(w))$, then $w \equiv w'$ implies $N_{\mathrm{rl}}(w) \equiv N_{\mathrm{rl}}(w')$ and $D_{\mathrm{rl}}(w) \equiv D_{\mathrm{rl}}(w')$. We obtain in this way a simple solution to the word problem of braids, *i.e.*, to the question of algorithmically recognizing whether a given braid word represents or not the unit braid: Indeed, $w \equiv \varepsilon$ is equivalent to $N_{\mathrm{rl}}(w) = D_{\mathrm{rl}}(w) = \varepsilon$, *i.e.*, a braid word w is trivial if and only if, when we reverse w to the left, and then the result to the right, we end up with an empty word [**38**]. It turns out that, for every braid

word w, the word $N_{\mathrm{rl}}(w)D_{\mathrm{rl}}(w)^{-1}$ is the shortest fraction equivalent to w. Using this and the normal form result of [**47, 1**] for positive braids leads to the greedy normal form of [**57**] and [**55**].

2.1.5. Action of braids on left cancellative LD-systems. — We are now ready to extend the action of B_n defined in Section 2.1 to left cancellative LD-systems that are not necessarily racks, *i.e.*, in which left division is not always possible.

Definition 2.1.14. — Assume that $(S, *)$ is an LD-system and w is an n-strand braid word, $n \leqslant \infty$. We say that a pair of sequences $(\vec{x}, \vec{y})$ in S^n is an *S-colouring* for w if colours from S can be attributed to each segment in the canonical diagram associated with w in such a way that $\vec{x}$ are the input (left) colours, $\vec{y}$ are the output (right) colours, and the rules

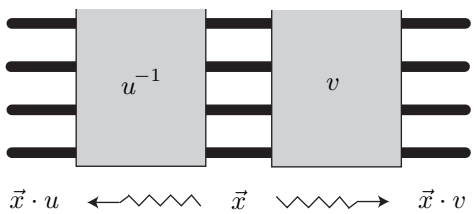


are obeyed at each crossing.

If w is a positive braid word, we know that, for every sequence $\vec{x}$ in S^n, the pair $(\vec{x}, \vec{x} \cdot w)$ is a colouring of w, and, if S is a rack, the same holds for every braid word. But, even if the LD-system we consider is not a rack, S-colourings exist for every braid word.

Lemma 2.1.15. — *Assume that $(S, *)$ is an LD-system and u, v are positive n-strand braid words. Then, for every sequence $\vec{x}$ in S^n, the pair $(\vec{x} \cdot u, \vec{x} \cdot v)$ is an S-colouring for $u^{-1}v$.*

Proof. — We apply the colours $\vec{x}$ in the middle of the diagram associated with $u^{-1}v$ and propagate them to the left through u^{-1} and to the right through v, as below:



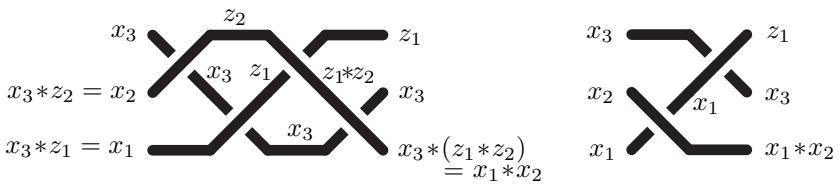
We obtain $\vec{x} \cdot u$ on the left of the diagram, and $\vec{x} \cdot v$ on the right. By construction, the rules of colouring are obeyed in each part of the diagram. □

We know that every braid can be represented by a braid word of the form $u^{-1}v$ with u, v positive, *i.e.*, every braid word w is equivalent to a word of the form above.

This, however, is *not* sufficient for concluding that there exists an S-colouring for w: indeed, an S-colouring for a braid word w need not be an S-colouring for every braid word w' that is equivalent to w. Now, we can use word reversing.

Lemma 2.1.16. — *Assume that $(S, *)$ is an LD-system. Assume that the braid word w is reversible to w' on the left, and that $(\vec{x}, \vec{y})$ is an S-colouring for w'. Then $(\vec{x}, \vec{y})$ is an S-colouring for w as well. In particular, if $(\vec{x}, \vec{y})$ is an S-colouring for $D(w)^{-1}N(w)$, it is an S-colouring for w as well.*

Proof. — It suffices to prove the result when w is reversible to w' in one step. If w' has been obtained from w by deleting some factor $\sigma_i\sigma_i^{-1}$, or replacing $\sigma_i\sigma_j^{-1}$ with $\sigma_j^{-1}\sigma_i$ in the case $|j-i| \geqslant 2$, the result is obvious. Assume that w' has been obtained from w by replacing some factor $\sigma_i\sigma_j^{-1}$ with $|j-i| = 1$ by $\sigma_j^{-1}\sigma_i^{-1}\sigma_j\sigma_i$. Assume for instance $i = 1$ and $j = 2$. So we start with a colouring of $\sigma_2^{-1}\sigma_1^{-1}\sigma_2\sigma_1$, and we wish to construct a colouring of $\sigma_1\sigma_2^{-1}$ with the same end colours. Let x_1, x_2, x_3 be the initial colours. The hypothesis that the rules of colouring are obeyed at the first crossing implies that there exists z_2 in S satisfying $x_2 = x_3 * z_2$. Similarly, there exists z_1 in S satisfying $x_1 = x_3 * z_1$. Then the colouring must be



Then, the colouring displayed on the right answers the question, as we have

$$x_3 * (z_1 * z_2) = (x_3 * z_1) * (x_3 * z_2) = x_1 * x_2. \qquad \square$$

A similar argument gives the symmetric result:

Lemma 2.1.17. — *Assume that $(S, *)$ is an LD-system. Assume that the braid word w is reversible to w' on the right, and that $(\vec{x}, \vec{y})$ is an S-colouring for w. Then $(\vec{x}, \vec{y})$ is an S-colouring for w' as well. In particular, if $(\vec{x}, \vec{y})$ is an S-colouring for w, it is an S-colouring for $\widetilde{N}(w)\widetilde{D}(w)^{-1}$ as well.*

If $(S, *)$ is an LD-system, $\vec{x}$ belongs to $S^{\mathbb{N}}$ and w is a braid word, we shall say that $\vec{x}$ is *admissible* for w if there exists at least one sequence $\vec{y}$ such that $(\vec{x}, \vec{y})$ is an S-colouring of w. It follows from Lemmas 2.1.15 and 2.1.16 that, for every LD-system S and every braid word w, there exists at least one sequence in $S^{\mathbb{N}}$ that is admissible for w, and it is not hard to extend this result to any finite family of words $w_1, \ldots, w_n$ by using a common denominator, namely a positive word u such that, for every i, the word uw_i is left reversible to some positive word. In this way, we deduce that there exists a sequence $\vec{x}$ in $S^{\mathbb{N}}$ that is admissible for all w_i's.

Now the point is the following result, which shows that all possible colourings lead to the same final colours provided we use a left cancellative LD-system:

Lemma 2.1.18. — *Assume that $(S, *)$ is a left cancellative LD-system, and w, w' are equivalent braid words. Then, for each sequence $\vec{x}$ in $S^{\mathbb{N}}$ that is admissible both for w and w', there exists exactly one sequence $\vec{y}$ such that $(\vec{x}, \vec{y})$ is both an S-colouring of w and w'.*

Proof. — Assume that $(\vec{x}, \vec{y})$ is an S-colouring of w, and $(\vec{x}, \vec{y}\,')$ is a colouring of w'.

We have observed w and w' being equivalent does not imply that their numerators and denominators are equivalent, but one shows that there must exist positive words v, v' satisfying

$$\widetilde{N}(w)v \equiv \widetilde{N}(w')v', \qquad \widetilde{D}(w)v \equiv \widetilde{D}(w')v'.$$

Now, by Lemma 2.1.17, $(\vec{x}, \vec{y})$ is also an S-colouring of the word $\widetilde{N}(w)\widetilde{D}(w)^{-1}$, and, therefore, it is an S-colouring of $\widetilde{N}(w)vv^{-1}\widetilde{D}(w)^{-1}$ too. By construction, we have

$$\vec{x} \cdot \widetilde{N}(w)v = \vec{y} \cdot \widetilde{D}(w)v.$$

The same argument gives $\vec{x} \cdot \widetilde{N}(w')v' = \vec{y}\,' \cdot \widetilde{D}(w')v'$, and, therefore, we have

$$\vec{y} \cdot \widetilde{D}(w)v = \vec{y}\,' \cdot \widetilde{D}(w')v' = \vec{y}\,' \cdot \widetilde{D}(w)v.$$

This implies $\vec{y} = \vec{y}\,'$. Indeed, it suffices for an induction to show that $\vec{x} \cdot \sigma_i = \vec{x}\,' \cdot \sigma_i$ implies $\vec{x} = \vec{x}\,'$. The hypothesis implies $x_i = x'_i$ and $x_i * x_{i+1} = x'_i * x'_{i+1}$, hence $x_{i+1} = x'_{i+1}$ provided $(S, *)$ admits left cancellation. □

The previous result applies in particular when the words w and w' coincide, and it tells us that there exists at most one colouring of w with a given initial sequence of colours. So, when $(S, *)$ is a left cancellative LD-system, there is no ambiguity in defining, for w a braid word and $\vec{x}$ in $S^{\mathbb{N}}$, the sequence $\vec{x} \cdot w$ to be the unique sequence $\vec{y}$ such that $(\vec{x}, \vec{y})$ is an S-colouring of w, if it exists. If β is a braid, we define $\vec{x} \cdot \beta$ to be $\vec{x} \cdot w$, where w is any representative braid word of β such that $\vec{x} \cdot w$ exists, if such a word exists.

We thus have extended the action of braids to all left cancellative LD-systems, at the expense of having a partial action, *i.e.*, one that need not be defined everywhere. In the case of a left cancellative LD-system that is not a rack, the hypothesis that the sequence $\vec{x} \cdot \beta$ exists does not imply that $\vec{x} \cdot w$ exists for every braid word w that represents β. The point is that, although the action is partial, there always exist sequences that are admissible for a given braid, or even, according to the remark after Lemma 2.1.17, for finitely many given braids simultaneously:

Proposition 2.1.19. — *Assume that $(S, *)$ is a left cancellative LD-system. Then, for every finite family of braids $\beta_1, \ldots, \beta_m$ in B_n, there exists at least one sequence $\vec{x}$ in S^n such that the sequence $\vec{x} \cdot \beta_i$ is defined for every i.*

2.2. Special braids

Owing to the results of Section 2.1, we can now use any left cancellative LD-system to colour braids. Free LD-systems happen to be left cancellative, and, therefore, they are relevant for this approach. On the other hand, in contrast to racks, free LD-systems are orderable, and using them to define an ordering on braids should not be a surprise. The reader could even expect a simple argument here: we use an orderable LD-system to colour braids, and deduce a braid ordering. Actually, things are not so simple, at least if we wish to justify the existence of an orderable LD-system. In the sequel, we shall simultaneously prove the orderability of braids *and* the existence of an orderable LD-system, which turn out to be two facets of one and the same result.

We shall begin with a proof of Property **C**—actually of the weaker form $\mathbf{C}_\infty$, *i.e.*, we prove that every nontrivial braid in B_∞ admits at least one representative that is σ-positive or σ-negative, but with no bound on the number of strands involved in the latter representative (such a bound can be established, but it is huge, and, in particular, not equal to the number of strands involved in the initial braid word).

2.2.1. A self-distributive operation on B_∞. — The argument used in [**34**] appeals to an LD-system consisting of braids, *i.e.*, it involves a left self-distributive operation, here denoted $*$, defined on B_∞. A crucial role is played by the closure of the singleton $\{1\}$ under $*$ in B_∞, *i.e.*, by those braids that can be expressed using 1 and $*$ exclusively. Such braids will be called *special*, and the main result is that every braid admits decompositions in terms of special braids.

We recall that sh denotes the shift endomorphism of the group B_∞, which maps σ_i to σ_{i+1} for every i.

Definition 2.2.1. — For β_1, β_2 in B_∞, we put

$$\beta_1 * \beta_2 = \beta_1 \ \mathrm{sh}(\beta_2) \ \sigma_1 \ \mathrm{sh}(\beta_1^{-1}). \tag{2.2.1}$$

We say that a braid is *special* if it can be obtained from the trivial braid 1 by using $*$ recursively; the set of all special braids is denoted B_{sp}.

So, for instance, 1, σ_1, $\sigma_2\sigma_1$, $\sigma_1^2\sigma_2^{-1}$ are special braids, as we have $\sigma_1 = 1 * 1$, $\sigma_2\sigma_1 = 1 * \sigma_1 = 1 * (1 * 1)$, $\sigma_1^2\sigma_2^{-1} = \sigma_1 * 1 = (1 * 1) * 1$.

At this point, Formula (2.2.1) comes as a rabbit out of a hat. Several justifications can be given. In particular, we shall see in Section 2.3 that (2.2.1) is the projection of some natural operation arising on an extension of B_∞ and explaining all its properties. At this point, it suffices to check the following result, a straightforward computation:

Lemma 2.2.2. — *The systems $(B_\infty, *)$ and $(B_{\mathrm{sp}}, *)$ are left cancellative LD-systems.*

So the LD-systems $(B_\infty, *)$, as well as $(B_{\mathrm{sp}}, *)$, which is, by definition, the sub-LD-system of $(B_\infty, *)$ generated by 1, are eligible for colouring braids. It is easy to check that neither $(B_\infty, *)$ nor $(B_{\mathrm{sp}}, *)$ is a rack, so the action of B_∞ on sequences from B_∞

is a partial action. As above, if $\vec{x}$ is a sequence of (special) braids, and β is a braid, we shall denote by $\vec{x} \cdot \beta$ the image of $\vec{x}$ under the action of β, when defined.

The first technically significant fact is that the action of braids on braids can be connected with a multiplication in B_∞.

Lemma 2.2.3. — *For $\vec{x}$ in $B_\infty^{\mathbb{N}}$ with finitely many nontrivial entries, define*

$$\prod^{sh}(\vec{x}) = \prod_{k=1}^{\infty} \mathrm{sh}^{k-1}(x_k) = x_1 \ \mathrm{sh}(x_2) \ \mathrm{sh}^2(x_3) \cdots . \tag{2.2.2}$$

Then, if β lies in B_∞ and $\vec{x} \cdot \beta$ exists, we have

$$\prod^{sh}(\vec{x} \cdot \beta) = \prod^{sh}(\vec{x}) \ \beta. \tag{2.2.3}$$

Proof. — We use induction on the minimal length of a braid word w representing β and such that $\vec{x} \cdot w$ exists. The result is true when w is empty. Assume $w = \sigma_i w'$. If β' is the braid represented by w', then $\vec{x} \cdot \sigma_i$ and $(\vec{x} \cdot \sigma_i) \cdot \beta'$ exist, and we find

$$\begin{aligned}
\prod^{sh}(\vec{x} \cdot \sigma_i) &= \prod^{sh}((x_1, \dots, x_i * x_{i+1}, x_i, \dots)) \\
&= x_1 \ \mathrm{sh}(x_2) \ \cdots \ \mathrm{sh}^{i-1}(x_i * x_{i+1}) \ \mathrm{sh}^i(x_i) \ \mathrm{sh}^{i+1}(x_{i+2}) \cdots \\
&= x_1 \ \mathrm{sh}(x_2) \ \cdots \ \mathrm{sh}^{i-1}(x_i) \ \mathrm{sh}^i(x_{i+1}) \ \sigma_i \ \mathrm{sh}^i(x_i)^{-1} \ \mathrm{sh}^i(x_i) \ \mathrm{sh}^{i+1}(x_{i+2}) \cdots \\
&= x_1 \ \mathrm{sh}(x_2) \ \cdots \ \mathrm{sh}^{i-1}(x_i) \ \mathrm{sh}^i(x_{i+1}) \ \sigma_i \ \mathrm{sh}^{i+1}(x_{i+2}) \cdots \\
&= x_1 \ \mathrm{sh}(x_2) \ \cdots \ \mathrm{sh}^{i-1}(x_i) \ \mathrm{sh}^i(x_{i+1}) \ \mathrm{sh}^{i+1}(x_{i+2}) \cdots \ \sigma_i = \prod^{sh}(\vec{x}) \ \sigma_i,
\end{aligned}$$

as σ_i commutes with $\mathrm{sh}^k(x)$ for $k \geqslant i+1$. Applying the induction hypothesis, we deduce $\prod^{sh}(\vec{x} \cdot \beta) = \prod^{sh}(\vec{x} \cdot \sigma_i) \ \beta' = \prod^{sh}(\vec{x}) \ \sigma_i \ \beta' = \prod^{sh}(\vec{x}) \ \beta$. The case $w = \sigma_i^{-1} w'$ is symmetric. □

It follows that the partial action of B_∞ on $B_\infty^{\mathbb{N}}$ is strongly faithful in the sense that, if there exists at least one sequence $\vec{x}$ in $B_\infty^{\mathbb{N}}$ such that $\vec{x} \cdot \beta$ and $\vec{x} \cdot \beta'$ are defined and equal, then β and β' are equal.

As the LD-system of special braids is not a rack, the action of braids on sequences of special braids is partial. However, Proposition 2.1.19 shows that, for every finite family $\beta_1, \dots, \beta_m$ in B_∞, there exists at least one sequence $\vec{x}$ of special braids such that $\vec{x} \cdot \beta_i$ is defined and consists of special braids for every i. Then applying Formula (2.2.3) in the case of special braids leads to a decomposition of every braid in terms of special braids.

Proposition 2.2.4. — *Every braid can be expressed as*

$$\cdots \ \mathrm{sh}^{i-1}(\beta_i')^{-1} \ \cdots \ \mathrm{sh}(\beta_2')^{-1} \ {\beta_1'}^{-1} \ \beta_1 \ \mathrm{sh}(\beta_2) \ \cdots \ \mathrm{sh}^{i-1}(\beta_i) \ \cdots \tag{2.2.4}$$

where $\beta_1', \beta_2', \dots, \beta_1, \beta_2, \dots$ are special braids.

Proof. — Assume first $\beta \in B_\infty^+$. Then the sequence $(1, 1, \dots) \cdot \beta$ is defined, since possible obstructions occur only with negative crossings, and it consists of special braids, as special braids are closed under $*$. Assume $(1, 1, \dots) \cdot \beta = (\beta_1, \beta_2, \dots)$. Then (2.2.3) gives

$$\prod^{sh}((\beta_1, \beta_2, \dots)) = \prod^{sh}((1, 1, \dots))\, \beta,$$

i.e., $\beta = \beta_1 \mathrm{sh}(\beta_2) \mathrm{sh}^2(\beta_3) \cdots$. By Proposition 1.1.5, every braid is the quotient of two positive braids, so the general result follows. □

It can be observed that, for β a positive braid, the expression of β as $\prod^{sh}(\vec{x})$ where $\vec{x}$ is a sequence of special braids is unique.

2.2.2. Property $\mathbf{C}_{LD}$. — We thus have obtained a decomposition of every braid in terms of special braids. Now, by construction, special braids form a *monogenic* LD-system, namely one generated by the unique braid 1. Then Property **C** for braids will be a consequence of a general property of monogenic LD-systems and the associated left divisibility relation.

Definition 2.2.5. — For $(S, *)$ a binary system, and x, y in S, we say that x is an *iterated left divisor* of y, and write $x \sqsubset y$, if there exists some positive integer p, and some elements $z_1, \dots, z_p$ in S satisfying

$$y = (\dots((x * z_1) * z_2) * \dots) * z_p. \tag{2.2.5}$$

Let us introduce the following statement, which will be proved below:

Property $\mathbf{C}_{LD}$ (Comparison for LD). — *Assume that $(S, *)$ is a monogenic LD-system. Then any two elements of S are comparable with respect to iterated left divisibility, i.e., for all x, y in S, at least one of $x = y$, $x \sqsubset y$, $y \sqsubset x$ is true.*

As $(B_{\mathrm{sp}}, *)$ is a monogenic LD-system, Property $\mathbf{C}_{LD}$ implies that any two special braids are comparable with respect to iterated divisibility. Now, due to the explicit definition of the operation $*$ on braids, iterated divisibility in $(B_{\mathrm{sp}}, *)$ is connected with occurrences of σ_1 and σ_1^{-1}:

Lemma 2.2.6. — *Assume that β_1, β_2 are special braids satisfying $\beta_1 \sqsubset \beta_2$. Then the braid $\beta_1^{-1}\beta_2$ admits a σ_1-positive representative braid word.*

Proof. — By construction, there exist special braids $\beta_1', \dots, \beta_p'$ satisfying

$$\beta_2 = (\cdots((\beta_1 * \beta_1') * \beta_2') * \cdots) * \beta_p'.$$

Expanding the latter product gives

$$\beta_2 = \beta_1 \ \mathrm{sh}(\beta_1') \ \sigma_1 \ \mathrm{sh}(\beta_1'') \ \sigma_1 \ \dots \ \sigma_1 \ \mathrm{sh}(\beta_p''),$$

with $\beta_1'' = \beta_1^{-1}\beta_2'$, $\beta_k'' = ((\dots((\beta_1 * \beta_1') * \beta_2') * \dots) * \beta_k')^{-1}\beta_{k+1}'$ for $2 \leqslant k < p$, and $\beta_p'' = ((\dots((\beta_1 * \beta_1') * \beta_2') * \dots) * \beta_p')^{-1}$. Hence the braid $\beta_1^{-1}\beta_2$ admits a representative braid word containing p letters σ_1 and no letter σ_1^{-1}. □

Proposition 2.2.7. — *Property* $\mathbf{C}_{LD}$ *implies Property* $\mathbf{C}_\infty$*, i.e., it implies that every braid in* B_∞ *admits a representative braid word that is* σ_1*-positive,* σ_1*-negative, or* σ_1*-free.*

Proof. — Let $\beta \in B_\infty$. By Proposition 2.2.4, the braid β admits a decomposition

$$\cdots \operatorname{sh}^{i-1}(\beta_i')^{-1} \ \cdots \ \operatorname{sh}(\beta_2')^{-1} \ {\beta_1'}^{-1} \ \beta_1 \ \operatorname{sh}(\beta_2) \ \cdots \ \operatorname{sh}^{i-1}(\beta_i) \cdots$$

where $\beta_1, \beta_2, \ldots, \beta_1', \beta_2', \ldots$ are special braids. By Property $\mathbf{C}_{LD}$, at least one of the following three cases occur:

(i) $\beta_1' \sqsubset \beta_1$: then, by Lemma 2.2.6, ${\beta_1'}^{-1}\beta_1$ admits a σ_1-positive representative, and, therefore, so does β, as adding letters $\sigma_i^{\pm 1}$ with $i \geqslant 2$ cannot destroy σ_1-positivity;

(ii) $\beta_1' = \beta_1$: then β belongs to the image of sh, so it has σ_1-free representatives;

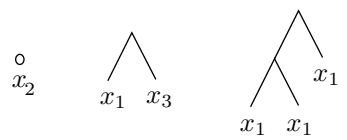
(iii) $\beta_1' \sqsupset \beta_1$: then, as in (i), β admits a σ_1-negative representative.

Thus Property $\mathbf{C}_\infty$ is proved. □

2.2.3. A proof of Property $\mathbf{C}_{LD}$. — It remains to establish Property $\mathbf{C}_{LD}$. To this end, we follow the argument of [**31**], building on results of [**30**].

Easy uniform arguments show that, if $\vec{I}$ is a family of algebraic identities and X is a nonempty set, then there exists a free $\vec{I}$-system based on X, *i.e.*, an algebraic system S consisting of a set equipped with operations satisfying $\vec{I}$ such that X generates S and every $\vec{I}$-system generated by X is a homomorphic image of S. Moreover, S is unique up to isomorphism. So, in particular, there exists for every n a free LD-system of rank n, that will be denoted by $\mathcal{F}_n$ here.

Free LD-systems can be described from absolutely free systems—or free magmas in the terminology of [**9**]. For $1 \leqslant n \leqslant \infty$, we denote by T_n the free magma generated by $x_1, \ldots, x_n$, *i.e.*, the set of all well-formed expressions constructed using the variables $x_1, \ldots, x_n$ and a binary operator, here $*$. The elements of T_n are called *terms*. Thus, x_2, $x_1 * x_3$, $(x_1 * x_1) * x_1$ are terms, while $x_2 x_1))*$ is not a term. It will be useful to see terms as *labelled binary trees*: every term that is not a variable has well defined left and right subterms, and we inductively define the tree associated with a term t to be the binary tree consisting of a root with two successors, namely a left subtree which is the tree associated with the left subterm of t, and a right subtree which is the tree associated with the right subterm of t. For instance, the trees associated with the terms above are



respectively.

Definition 2.2.8. — For t, t' terms in T_∞, we say that t' is an *LD-expansion* of t (*resp.* is *LD-equivalent* to t, denoted $t' =_{LD} t$) if we can go from t to t' by applying

finitely many transformations consisting in replacing a subterm of the form $t_1 * (t_2 * t_3)$ with the corresponding term $(t_1 * t_2) * (t_1 * t_3)$ (*resp.* in applying the previous transformations and their inverses).

By definition, t' being LD-equivalent to t means that we can transform t to t' by applying Identity (LD), *i.e.*, (2.0.1), while t' being an LD-expansion of t means that we transform t to t' by applying Identity (LD), but only in the expanding direction, *i.e.*, from $x * (y * z)$ to $(x * y) * (x * z)$, but not in the converse, contracting direction.

The following result is straightforward:

Lemma 2.2.9. — *For every n, LD-equivalence is a congruence on T_n, and $T_n/{=_{LD}}$ is a free LD-system of rank n, i.e., it is a realization of $\mathcal{F}_n$.*

Our aim is to prove Property $\mathbf{C}_{LD}$ in every monogenic LD-system. As every monogenic LD-system is a homomorphic image of the free LD-system $\mathcal{F}_1$, and Property $\mathbf{C}_{LD}$ is preserved under homomorphism, it suffices that we prove $\mathbf{C}_{LD}$ in $\mathcal{F}_1$. To this end, we shall work with terms.

Definition 2.2.10. — For t a term and k sufficiently small, we denote by $\mathrm{left}^k(t)$ the kth iterated left subterm of t: we have $\mathrm{left}^0(t) = t$ for every t, and $\mathrm{left}^k(t) = \mathrm{left}^{k-1}(t_1)$ for $t = t_1 * t_2$ and $k \geqslant 1$. For t_1, t_2 in T_∞, we say that $t_1 \sqsubset_{LD} t_2$ is true if we have $t'_1 = \mathrm{left}^k(t'_2)$ for some k, t'_1, t'_2 satisfying $k \geqslant 1$, $t'_1 =_{LD} t_1$, and $t'_2 =_{LD} t_2$.

By construction, saying that $t_1 \sqsubset_{LD} t_2$ is true in T_1 is equivalent to saying that the class $\overline{t_1}$ of t_1 in $\mathcal{F}_1$ is an iterated left divisor of the class $\overline{t_2}$ of t_2. So proving Property $\mathbf{C}_{LD}$ amounts to proving that, for any two terms t_1, t_2 in T_1, at least one of

$$t_1 \sqsubset_{LD} t_2, \quad t_1 =_{LD} t_2, \quad t_1 \sqsupset_{LD} t_2$$

is true. The argument relies on three specific properties of left self-distributivity.

If t is a term, we define inductively $t^{[1]} = t$, $t^{[k+1]} = t * t^{[k]}$, and, in the sequel, we write x instead of x_1 for the variable of terms in T_1.

Lemma 2.2.11. — *For every term t in T_1, we have $x^{[k+1]} =_{LD} t * x^{[k]}$ for k sufficiently large.*

Proof. — We use induction on t. For $t = x$, we have $x^{[k+1]} = x * x^{[k]}$ for every k, by definition. Assume now $t = t_1 * t_2$. Assuming that the result is true for t_1 and t_2, we obtain for k sufficiently large

$$\begin{aligned} x^{[k+1]} =_{LD} t_1 * x^{[k]} &=_{LD} t_1 * (t_2 * x^{[k-1]}) \\ &=_{LD} (t_1 * t_2) * (t_1 * x^{[k-1]}) =_{LD} (t_1 * t_2) * x^{[k]} = t * x^{[k]}, \end{aligned}$$

which is the result for t. □

Lemma 2.2.12. — *If t is a term in T_∞ such that* $\mathrm{left}^k(t)$ *is defined, and if t' is an LD-expansion of t, then there exists $k' \geqslant k$ such that* $\mathrm{left}^{k'}(t')$ *is an LD-expansion of* $\mathrm{left}^k(t)$.

Proof. — It suffices to prove the result when t' is obtained by replacing exactly one subterm t_0 of t of the form $t_1 * (t_2 * t_3)$ with the corresponding $(t_1 * t_2) * (t_1 * t_3)$. If t_0 is $\mathrm{left}^j(t)$ with $j \leqslant k$, then $\mathrm{left}^{k+1}(t')$ is equal to $\mathrm{left}^k(t)$; if t_0 is $\mathrm{left}^j(t)$ with $j > k$, then $\mathrm{left}^k(t')$ is an LD-expansion of $\mathrm{left}^k(t)$; otherwise, we have $\mathrm{left}^k(t') = \mathrm{left}^k(t)$. □

Lemma 2.2.13. — *Two LD-equivalent terms in T_∞ admit a common LD-expansion.*

Proof (sketch, see also Lemma 2.3.5). — The point is to prove that, if t' and t'' are any two LD-expansions of some term t, then t' and t'' admit a common LD-expansion. Now, let us say that t' is a k-expansion of t if t' is obtained from t by applying Identity (LD) at most k times in the expanding direction. Then, for every term t, one can explicitly define an LD-expansion ∂t of t that is a common LD-expansion of all 1-expansions of t, and check that, if t' is an LD-expansion of t, then $\partial t'$ is an LD-expansion of ∂t. The definition of the term ∂t is inductive: ∂t equals t when t is a variable, and $\partial(t_1 * t_2)$ is obtained by replacing each variable x_i in ∂t_2 with the term $\partial t_1 * x_i$. The reader can check for instance that $\partial x^{[3]}$ is equal to $((x*x)*(x*x))*((x*x)*(x*x))$—and that $\partial^2 x^{[3]}$ is a complicated term with 42 variables. Then, one shows using an induction that, for every k, the term $\partial^k t$ is an LD-expansion of all k-expansions of t. It follows that, if t' and t'' are any two LD-expansions of some term t, then t' and t'' admit common LD-expansions, namely all terms $\partial^k t$ with k sufficiently large. □

Remark 2.2.14. — The previous argument has much in common with the property that any two braids in the monoid B_n^+ admit a least common right multiple—the latter result easily follows from Lemma 2.2.13 using the group G_{LD} introduced in Section 2.3 below. Here the terms $\partial^k t$ play the role of Garside's fundamental braids Δ_n^k mentioned in Proposition 1.1.5. Garside's argument relies on the fact that Δ_n is a common right multiple of all braids in B_n^+ where any two strands cross at most once ("simple braids"), which then implies that Δ_n^k is a common right multiple for all braids in B_n^+ that are products of at most k simple braids. Going to a 1-expansion is analogous to applying a simple braid, and, similarly, ∂t is a common LD-expansion of all 1-expansions of t, and, then, $\partial^k t$ is a common LD-expansion of all LD-expansions of t that can be decomposed in at most k 1-expansions. One of the main differences—explaining some technical difficulties—is that, in the case of B_n^+, the index n remains fixed, while, in the case of LD, the analogous parameter, namely the term t, changes at each step—see [**40**] for a more complete explanation.

We are now ready to conclude.

***Proposition 2.2.15* (Property $\mathbf{C}_{LD}$).** — *If $(S,*)$ is a monogenic LD-system, then any two elements of S are comparable with respect to iterated left division.*

Proof. — Let t_1, t_2 be arbitrary terms in T_1. By Lemma 2.2.11, we have $t_1 * x^{[k]} =_{LD} x^{[k+1]} =_{LD} t_2 * x^{[k]}$ for k sufficiently large. Fix such a k. By Lemma 2.2.13, the terms $t_1 * x^{[k]}$ and $t_2 * x^{[k]}$ admit a common LD-expansion, say t. By Lemma 2.2.12, there exist nonnegative integers k_1, k_2 such that, for $i = 1, 2$, the term $\mathrm{left}^{k_i}(t)$ is an LD-expansion of $\mathrm{left}(t_i * x^{[k]})$, *i.e.*, of t_i. Thus we have $t_1 =_{LD} \mathrm{left}^{k_1}(t)$, and $t_2 =_{LD} \mathrm{left}^{k_2}(t)$. Then, up to swapping the indices 1 and 2, the term $\mathrm{left}^{k_1}(t)$ is an iterated left subterm of $\mathrm{left}^{k_2}(t)$, *i.e.*, we have $\mathrm{left}^{k_1}(t) \sqsubseteq \mathrm{left}^{k_2}(t)$, and, therefore, $t_1 \sqsubset_{LD} t_2$ or $t_1 =_{LD} t_2$. □

By Proposition 2.2.7, we deduce

***Corollary 2.2.16* (Property $\mathbf{C}_\infty$).** — *Every braid in B_∞ admits a representative braid word that is σ_1-positive, σ_1-negative, or σ_1-free.*

2.2.4. A proof of Property C. — The previous argument establishes Property $\mathbf{C}_\infty$ only: starting with a braid β in B_n, we can find a special decomposition of β in terms of special braids, but there is no reason that the involved special braids lie in B_n—and they do not in general. Let us mention that Richard Laver, using normal forms methods in free LD-systems analogous to those mentioned in Section 2.4 below, has given (around 1994) a proof of Property **C** itself (for each fixed group B_n) [unpublished work].

2.3. The group of left self-distributivity

We turn to Property **A**, *i.e.*, to the result that a σ_1-positive braid word never represents 1. We sketch the argument first given in [**34**]. This argument is a little delicate, but, on the other hand, it may be argued that it is beautifully structural, and that it deeply analyses the connection between braids and left self-distributivity.

2.3.1. Property $\mathbf{A}_{LD}$. — Let us introduce the following statement:

Property $\mathbf{A}_{LD}$ (Acyclicity for LD). — *There exists an acyclic left cancellative LD-system, i.e., one where no equality of the form* (2.1.6) *is true.*

Proposition 2.3.1. — *Property $\mathbf{A}_{LD}$ implies Property **A**, i.e., it implies that a σ_1-positive braid word never represents* 1*; more generally, Property $\mathbf{A}_{LD}$ implies Property $\mathbf{A}_i$ for every i, i.e., it implies that a braid word containing at least one letter σ_i and no letter σ_i^{-1} never represents* 1.

Proof. — Assume that $(S,*)$ is an acyclic left cancellative LD-system. Let w be a σ_1-positive braid word. By Proposition 2.1.19, there exists a sequence of colours $\vec{x}$ in S such that $\vec{x} \cdot w$ is defined. Call the initial left colour z_0, call the colour of the

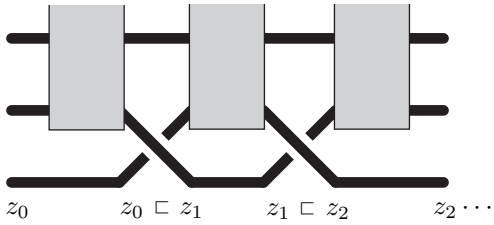


FIGURE 2.3. A σ_1-positive braid word is never trivial

first strand after the first σ_1 crossing z_1, etc. (see Figure 2.3). By construction, z_i is a left divisor of z_{i+1} for each i. The hypothesis that left division in $(S, *)$ has no cycle, implies the final colour z_n cannot be equal to the initial colour z_0, hence, by Lemma 2.1.18, w cannot represent 1.

The argument for Property $\mathbf{A}_i$ is similar: the parameter that increases at each letter σ_i and is preserved under each letter $\sigma_j^{\pm 1}$ with $j \neq i$ is the $*$-product of the i first colours, *i.e.*, $x_1 * (x_2 * (\dots (x_{i-1} * x_i) \dots))$ for $\vec{x}$. □

So we are left with proving Property $\mathbf{A}_{LD}$. Before turning to the proof, let us observe that the implication of Proposition 2.3.1 is actually an equivalence:

Proposition 2.3.2. — *Property* $\mathbf{A}$ *implies Property* $\mathbf{A}_{LD}$.

Proof. — We know that $(B_\infty, *)$ is a left cancellative LD-system, and we have seen in Lemma 2.2.6 that the braid ${\beta'}^{-1}\beta$ admits a σ_1-positive representative whenever $\beta' \sqsubset \beta$ holds. If Property $\mathbf{A}$ is known to be true, we deduce that $\beta' \sqsubset \beta$ implies $\beta' \neq \beta$, *i.e.*, that left division in $(B_\infty, *)$ has no cycle. So $(B_\infty, *)$ is an acyclic left cancellative LD-system. □

It follows that any of the proofs of Property $\mathbf{A}$, *i.e.*, of Property $\mathbf{A}_1$, developed in the subsequent chapters gives a new proof of the result that left division in the LD-system $(B_\infty, *)$ has no cycle, and, therefore, of Properties $\mathbf{A}_{LD}$ and $\mathbf{A}_i$ for each i.

2.3.2. The geometry of Identity (LD)**.** — The original proof of Property $\mathbf{A}_{LD}$ given in [**34**] consists in showing that free LD-systems satisfy the required conditions, *i.e.*, they are left cancellative and acyclic, and actually orderable. As one example only is needed, we shall consider monogenic free LD-systems.

We use the formalism of Section 2.2. So T_1 denotes the set of all formal terms built using the variable x and the operator $*$, and write $\mathcal{F}_1$ for $T_1/{=_{LD}}$. Then $\mathcal{F}_1$ is a free monogenic LD-system, and saying that it is acyclic amounts to proving

Proposition 2.3.3. — *For* t, t' *in* T_1*, the relations* $t \sqsubset_{LD} t'$ *and* $t =_{LD} t'$ *exclude each other.*

Our aim until the end of this section is to prove Proposition 2.3.3. To this end, we shall analyse what can be called the *geometry* of Identity (LD). Saying that two terms t, t' are LD-equivalent means that we can transform t into t' by repeatedly replacing a subterm of the form $t_1*(t_2*t_3)$ with the corresponding term $(t_1*t_2)*(t_1*t_3)$, or vice versa, *i.e.*, by applying Identity (LD). The idea will be to take into account the *position* in the terms where the identity is applied. To this end, we fix a system of addresses for the subterms of a term. A simple system is obtained by viewing terms as binary trees and describing the path that goes from the root of the tree to the root of the considered subterm using for instance 0 for forking to the left and 1 for forking to the right. We denote by $\mathbf{A}$ the set of all such addresses, *i.e.*, the set of all finite sequences of 0's and 1's, and we use $\varnothing$ for the empty address. This allows us to speak of the αth subterm of a given term t. Notice that, for each term t, the αth subterm of t is defined for finitely many addresses α only.

Definition 2.3.4. — For α an address, we denote by LD_α the (partial) operator on terms corresponding to applying Identity (LD) to the αth subterm, in the expanding direction. We denote by $\mathcal{G}_{LD}$ the monoid generated by all operators LD_α and their inverses using reversed composition (we think of $\mathcal{G}_{LD}$ as acting on terms on the right, writing $t \cdot f$ for the result of applying f to t).

The operator LD_α is a partial operator: the term t belongs to the domain of LD_α if and only if the αth subterm of t exists, and it can be decomposed as $t_1 * (t_2 * t_3)$, in which case $t \cdot \mathit{LD}_\alpha$ is the term obtained by replacing the above subterm with $(t_1 * t_2) * (t_1 * t_3)$ in t. The operator LD_α is injective, and its inverse LD_α^{-1} corresponds to applying Identity (LD) at α in the contracting direction. For every term t, the set of those addresses α's such that $t \cdot \mathit{LD}_\alpha$ exists is finite. For instance, if t is the term $x_1 * (x_2 * (x_3 * x_4))$, then $t \cdot \mathit{LD}_\varnothing$ and $t \cdot \mathit{LD}_1$ only are defined, and the values are $(x_1 * x_2) * (x_1 * (x_3 * x_4))$ and $x_1 * ((x_2 * x_3) * (x_2 * x_4))$ respectively.

By construction, two terms t, t' are LD-equivalent if and only if some element of the monoid $\mathcal{G}_{LD}$ maps t to t'.

Now, we would like to replace the monoid $\mathcal{G}_{LD}$ with a group. Because it consists of partial operators, the monoid $\mathcal{G}_{LD}$ is an inverse monoid, but not a group. Moreover, some composed operators are empty: for instance, the branches $00\ldots$ and $10\ldots$ must have the same length in any term belonging to the image of $\mathit{LD}_\varnothing$, *i.e.*, to the domain of $\mathit{LD}_\varnothing^{-1}$, and this is never the case for a term in the image of $\mathit{LD}_\varnothing \cdot \mathit{LD}_1$, so the operator $\mathit{LD}_\varnothing \cdot \mathit{LD}_1 \cdot \mathit{LD}_\varnothing^{-1}$ is just empty. It follows that there is no way to quotient $\mathcal{G}_{LD}$ into a nontrivial group without distorting it completely. So, we resort to an indirect approach, namely guessing a presentation of $\mathcal{G}_{LD}$, and then introducing the group G_{LD} defined by this presentation: the idea is that, if we guessed the presentation correctly, then the group G_{LD} should resemble the monoid $\mathcal{G}_{LD}$, and all results about the action of $\mathcal{G}_{LD}$ on terms should admit purely syntactic counterparts in G_{LD}. The first step is therefore to find relations between the various operators LD_α.

Lemma 2.3.5. — *For all addresses α, β, γ, the following relations hold in $\mathcal{G}_{LD}$:*

$$LD_{\alpha 0\beta} \cdot LD_{\alpha 1\gamma} = LD_{\alpha 1\gamma} \cdot LD_{\alpha 0\beta},$$
$$LD_{\alpha 0\beta} \cdot LD_{\alpha} = LD_{\alpha} \cdot LD_{\alpha 00\beta} \cdot LD_{\alpha 10\beta},$$
$$LD_{\alpha 10\beta} \cdot LD_{\alpha} = LD_{\alpha} \cdot LD_{\alpha 01\beta},$$
$$LD_{\alpha 11\beta} \cdot LD_{\alpha} = LD_{\alpha} \cdot LD_{\alpha 11\beta},$$
$$LD_{\alpha 1} \cdot LD_{\alpha} \cdot LD_{\alpha 1} \cdot LD_{\alpha 0} = LD_{\alpha} \cdot LD_{\alpha 1} \cdot LD_{\alpha}.$$

The verification is easy. The above relations are quite natural: they are nothing but a syntactic counterpart to Lemma 2.2.13. Notice that, for each pair of addresses α, β, there exists exactly one relation in the list above taking the form $LD_{\alpha} \cdots = LD_{\beta} \cdots$, *i.e.*, explaining how to obtain a common LD-expansion for $t \cdot LD_{\alpha}$ and $t \cdot LD_{\beta}$. According to the strategy sketched above, we introduce

Definition 2.3.6. — We denote by G_{LD} the group generated by an infinite sequence of generators τ_{α} indexed by addresses, *i.e.*, by finite sequence of 0's and 1's, subject to the relations of Lemma 2.3.5, *i.e.*, $\tau_{\alpha 0\beta} \cdot \tau_{\alpha 1\gamma} = \tau_{\alpha 1\gamma} \cdot \tau_{\alpha 0\beta}$, etc.

Remark 2.3.7. — The braid group B_{∞} is a quotient of the group G_{LD}: indeed, mapping τ_{1^n} to σ_{n-1} and collapsing all generators τ_{α} such that the address α contains at least one 0 defines a surjective homomorphism, as can be seen by comparing the relations in Lemma 2.3.5 with the braid relations. The existence of this homomorphism is the core of the connection between braids and left self-distributivity.

Let us denote by $\mathcal{G}_{LD}^{+}$ (*resp.* G_{LD}^{+}) the submonoid of $\mathcal{G}_{LD}$ (*resp.* G_{LD}) generated by the elements LD_{α} (*resp.* τ_{α}), *i.e.*, we forbid inverses. Lemma 2.3.5 implies that $\mathcal{G}_{LD}^{+}$ is a quotient of G_{LD}^{+}, so the action of $\mathcal{G}_{LD}^{+}$ on terms factors through an action of G_{LD}^{+} (we shall denote by $t \cdot x$ the result of letting x act on t). This action however does not extend to the group G_{LD}, as $\mathcal{G}_{LD}$ is not a quotient of G_{LD}, due to the fact that the composition of two operators in $\mathcal{G}_{LD}$ may be empty.

Remark 2.3.8. — The previous approach applies to every algebraic identity, and, more generally, to every family of algebraic identities [**33, 35, 43**]: in each case, some monoid describes the associated geometry, and, in good cases, a group is involved. In the case of associativity, the involved group happens to be Richard Thompson's group F investigated in [**107, 20**].

2.3.3. The blueprint of a term. — The core of the argument for proving Property $\mathbf{A}_{LD}$ is the following observation: Lemma 2.2.11 tells us that, for every term t in T_1, the relation $x^{[k+1]} =_{LD} t * x^{[k]}$ holds for k sufficiently large. By construction, this means that some element f of the monoid $\mathcal{G}_{LD}$ (depending on t and, *a priori*, on k) maps $x^{[k+1]}$ to $t * x^{[k]}$, *i.e.*, in some sense, constructs the term t from the universal term $x^{[k+1]}$. Moreover, the inductive proof of Lemma 2.2.11 gives us an explicit

definition of f. Indeed, assume that t equals $t_1 * t_2$, that f_1 maps $x^{[k+1]}$ to $t_1 * x^{[k]}$, and that f_2 maps $x^{[k]}$ to $t_2 * x^{[k-1]}$. For g in $\mathcal{G}_{LD}$, let $\mathrm{sh}_1(g)$ denote the shifted version of g consisting of applying g to the right subterm of its argument. Then we read on the sequence

$$x^{[k+1]} \xrightarrow{f_1} t_1 * x^{[k]} \xrightarrow{\mathrm{sh}_1(f_2)} t_1 * (t_2 * x^{[k-1]}) \xrightarrow{LD_\varnothing} (t_1 * t_2) * (t_1 * x^{[k-1]}) \xrightarrow{\mathrm{sh}_1(f_1^{-1})} t * x^{[k]}$$

that the operator

$$f_1 \cdot \mathrm{sh}_1(f_2) \cdot LD_\varnothing \cdot \mathrm{sh}_1(f_1^{-1})$$

maps $x^{[k+1]}$ to $t * x^{[k]}$. In this way, we obtain

Lemma 2.3.9. — *For t in T_1, define χ_t in $\mathcal{G}_{LD}$ inductively by $\chi_x = \mathrm{id}$ and*

$$\chi_{t_1 * t_2} = \chi_{t_1} \cdot \mathrm{sh}_1(\chi_{t_2}) \cdot LD_\varnothing \cdot \mathrm{sh}_1(\chi_{t_1}^{-1}). \tag{2.3.1}$$

Then, for every t, and for k sufficiently large, the operator χ_t maps $x^{[k+1]}$ to $t * x^{[k]}$.

According to our strategy, we introduce the counterpart $[\![t]\!]$ of the operator χ_t in G_{LD}: $[\![t]\!]$ should be seen as a sort of copy of t inside G_{LD}. We denote by sh_1 (*resp.* sh_0) the (left) shift endomorphism of G_{LD} that maps τ_α to $\tau_{1\alpha}$ (*resp.* $\tau_{0\alpha}$) for every α.

Definition 2.3.10. — For t in T_1, the *blueprint* of t is the element $[\![t]\!]$ of G_{LD} inductively defined by $[\![x]\!] = 1$ and

$$[\![t_1 * t_2]\!] = [\![t_1]\!] \cdot \mathrm{sh}_1([\![t_2]\!]) \cdot \tau_\varnothing \cdot \mathrm{sh}_1([\![t_1]\!]^{-1}). \tag{2.3.2}$$

We recall that our aim is to prove that the relations $t =_{LD} t'$ and $t \sqsubset_{LD} t'$ exclude each other. To see that, we translate them to G_{LD} using $[\![t]\!]$ as a counterpart to t.

Lemma 2.3.11. — (i) *If $t =_{LD} t'$ is satisfied, then $[\![t]\!]^{-1} \cdot [\![t']\!]$ lies in the subgroup of G_{LD} generated by the elements $\tau_{0\alpha}$.*
(ii) *If $t \sqsubset_{LD} t'$ is satisfied, then $[\![t]\!]^{-1} \cdot [\![t']\!]$ admits an expression where $\tau_\varnothing$ occurs, but $\tau_\varnothing^{-1}$ does not.*

Proof (sketch). — (i) First, observe that the statement is natural. Indeed, we know that, if $t =_{LD} t'$ holds, then some element f of $\mathcal{G}_{LD}$ maps t to t', and, therefore, $\mathrm{sh}_0(f)$ maps $t * x^{[k]}$ to $t' * x^{[k]}$. Now, by construction, the operator $\chi_t^{-1} \cdot \chi_{t'}$ also maps $t * x^{[k]}$ to $t' * x^{[k]}$. This means that the operator $\chi_t^{-1} \cdot \chi_{t'}$ coincides with some operator $\mathrm{sh}_0(f)$ on at least one term, which, by substitution arguments, implies that it does on every term where defined. If our axiomatization is correct, it should therefore be true that the element $[\![t]\!]^{-1} \cdot [\![t']\!]$ of G_{LD} coincides with some element of the form $\mathrm{sh}_0(a)$, *i.e.*, belongs to the subgroup $\mathrm{sh}_0(G_{LD})$ of G_{LD}. The point is that, if the latter statement is true, it must be provable by a direct verification. This is exactly what happens: for instance, the verification corresponding to $f = LD_\varnothing$ is the identity

$$[\![(t_1 * t_2) * (t_1 * t_3)]\!] = [\![t_1 * (t_2 * t_3)]\!] \cdot \tau_\varnothing,$$

which follows from the definition of the blueprint and the defining relations of G_{LD}.

(ii) By (i), it is sufficient to prove the relation for $t \sqsubset t'$, and even for $t' = t * t_0$, as an induction then gives the result. Now, the definition gives

$$[\![t]\!]^{-1} \cdot [\![t']\!] = \mathrm{sh}_1([\![t_0]\!]) \cdot \tau_{\varnothing} \cdot \mathrm{sh}_1([\![t]\!]^{-1}),$$

which has the desired form: $\tau_{\varnothing}$ occurs, but $\tau_{\varnothing}^{-1}$ does not. □

2.3.4. A preordering on G_{LD}. — It therefore remains to *separate* the elements of G_{LD} according to whether they can be expressed with or without $\tau_{\varnothing}$ and $\tau_{\varnothing}^{-1}$. To this end, we use an order argument in G_{LD} relying on a study of where left subterms are mapped.

By definition, a term t' in T_∞ is an LD-expansion of another term t if some element of $\mathcal{G}_{LD}^+$ maps t to t'. In this framework, Lemma 2.2.13 can be rephrased as the result that every element in $\mathcal{G}_{LD}$ can be expressed as a fraction fg^{-1} with f, g in $\mathcal{G}_{LD}^+$. The proof uses the relations of Lemma 2.3.5 only, so it applies to the group G_{LD}:

Lemma 2.3.12. — *Every element of the group G_{LD} can be expressed as ab^{-1} with a, b in G_{LD}^+.*

Then, Lemma 2.2.12 tells us that, if t' is an LD-expansion of t, then, for every k such that t has a kth iterated left subterm, there exists k' such that the k'th iterated left subterm of t' is an LD-expansion of the kth iterated left subterm of t. It is easy to see that the index k' only depends on k and on the element f of $\mathcal{G}_{LD}^+$ that describes the passage from t to t'. More precisely, there exist two maps $\delta : \mathbb{N} \times \mathcal{G}_{LD}^+ \to \mathbb{N}$ and $\pi : \mathcal{G}_{LD}^+ \times \mathbb{N} \to \mathcal{G}_{LD}^+$ such that, for every f in $\mathcal{G}_{LD}^+$, every term t, and every k sufficiently small, we have

$$\mathrm{left}^{\delta(k,f)}(t \cdot f) = \mathrm{left}^k(t) \cdot \pi(f, k). \tag{2.3.3}$$

This once again can be proved using the relations of Lemma 2.3.5 exclusively, so we can obtain a counterpart in G_{LD}^+ (once again, if the property is true—and it is!—its proof is a simple verification):

Lemma 2.3.13. — *There exists two maps $d : \mathbb{N} \times G_{LD}^+ \to \mathbb{N}$ and $p : G_{LD}^+ \times \mathbb{N} \to G_{LD}^+$ such that, for every a in G_{LD}^+, every term t, and every k sufficiently small, we have*

$$\mathrm{left}^{d(k,a)}(t \cdot a) = \mathrm{left}^k(t) \cdot p(a, k). \tag{2.3.4}$$

We are ready to conclude the argument.

Proof of Proposition 2.3.3 (sketch). — Let us introduce two subsets of G_{LD} as follows: We say that an element c of G_{LD} belongs to $P_<$ (*resp.* $P_=$) if there exists a decomposition $c = ab^{-1}$ with a, b in G_{LD}^+ satisfying

$$d(1, b) < d(1, a) \quad (\text{resp.} =). \tag{2.3.5}$$

It is not hard to prove that the sets $P_<$ and $P_=$ are disjoint and that they are closed under product. The point is to prove that, if an element c of G_{LD} decomposes into $c = ab^{-1} = a'b'^{-1}$ with a, b, a', b' in G_{LD}^+ and we have, say, $d(1,b) < d(1,a)$, then we necessarily have $d(1,b') < d(1,a')$. Now, under the above hypotheses, we have $ag = a'g'$ and $bg = b'g'$ for some g, g' in G_{LD}^+, which implies $d(1,ag) = d(1,a'g')$, *i.e.*,

$$d(d(1,a),g) = d(d(1,a'),g'),$$

and, similarly,

$$d(d(1,b),g) = d(d(1,b'),g').$$

As the mappings $d(\cdot,g)$ and $d(\cdot,g')$ are increasing—this is the point, and it is natural as its geometric counterpart for the map δ is obvious—it is clear that $d(1,b) < d(1,a)$ is equivalent to $d(1,b') < d(1,a')$.

Then, one can easily show that $P_=$ is closed under inverse, and prove the inclusions $P_= \cdot P_< \subseteq P_<$, and $P_< \cdot P_= \subseteq P_<$ (thus the relation $c^{-1}c' \in P_< \cup P_=$ defines a preordering on G_{LD},*i.e.*, a reflexive and transitive relation,that is compatible with multiplication on the left, and $P_=$ is the associated equivalence relation). Let us consider the generators τ_α. First, we decompose $\tau_\varnothing$ as the fraction with trivial denominator $\tau_\varnothing = \tau_\varnothing \cdot 1^{-1}$, and find

$$d(1,1) = 1 < d(1,\tau_\varnothing) = 2,$$

hence $\tau_\varnothing$ belongs to $P_<$. On the other hand, we have for each α

$$d(1,1) = 1 = d(1,\tau_{0\alpha}) = d(1,\tau_{1\alpha}),$$

hence $\tau_{0\alpha}$ and $\tau_{1\alpha}$ belong to $P_=$. Now, if c belongs to the subgroup $\mathrm{sh}_0(G_{LD})$ of G_{LD}, then, by the above computation, c belongs to $P_=$. On the other hand, if c admits an expression where $\tau_\varnothing$ occurs but $\tau_\varnothing^{-1}$ does not, c belongs to $P_<$. We conclude that the two cases exclude each other. □

We are ready to conclude the argument.

Proposition 2.3.14. — *The free LD-system $\mathcal{F}_1$ is acyclic and cancellative.*

Proof. — Proposition 2.3.3 tells us that $\mathcal{F}_1$ is acyclic. It remains to prove that $\mathcal{F}_1$ is left cancellative, *i.e.*, to prove that $t * t_1 =_{LD} t * t_2$ implies $t_1 =_{LD} t_2$ for all terms t, t_1, t_2 in T_1. But, once Property $\mathbf{C}_{LD}$ is known to be true, we know that $t_1 \neq_{LD} t_2$ implies that at least one of $t_1 \sqsubset_{LD} t_2$ or $t_2 \sqsubset_{LD} t_1$ is true, and therefore so is at least one of $t * t_1 \sqsubset_{LD} t * t_2$ or $t * t_2 \sqsubset_{LD} t * t_1$ (by the explicit definition of $\sqsubset_{LD}$ and left self-distributivity). By Proposition 2.3.3, both contradict $t * t_1 =_{LD} t * t_2$. □

***Corollary 2.3.15* (Property $\mathbf{A}_{LD}$).** — *There exists an acyclic left cancellative LD-system.*

Finally, by applying Proposition 2.3.1, we deduce:

***Corollary 2.3.16* (Property $\mathbf{A}_i$).** — *A braid that admits a representative braid word containing at least one letter σ_i but no letter σ_i^{-1} is not trivial.*

In particular, for $i = 1$, we obtain Property **A**.

2.3.5. A characterization of the braid ordering. — We have seen above that the free LD-system of rank 1 is acyclic. Actually, it turns out that this LD-system is even orderable, a result that extends to every free LD-system:

Proposition 2.3.17. — *Every free LD-system is orderable, and, conversely, every orderable LD-system is free.*

Proof (sketch). — In the case of $\mathcal{F}_1$, all needed ingredients are already at hand. Indeed, Property $\mathbf{A}_{LD}$ tells us that the relation $\sqsubset_{LD}$ on T_1 induces an acyclic relation $<$ on $\mathcal{F}_1$; by construction, $<$ is transitive, so it is a partial ordering. Moreover, $x < x * y$ is true by definition, and, using self-distributivity, we easily see that $y < z$ implies $x * y < x * z$. Then Property $\mathbf{C}_{LD}$ tells us that the partial ordering $<$ is actually linear, and we conclude that $(\mathcal{F}_1, <)$ is an ordered LD-system.

Conversely, assume that $(S, <)$ is a monogenic ordered LD-system. By definition, there exists a canonical surjection π of $\mathcal{F}_1$ onto S. By construction, the relation $x < y$ is true in $\mathcal{F}_1$ if and only if we have

$$y = (\dots((x * z_1) * z_2) * \dots) * z_p \tag{2.3.6}$$

for some p and some $z_1, \dots, z_p$ in $\mathcal{F}_1$. Now such a relation implies

$$\pi(y) = (\dots((\pi(x) * \pi(z_1)) * \pi(z_2)) * \dots) * \pi(z_p),$$

hence $\pi(x) < \pi(y)$ in S (as $a < a * b$ is supposed to be true in S). This shows that π is increasing, hence it is an isomorphism.

The extension to the case of LD-systems with more than one generator is easy. First, it is always true that the relation $\sqsubset_{LD}$ on terms induces a partial ordering on every free LD-system $\mathcal{F}$. This order is not linear in general, but there is a unique way to combine it with a linear ordering of the generators in a sort of lexicographical extension so as to order $\mathcal{F}$.

Finally, if $(S, <)$ is an arbitrary ordered LD-system generated by X, one shows as above that the canonical projection of the X-based free LD-system onto S is an isomorphism (using the uniqueness of the ordering). □

Remark 2.3.18. — The previous result applies to the LD-system $(B_{\mathrm{sp}}, *)$: when equipped with the restriction of the braid ordering, the latter becomes an ordered LD-system, and, therefore, a free one: we obtain in this way a concrete realization of the free LD-system of rank 1. Observe that, if we replace Identity (LD) with associativity, *i.e.*, if we consider semigroups instead of LD-systems, then the free system of rank 1 is $(\mathbb{Z}^+, +)$; here also, there exists a (unique) compatible linear ordering satisfying $x < x + y$, namely the standard ordering of the positive integers. Thus, the cases of

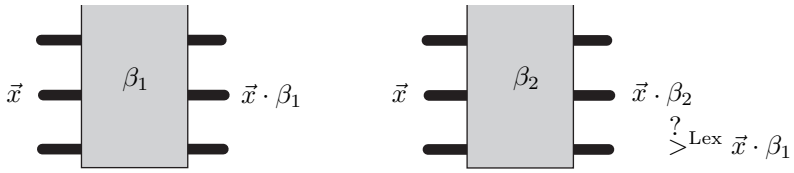


FIGURE 2.4. The braid ordering in terms of colourings

semigroups and LD-systems are not so different—yet the free monogenic LD-system, namely $(B_{\mathrm{sp}}, *)$, is combinatorially much more complicated than the free monogenic semigroup $(\mathbb{Z}^+, +)$. Let us mention that similar realizations for the free LD-systems of rank bigger than 1 can be constructed inside some extension of B_∞ ("charged braids", in the sense of [**41**]).

With the previous results at hand, it is easy to characterize the braid ordering in terms of an action on an ordered LD-system S: the braid β_1 is smaller than the braid β_2 if, when we put the same colours from S on the left ends of the strands of β_1 and β_2, and compare the sequences of colours obtained at the right ends, then the final sequence of β_1 is smaller than that of β_2 with respect to the lexicographical ordering on sequences of colours (Figure 2.4). We thus obtain the second characterization of the braid ordering mentioned in the introduction:

Proposition 2.3.19. — *Assume that $(S, *, <)$ is a left cancellative ordered LD-system. Then, for all braids β_1, β_2 in B_n $(2 \leqslant n \leqslant \infty)$, the relation $\beta_1 < \beta_2$ is true if and only if, for some/any sequence $\vec{x}$ in S^n such that $\vec{x} \cdot \beta_1$ and $\vec{x} \cdot \beta_2$ are defined, we have $\vec{x} \cdot \beta_1 <^{\mathrm{Lex}} \vec{x} \cdot \beta_2$, where $<^{\mathrm{Lex}}$ denotes the lexicographical extension of $<$ to S^n.*

Proof. — As the relation $<$ is a linear ordering on B_∞, it suffices to show that, if $\vec{x} \cdot \beta_1 <^{\mathrm{Lex}} \vec{x} \cdot \beta_2$ is satisfied for at least one sequence $\vec{x}$, then we have $\beta_1 < \beta_2$.

By Proposition 2.3.17, the LD-system $(S, *)$ must be a free LD-system. Assume first that S is monogenic. Then we may assume that S is the set B_{sp} of special braids equipped with the operation of (2.2.1), so, in particular, $\vec{x}$ is a sequence of special braids. By Lemma 2.2.3, the hypothesis that the sequences $\vec{x} \cdot \beta_1$ and $\vec{x} \cdot \beta_2$ are defined implies that the braid $\beta_1^{-1}\beta_2$ is equal to

$$\prod^{sh}(\vec{x} \cdot \beta_1)^{-1} \prod^{sh}(\vec{x} \cdot \beta_2). \tag{2.3.7}$$

Then, as in the proof of Proposition 2.2.7, the hypothesis that $\vec{x} \cdot \beta_1 <^{\mathrm{Lex}} \vec{x} \cdot \beta_2$ is true implies that the braid in (2.3.7) admits a σ-positive representative, and we deduce $\beta_1 < \beta_2$.

In the general case, *i.e.*, when S is not monogenic, we consider the canonical projection p of S to a monogenic LD-system S' that maps all generators of S to the

(unique) generator of S'. Then one can show that the inequality $\vec{x} \cdot \beta_1 <^{\mathrm{Lex}} \vec{x} \cdot \beta_2$ in S must project to a similar inequality $p(\vec{x}) \cdot \beta_1 <^{\mathrm{Lex}} p(\vec{x}) \cdot \beta_2$ in S', from which we deduce $\beta_1 < \beta_2$ as above. □

2.4. Normal forms in free LD-systems

Here we sketch the argument developed by Richard Laver in [**94**] to prove the subword property **S** by using left self-distributivity.

The problem is as follows. We have to prove that, for each braid β and each i, the inequality $\beta < \sigma_i\beta$ is satisfied. To this end, owing to Proposition 2.3.19, we could use a convenient ordered LD-system, and prove $\vec{x} \cdot \beta <^{\mathrm{Lex}} \vec{x} \cdot \sigma_i\beta$ for some sequence $\vec{x}$. We could for instance try to use the ordered LD-systems $(B_\infty, *, <)$ and $(B_{\mathrm{sp}}, *, <)$, which are eligible. The problem is that, if S is any of these ordered LD-systems, the action of B_n on S^n is not order preserving: For instance, we have $(1, \sigma_1) <^{\mathrm{Lex}} (\sigma_1, 1)$, but

$$(1, \sigma_1) \cdot \sigma_1 = (\sigma_2\sigma_1, 1) >^{\mathrm{Lex}} (\sigma_1, 1) \cdot \sigma_1 = (\sigma_1^2\sigma_2^{-1}, \sigma_1),$$

as $\sigma_2\sigma_1 > \sigma_1^2\sigma_2^{-1}$ is true since the braid $\sigma_1^{-1}\sigma_2^{-1}\sigma_1^2\sigma_2^{-1}$ admits the σ_1-negative expression $\sigma_2^2\sigma_1^{-1}\sigma_2^{-2}$. So, when using the ordered LD-system $(B_\infty, *, <)$, we certainly have $\vec{x} <^{\mathrm{Lex}} \vec{x} \cdot \sigma_i$ for every $\vec{x}$, but it is not clear how to deduce $\vec{x} \cdot \beta <^{\mathrm{Lex}} \vec{x} \cdot \sigma_i\beta$ in general.

The solution proposed by Laver consists in working with the free LD-system on countably many generators $\mathcal{F}_\infty$ to have more space, and to use a partial action of B_n on some proper subset $\mathcal{D}$ of (some superset of) $\mathcal{F}_\infty$, one whose elements are sparse enough to obtain an order-preserving action.

2.4.1. LD-monoids. — Actually, the argument requires using a still larger structure, namely a free *LD-monoid* with countably many generators.

Definition 2.4.1. — An *LD-monoid* is defined to be a monoid $(M, \cdot, 1)$ equipped with a second binary operation $*$ so that the following mixed identities are satisfied:

$$x \cdot y = (x * y) \cdot x, \tag{2.4.1}$$

$$(x \cdot y) * z = x * (y * z), \tag{2.4.2}$$

$$x * (y \cdot z) = (x * y) \cdot (x * z), \tag{2.4.3}$$

$$1 * x = x, \quad x * 1 = 1. \tag{2.4.4}$$

Observe that every LD-monoid is an LD-system, as the second operation $*$ must be left self-distributive:

$$x * (y * z) = (x \cdot y) * z = ((x * y) \cdot x) * z = (x * y) * (x * z).$$

LD-monoids occur naturally in the study of LD-systems [**29, 90**]. Many examples of LD-systems are in fact LD-monoids: for instance, if G is a group, G equipped with its group multiplication and with conjugation is an LD-monoid. Moreover, there

exists an easy uniform way for embedding a given LD-system S into an LD-monoid $\widehat{S}$ built from the free monoid generated by S (a construction closely connected to that used in [**58, 99**] to study the set-theoretical Yang-Baxter equation); in particular, the completion of the free LD-system of rank n is a free LD-monoid of rank n. In the sequel, for $1 \leqslant n \leqslant \infty$, we denote by $\mathcal{F}_n$ the free LD-system of rank n, and, therefore, by $\widehat{\mathcal{F}}_n$ the free LD-monoid of rank n. Such structures are eligible for our approach, as we have

Proposition 2.4.2. — *For every n, $1 \leqslant n \leqslant \infty$, the free LD-monoid $\widehat{\mathcal{F}_n}$ is an acyclic left cancellative LD-system. There exists a unique ordering on $\widehat{\mathcal{F}_n}$ such that $x < x * y$ and $x \leqslant x \cdot y$ always hold and we have $x_1 > x_2 > \ldots > x_n$.*

2.4.2. Decreasing division form. — The problem with the action on say $\mathcal{F}_1$ is that we lack space for separating the elements: typically, Lemma 2.2.11 shows that any two elements of $\mathcal{F}_1$ become equal when multiplied on the right by some sufficiently large power of the generator. To avoid such phenomena, which discard the possibility of an order-preserving braid action, we consider a convenient subset $\mathcal{D}$ of $\widehat{\mathcal{F}}_\infty$.

The construction of $\mathcal{D}$ is rather delicate, and it appeals to the normal form results established in [**91, 92**]. The elements of $\widehat{\mathcal{F}}_\infty$ can be represented as equivalence classes of terms constructed using an infinite series of variables x_1, x_2, ..., and two binary operators $*$ and $\cdot$, with respect to the congruence $=_{LDM}$ corresponding to Identities (2.4.1), ..., (2.4.4) together with the usual identities of a monoid, *i.e.*, associativity and neutral element. Defining a normal form means selecting in each equivalence class of $=_{LDM}$ a distinguished term.

Definition 2.4.3. — We say that the term t is in *division normal form*, or, for short, is *normal*, if it has the form

$$((\cdots((x_i * t_1) * t_2)\cdots) * t_{n-1}) \bullet t_n, \tag{2.4.5}$$

where $\bullet$ is either $*$ or $\cdot$, t_1, ..., t_n are normal, and we have $t_{k+2} \sqsubseteq_{LD} (\cdots((x_i * t_1) * t_2)\cdots) * t_k$ for $k \leqslant n-2$, and, if $\bullet$ is $\cdot$, $t_n \sqsubset_{LD} (\cdots((x_i * t_1) * t_2)\cdots) * t_{n-2}$.

Normal terms make distinguished representatives for all $=_{LDM}$-classes:

Proposition 2.4.4. — *Every element of $\widehat{\mathcal{F}}_\infty$ is represented by a unique normal term; moreover, we have $x < y$ with respect to the ordering of Proposition 2.4.2 if and only if the normal term representing x precedes the normal term representing y in the lexicographical extension of $x_1 > x_2 > \cdots$.*

Let us now introduce a subset of $\widehat{\mathcal{F}}_\infty$.

Definition 2.4.5. — We say that a normal term t is *decreasing* if, for each subterm $t_1 * t_2$ or $t_1 \cdot t_2$ of t, we have $t_1 > t_2$ with respect to the order mentioned in Proposition 2.4.4. We denote by $\mathcal{D}$ the subset of $\widehat{\mathcal{F}}_\infty$ consisting of those elements whose normal form is decreasing.

It is easy to see that $\mathcal{D}$ is a proper subset of $\widehat{\mathcal{F}}_\infty$: for instance, the element $x_2 * x_1$ does not belong to $\mathcal{D}$, as its normal form is the non-decreasing normal term $x_2 * x_1$.

The main result is then the existence of an action of B_∞ on $\mathcal{D}$—*not* on $\mathcal{D}^{\mathbb{N}}$: the action is not that of Section 2.1, and it should rather be thought of as an analog of the action of B_∞ on a free group considered in Chapter 5.

Proposition 2.4.6. — (i) *The formulas*

$$x_{i-1} \cdot \sigma_i = x_{i-1} * x_i, \quad x_i \cdot \sigma_i = x_{i-1}, \quad x_j \cdot \sigma_i = x_j \text{ for } j \neq i-1, i \tag{2.4.6}$$

induce a well-defined and faithful partial action of B_∞ on $\mathcal{D}$ in the following sense: for any two distinct braids β_1, β_2 in B_∞, there exists at least one element x of $\mathcal{D}$ such that $x \cdot \beta_1$ and $x \cdot \beta_2$ are defined and distinct.

(ii) *Moreover, for all x, y in $\mathcal{D}$, we have $x \leqslant x \cdot \sigma_i$, and $x < y$ is equivalent to $x \cdot \sigma_i < y \cdot \sigma_i$, whenever these expressions are defined.*

We skip the proof, which is an intricate induction on normal terms (the argument for the existence of at least one element x in $\mathcal{D}$ such that $x \cdot \beta_1$ and $x \cdot \beta_2$ are defined is the same as the one of Section 1 for the partial action on the powers of an LD-system). It is then easy to conclude.

***Proposition 2.4.7* (Property S)**. — *Every braid of the form $\beta^{-1}\sigma_i\beta$ admits a σ-positive representative braid word.*

Proof (sketch). — By definition of the action of B_∞ on $\mathcal{D}$, if β_1, β_2 belong to the image of sh^i, *i.e.*, can be represented without using $\sigma_1, \ldots, \sigma_i$, then $x \cdot \beta_1 < x \cdot \sigma_i\beta_2$ is true (whenever the terms are defined), and one can deduce that any inequality $\beta_1 < \beta_2$ in B_∞ implies $x \cdot \beta_1 < x \cdot \beta_2$ in $\mathcal{D}$ when the terms are defined.

Now, let β be an arbitrary braid. By Proposition 2.4.6(i), there exists x in $\mathcal{D}$ such that $x \cdot \beta$ and $x \cdot \sigma_i\beta$ are defined and distinct. By Proposition 2.4.6(ii), we obtain $x \leqslant x \cdot \sigma_i$, and then, inductively, $x \cdot \beta \leqslant x \cdot \sigma_i\beta$, hence $x \cdot \beta < x \cdot \sigma_i\beta$ as we assumed $x \cdot \beta \neq x \cdot \sigma_i\beta$. Two cases are possible *a priori*, namely $\beta < \sigma_i\beta$ and $\beta > \sigma_i\beta$. By the remark above, the latter would imply $x \cdot \beta > x \cdot \sigma_i\beta$, so it is impossible. Hence we have $\beta < \sigma_i\beta$, *i.e.*, the braid $\beta^{-1}\sigma_i\beta$ admits a σ-positive representative. □

2.5. Appendix: Iterations of elementary embeddings in set theory

It might be of interest to mention the connection between the results described in this chapter and some questions in set theory, centered around Property $\mathbf{A}_{LD}$.

Contrary to algebraic systems containing an operation that is self-distributive both on the left and on the right, which had been studied in particular in the 1960's and 70's by Belousov in Kichinev, and Jaroslav Ježek, Tomáš Kepka, Petr Nemeč and *al.* in Prague, LD-systems, and, in particular, free LD-systems, had received rather little attention until the beginning of the 1980's. Then set theory provided a new, puzzling example involving the iterations of an elementary embedding of a self-similar rank. An elementary embedding is a sort of strong homomorphism—see for instance [**81**]—and a rank is a set with the special property that every mapping of R to R can be seen as an element of R. It follows that, if i, j are two mappings of a rank into itself, then i may be *applied* to j (as j is an element of R). Playing with this situation, it is easy to see that this application operation, denoted $*$ in the sequel, satisfies Identity (LD), *i.e.*, the set I of all elementary embeddings of a rank R into itself equipped with this operation $*$ is an LD-system. Early results, in particular in [**90, 30**], showed that the LD-system $(I, *)$ has complicated and presumably deep properties, motivating further investigations. In 1989, the following results were proved independently:

Proposition 2.5.1. — [**31**] *Assume that Property* $\mathbf{A}_{LD}$ *is true, i.e., there exists an acyclic LD-system. Then the word problem of Identity* (LD) *is decidable, i.e., there exists an effective algorithm to decide for arbitrary given terms* t, t' *whether* $t =_{LD} t'$ *is true or not.*

(The result is based on Property $\mathbf{C}_{LD}$: starting with two terms t, t', we can enumerate all terms LD-equivalent to t and t'; after finitely many steps, we shall obtain a proof of $t =_{LD} t'$, $t \sqsubset_{LD} t'$, or $t' \sqsubset_{LD} t$; if Property $\mathbf{A}_{LD}$ is true, we can conclude in the last two cases that $t =_{LD} t'$ is false.)

Proposition 2.5.2. — [**91**] *The LD-system* $(I, *)$ *is acyclic.*

The conjunction of the above two results seems to give a proof of the decidability of the word problem for Identity (LD), but *it does not.* Indeed, the existence of the system $(I, *)$ is an unprovable statement, one whose logical status is to remain open: the construction of $(I, *)$ requires starting with a very large rank, called *self-similar*, and, like the existence of an inaccessible cardinal or of a measurable cardinal, the existence of a self-similar rank cannot be deduced from the usual axioms of set theory—and it cannot even be proved to be non-contradictory. Thus, the only consequence one can deduce from Propositions 2.5.1 and 2.5.2 was:

Corollary 2.5.3. — *If there exists a self-similar rank, then the word problem for Identity* (LD) *is decidable.*

This was a quite paradoxical statement, as the existence of a connection between a syntactic, finitistic question such as the word problem of an algebraic identity and huge objects of set theory appears as unlikely—though not *a priori* impossible. Between 1989 and 1992, two possible conclusions were possible: either Property $\mathbf{A}_{LD}$

is inevitably connected with some strong logical axiom—as are certain combinatorial properties of the integers studied by Harvey Friedman in his reverse mathematics program [**63, 64**]—or there exists a new, direct proof of Property $\mathbf{A}_{LD}$ that does not require using any weird logical assumption. The latter happened: by studying free LD-systems along the lines described in Sections 2.2 and 2.3—which entirely take place in an ordinary mathematical framework—one could build the desired proof of Property $\mathbf{A}_{LD}$, with the additional benefit of introducing braids in the picture and deducing unexpected braid orderability results.

It is not clear that proving Property $\mathbf{A}_{LD}$ would have been considered an interesting challenge if set theory had not given some strong hint that this property should be true. So all further developments about braid orderings and induced results can be seen as *applications* of set theory. However, it should be stressed that these are applications of a particular type, as they precisely appeared in the process of removing set theory from some earlier results. We can compare the role of set theory here with the role of physics when it gives evidence for some formulas that remain then to be proved rigorously: in some sense, adding an unprovable logical statement is not so far from, say, liberally using diverging series or infinite integrals.

Let us mention that Laver investigated in [**93**] some finite quotients of the set-theoretical LD-system $(I, *)$. Under the hypothesis that a self-similar rank exists, he deduces nontrivial combinatorial properties of these finite LD-systems. A puzzling point is that, contrary to the case of Property $\mathbf{A}_{LD}$, no alternative proof avoiding set theoretical hypotheses has been found so far, nor has it either been proved that these hypotheses are inevitable, despite strong attempts by Randall Dougherty [**49**], Thomas Jech [**50**], and Aleš Drápal [**51**]—also see Chapter XIII of [**40**].

CHAPTER 3

HANDLE REDUCTION

Handle reduction is a combinatorial method (with a natural geometrical content) that gives a proof of Property **C**, provided Property **A** is known. It was developed in [**37**]. The main interest of the method is that it does not only give a proof that every nontrivial braid word admits a representative that is σ-positive or σ-negative, but it also gives an algorithm for finding such a representative, *i.e.*, for transforming an arbitrary braid word into an equivalent σ-positive or σ-negative braid word, and, therefore, for comparing braids with respect to the order $<$. Moreover, this algorithm turns out to be extremely efficient in pratice, yet no theoretical confirmation of that efficiency has been found so far.

The techniques in this chapter mainly belong to combinatorial group theory, and, in particular, the Cayley graph of the braid group plays a central role.

3.1. Handles

The idea of handle reduction is simple. Assume that w is a nonempty braid word that is neither σ-positive nor σ-negative: this means that, if i is the smallest index such that $\sigma_i^{\pm 1}$ appears in w, then both σ_i and σ_i^{-1} appear in w. So, necessarily, w contains some subword of the form $\sigma_i^e \mathrm{sh}^i(u)\sigma_i^{-e}$ with $e = \pm 1$, where we recall sh denotes the word homomorphism that maps every letter $\sigma_i^{\pm 1}$ to $\sigma_{i+1}^{\pm 1}$—as well as the induced endomorphism of B_∞.

Definition 3.1.1. — A braid word of the form $\sigma_i^e \mathrm{sh}^i(u)\sigma_i^{-e}$ with $e = \pm 1$ is called a *σ_i-handle*.

Thus, every braid word that is neither σ-positive nor σ-negative must contain a σ_i-handle. The name refers to the handle formed by the $(i+1)$st strand in the associated braid diagram, as shown in Figure 3.1.

Now, we can get rid of a handle by pushing the strand involved in the handle, so that it skirts above the next crossings (in the case of a handle $\sigma_i \dots \sigma_i^{-1}$) or below them

FIGURE 3.1. A σ_1-handle

(in the case of a handle $\sigma_i^{-1} \dots \sigma_i$), according to the scheme of Figure 3.2. We call this transformation *reduction* of the handle. We can then iterate handle reduction until no handle is left: if the process converges, then, by construction, the final word contains no handle, which implies that it is either σ-positive, or σ-negative, or empty. This naive approach does not work readily: when applied to the word $w = \sigma_1\sigma_2\sigma_3\sigma_2^{-1}\sigma_1^{-1}$, it leads in one step to the word $w' = \sigma_2^{-1}w\sigma_2$: the initial handle is still there, and iterating the process leads to nothing but longer and longer words. Now, we see that the handle in w' is not the original handle of w, but it comes from the σ_2-handle $\sigma_2\sigma_3\sigma_2^{-1}$ of w. If we reduce the latter handle into $\sigma_3^{-1}\sigma_2\sigma_3$ before reducing the main handle of w, *i.e.*, if we first go from w to $w'' = \sigma_1\sigma_3^{-1}\sigma_2\sigma_3\sigma_1^{-1}$, then applying handle reduction yields $\sigma_3^{-1}\sigma_2^{-1}\sigma_1\sigma_2\sigma_3$, a σ-positive word equivalent to w. We shall see in the sequel that the previous obstruction, namely the existence of nested handles, is the only possible problem: provided we first reduce all nested handles, handle reduction always comes to an end in a finite number of steps.

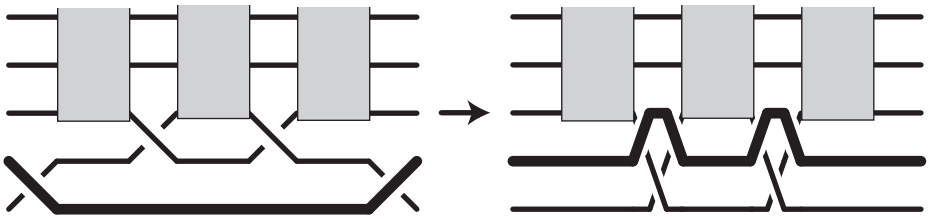


FIGURE 3.2. Reduction of a handle

Definition 3.1.2. — A handle $\sigma_i^e v \sigma_i^{-e}$ is said to be *permitted* if the word v includes no σ_{i+1}-handle. We say that the braid word w' is obtained from the braid word w *by a one-step handle reduction* if some subword of w is a permitted σ_i-handle, say $\sigma_i^e v \sigma_i^{-e}$, and w' is obtained from w by applying in the latter handle the alphabetical homomorphism

$$\sigma_i^{\pm 1} \mapsto \varepsilon, \quad \sigma_{i+1}^{\pm 1} \mapsto \sigma_{i+1}^{-e}\sigma_i^{\pm 1}\sigma_{i+1}^{e}, \quad \sigma_k^{\pm 1} \mapsto \sigma_k^{\pm 1} \text{ for } k \geqslant i+2.$$

We say that w' is obtained from w by m steps of handle reduction if there exists a sequence of length $m+1$ from w to w' such that each term is obtained from the previous one by a one-step handle reduction.

The general form of a σ_i-handle is

$$\sigma_i^e\, v_0\, \sigma_{i+1}^{d_1}\, v_1\, \sigma_{i+1}^{d_2} \dots \sigma_{i+1}^{d_k}\, v_k\, \sigma_i^{-e}$$

with $d_j = \pm 1$ and $v_j \in \mathrm{sh}^{i+1}(B_\infty)$. Saying that this handle is permitted amounts to saying that all exponents d_j have a common value d. Then, reducing the handle means replacing it with

$$v_0\, \sigma_{i+1}^{-e}\sigma_i^d\sigma_{i+1}^e\, v_1\, \sigma_{i+1}^{-e}\sigma_i^d\sigma_{i+1}^e \dots \sigma_{i+1}^{-e}\sigma_i^d\sigma_{i+1}^e\, v_k :$$

we remove the initial and final $\sigma_i^{\pm 1}$, and replace each σ_{i+1}^d with $\sigma_{i+1}^{-e}\sigma_i^d\sigma_{i+1}^e$. The following result is clear:

Lemma 3.1.3. — (i) *Handle reduction transforms a word into an equivalent word;* (ii) *If a nonempty braid word w is terminal w.r.t. handle reduction, i.e., if w contains no handle, then w is σ-positive or σ-negative.*

Remark 3.1.4. — Observe that handle reduction generalizes free reduction, defined to be the deletion of some subword $\sigma_i\sigma_i^{-1}$ or $\sigma_i^{-1}\sigma_i$. Indeed such subwords are particular σ_i-handles, and reducing them amounts to deleting them. Free reduction solves the word problem for free groups, but it does not solve it for any group that is not free; handle reduction is an extension of free reduction that we shall see is relevant for braid groups.

3.2. Convergence of handle reduction

This section is devoted to proving that handle reduction always converges, which in particular will give us a new proof of Property **C**. To state our complexity bound precisely, we introduce the following parameter.

Definition 3.2.1. — We say that a nonempty braid word w has *width* n if the difference between the smallest and the largest indices i such that σ_i or σ_i^{-1} occurs in w is $n-2$.

The width of w is the size of the smallest interval containing the indices of all strands really braided in w. So, every n-strand braid word has width at most n, but the inequality may be strict: for instance, the 8-strand braid word $\sigma_3\sigma_7^{-1}$ has width 6.

Proposition 3.2.2. — *Let w be a braid word of length ℓ and width n. Then every sequence of handle reductions from w converges in at most $2^{n^4\ell}$ steps.*

Corollary 3.2.3 **(Property C)**. — *Every n-strand braid word is equivalent to some n-strand braid word that is σ-positive, σ-negative, or empty.*

Indeed, Propostion 3.2.2 tells us that every n-strand braid word w is equivalent to some braid word w' that contains no handle: by definition, such a braid word w' is either empty, or σ-positive, or σ-negative.

Observe that handle reduction gives a solution both for the word problem of B_∞ and for the decision problem of the linear ordering on B_∞.

Proposition 3.2.4. — *Let w be a braid word, and β be the braid represented by w.*
(i) *We have $\beta = 1$ in B_∞ if and only if w is reducible to the empty word;*
(ii) *We have $\beta > 1$ in B_∞ if w is reducible to some σ-positive braid word.*

Proof. — The only problem is that a given braid word may contain several handles, and reduction need not be confluent in general, *i.e.*, various sequences of reductions from a given word may lead to distinct final words. However, Property **A** tells us that the cases in Proposition 3.2.4 do not depend of the considered sequence of reductions: all words obtained by reduction from w are pairwise equivalent, and Property **A** asserts that no equivalence class may contain the empty word and a σ-positive word at the same time. □

We deduce the third equivalent definition of the braid ordering mentioned in Introduction:

Corollary 3.2.5. — *Let β_1, β_2 be any braids. Then $\beta_1 < \beta_2$ is true if and only if any sequence of handle reductions from any expression of $\beta_1^{-1}\beta_2$ ends up with a σ-positive word.*

3.2.1. The absolute value of a braid word. — Handle reduction may increase the length of the braid word it is applied to. Our first task for proving convergence of handle reduction will be to show that all words obtained from a word w using handle reduction remain traced in some finite region of the Cayley graph of B_∞ depending on w only. To this end, we introduce the notion of the absolute value of a braid word using the techniques of Section 2.1, and we connect the operation of handle reduction with the operations of left and right reversing. We recall that, for every braid word w, there exist a unique pair of positive braid words $D(w), N(w)$ such that w is left reversible to $D(w)^{-1}N(w)$, this meaning that we iteratively substitute positive–negative subwords of the form $\sigma_i\sigma_j^{-1}$ with equivalent negative–positive words $f(\sigma_j, \sigma_i)^{-1}f(\sigma_i, \sigma_j)$ specified by the braid relations; symmetrically, w is right reversible to $\widetilde{N}(w)\widetilde{D}(w)^{-1}$, where we now substitute negative–positive factors with equivalent positive–negative words. The equivalences

$$w \equiv D(w)^{-1}N(w) \equiv \widetilde{N}(w)\widetilde{D}(w)^{-1}$$

then imply $D(w)\,\widetilde{N}(w) \equiv N(w)\,\widetilde{D}(w)$. We recall that $\equiv$ denotes braid word equivalence: $w \equiv w'$ is true if w and w' are two braid word representatives of the sme braid.

Definition 3.2.6. — For w a braid word, the *absolute value* $|w|$ of w is defined to be the positive braid represented by $D(w)\widetilde{N}(w)$.

Example 3.2.7. — The braid word $w = \sigma_1^{-1}\sigma_2\sigma_2^{-1}\sigma_3$ is left reversible to $\sigma_1^{-1}\sigma_3$, so we have $N(w) = \sigma_3$ and $D(w) = \sigma_1$. On the other hand, w is right reversible to $\sigma_2\sigma_1\sigma_3\sigma_2\sigma_1\sigma_3^{-1}\sigma_2^{-1}\sigma_1^{-1}\sigma_3^{-1}\sigma_2^{-1}$, *i.e.*, we have $\widetilde{N}(w) = \sigma_2\sigma_1\sigma_3\sigma_2\sigma_1$ and $\widetilde{D}(w) = \sigma_2\sigma_3\sigma_1\sigma_2\sigma_3$. So the absolute value of w is the positive braid $\sigma_1\sigma_2\sigma_1\sigma_3\sigma_2\sigma_1$, which is also $\sigma_3\sigma_2\sigma_3\sigma_1\sigma_2\sigma_3$.

In the case of B_2, *i.e.*, of $\mathbb{Z}$, the absolute value of σ_1^k is $\sigma_1^{|k|}$. Observe that, by construction, $|w| = |w^{-1}|$ is true for every w.

The general notion of the Cayley graph of a group (with respect to specified generators) is well-known. Here we consider finite fragments of such graphs.

Definition 3.2.8. — Assume that β is a positive braid. The *Cayley graph of* β is the finite labelled oriented graph $\Gamma(\beta)$ defined as follows: the vertices are the left divisors of β, and there exists an edge labelled σ_i from the vertex β_1 to the vertex β_2 if $\beta_2 = \beta_1\sigma_i$ holds.

When we are given a graph Γ whose (oriented) edges are labelled using letters from some alphabet A, we have the natural notion of a word *traced* in Γ: for w a word on the alphabet A, we say that w is traced in Γ from the vertex β_0 if there exists in Γ a path starting at β_0 labelled w, *i.e.*, w is the word obtained by concatenating the labels of the edges in that path, with exponents ± 1 according as the edge orientation agrees or disagrees with that of the path. This corresponds to the following formal definition:

Definition 3.2.9. — Assume that Γ is a fragment of the Cayley graph of B_∞, and that β_0 is a vertex of Γ. For w a braid word, we say that w is *traced in* Γ *from* β_0 if either w is the empty word, or w is $\sigma_i v$ (*resp.* w is $\sigma_i^{-1}v$), there is an σ_i-labelled edge from β_0 to another vertex β_0' (*resp.* from another vertex β_0' to β_0) in Γ and v is traced from β_0' in Γ.

Infinitely many words are traced in every graph Γ containing at least one edge: if σ_i is the label of an edge from β, the word $(\sigma_i\sigma_i^{-1})^k$ is traced from β for every integer k. For our current purpose, the point is that, for every positive braid β, the set of all words traced in the Cayley graph of β enjoys good closure properties.

Lemma 3.2.10. — *Assume that β is a positive braid. Then the set of all words traced in $\Gamma(\beta)$ from a given point is closed under left and right reversing.*

Proof. — Let us consider left reversing. So we assume that some word $v\,\sigma_i\,\sigma_j^{-1}\,v'$ is traced from β_0 in $\Gamma(\beta)$, and we have to show that the word

$$v\ f(\sigma_j,\sigma_i)^{-1}\ f(\sigma_i,\sigma_j)\ v'$$

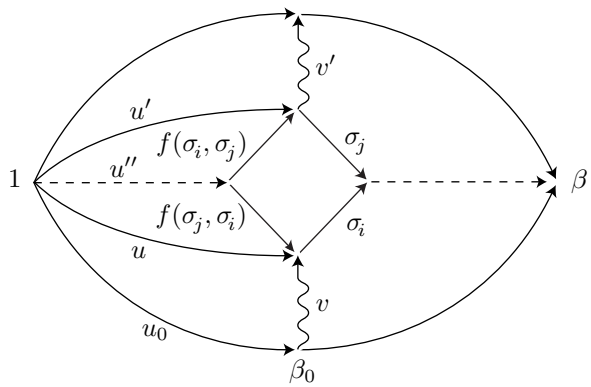


FIGURE 3.3. Closure of words traced under left reversing

is also traced from β_0 in $\Gamma(\beta)$. Let u_0 be a positive word representing β_0. The hypothesis means that there exist positive braid words u, u' such that both $u\sigma_i$ and $u'\sigma_j$ are traced from 1 in $\Gamma(\beta)$, the equivalence $u\sigma_i \equiv u'\sigma_j$ is satisfied, and, moreover, we have $u \equiv u_0 v$ (Figure 3.3). According to Garside's result that right least common multiples exist in the monoid B_∞^+ [**66**], $u\sigma_i \equiv u'\sigma_j$ implies that there exists a positive braid word u'' satisfying $u \equiv u'' f(\sigma_j, \sigma_i)$ and $u' \equiv u'' f(\sigma_i, \sigma_j)$. By definition of the Cayley graph of β, the words $u'' f(\sigma_j, \sigma_i)\sigma_i$ and $u'' f(\sigma_i, \sigma_j)\sigma_j$ are traced in $\Gamma(\beta)$, since they are both equivalent to $u\sigma_i$. This shows that the edges $f(\sigma_j, \sigma_i)^{-1}$ and $f(\sigma_i, \sigma_j)$ needed to complete the path labelled $v f(\sigma_j, \sigma_i)^{-1} f(\sigma_i, \sigma_j) v'$ from β_0 are in $\Gamma(\beta)$, as was expected. The argument is symmetric for right reversing. □

Definition 3.2.11. — We say that two (not necessarily positive) braid words w, w' are *positively* (*resp. negatively*) equivalent if one can transform w into w' using the positive braid relations $\sigma_1\sigma_2\sigma_1 = \sigma_2\sigma_1\sigma_2$, etc. (*resp.* the inversed braid relations $\sigma_1^{-1}\sigma_2^{-1}\sigma_1^{-1} = \sigma_2^{-1}\sigma_1^{-1}\sigma_2^{-1}$ etc.) but no relation $\sigma_i\sigma_i^{-1} = 1$ or $\sigma_i^{-1}\sigma_i = 1$.

Lemma 3.2.12. — *Assume that β is a positive braid. Then the set of all words traced in $\Gamma(\beta)$ from a given point is closed under positive and negative equivalence.*

Proof. — Assume, for instance, that $v\sigma_i\sigma_{i+1}\sigma_i v'$ is traced in $\Gamma(\beta)$ from β_0. Let u_0 be a positive word representing β_0. Now v is not necessarily a positive word, but, by definition, there exists a positive word u such that $u\sigma_i\sigma_{i+1}\sigma_i$ is traced in $\Gamma(\beta)$ from 1 and $u_0 v \equiv u$ holds. Now $u\sigma_{i+1}\sigma_i\sigma_{i+1}$ is a positive word equivalent to $u\sigma_i\sigma_{i+1}\sigma_i$, so it is traced from 1 in $\Gamma(\beta)$, and, therefore, $v\sigma_{i+1}\sigma_i\sigma_{i+1}v'$ is still traced in $\Gamma(\beta)$ from β_0. The case of negative equivalence is similar and corresponds to traversing the edges with reversed orientation. □

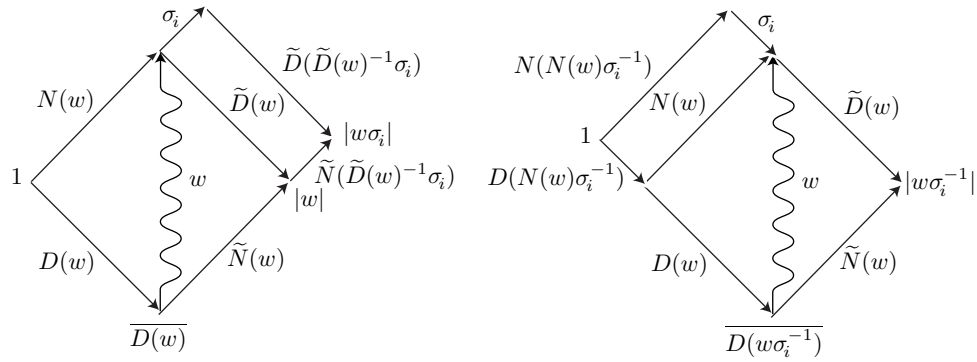


FIGURE 3.4. The word w is traced in the Cayley graph of $|w|$.

If w is a braid word, we use $\overline{w}$ for the braid represented by w. The following result gives a sort of upper bound for the words that can be deduced from a given braid word using reversing and signed equivalence, *i.e.*, essentially, when introducing new patterns $\sigma_i\sigma_i^{-1}$ or $\sigma_i^{-1}\sigma_i$ is forbiden.

Proposition 3.2.13. — *Assume that w is an arbitrary braid word. Then every word that can be obtained from w using left reversing, right reversing, positive equivalence, and negative equivalence is traced from $\overline{D(w)}$ in $\Gamma(|w|)$.*

Proof. — Owing to Lemmas 3.2.10 and 3.2.12, it only remains to show that w itself is traced from $\overline{D(w)}$ in $\Gamma(|w|)$. We use induction on the length of w. The result is straightforward for an empty word. So, let us assume that the result is proved for w, and consider the words $w\sigma_i$ and $w\sigma_i^{-1}$ successively.

By construction, we have (Figure 3.4 left)

$$D(w\sigma_i) = D(w), \quad N(w\sigma_i) = N(w)\sigma_i, \tag{3.2.1}$$

$$\widetilde{N}(w\sigma_i) = \widetilde{N}(w)\widetilde{N}(\widetilde{D}(w)^{-1}\sigma_i), \quad \widetilde{D}(w\sigma_i) = \widetilde{D}(\widetilde{D}(w)^{-1}\sigma_i), \tag{3.2.2}$$

hence $|w\sigma_i| = |w|\widetilde{N}(\widetilde{D}(w)^{-1}\sigma_i)$. It follows that $\Gamma(|w|)$ is, in the obvious sense, an initial subgraph of $\Gamma(|w\sigma_i|)$. So the induction hypothesis that w is traced from $\overline{D(w)}$ in $\Gamma(|w|)$ implies that w is also traced from $\overline{D(w)}$ in $\Gamma(|w\sigma_i|)$. So, in order to show that $w\sigma_i$ is traced in $\Gamma(|w\sigma_i|)$, it only remains to check that the final σ_i-edge is traced in the latter graph from the endpoint of the w-path from $\overline{D(w)}$, which is $\overline{D(w)w}$, *i.e.*, $\overline{N(w)}$. Now, by (3.2.1) and (3.2.2), $|w\sigma_i|$ is represented by $N(w)\sigma_i\widetilde{D}(w\sigma_i)$, which gives the expected result for $w\sigma_i$.

The argument for $w\sigma_i^{-1}$ is symmetric, and it is illustrated on Figure 3.4 (right). $\square$

We have already observed that, if the braid word w' is obtained from w using handle reduction, then w' is equivalent to w. This obvious fact can be refined into the following result.

Lemma 3.2.14. — *Assume that w' is obtained from w using handle reduction. Then one can transform w into w' using right reversing, left reversing, positive equivalence, and negative equivalence.*

Proof. — The point is to show that, for $v_0, \dots, v_k$ in $\mathrm{sh}^{i+1}(B_\infty)$, we can go from

$$\sigma_i^e\, v_0\, \sigma_{i+1}^d\, v_1\, \sigma_{i+1}^d\, \dots\, \sigma_{i+1}^d\, v_k\, \sigma_i^{-e} \tag{3.2.3}$$

to

$$v_0\, \sigma_{i+1}^{-e}\sigma_i^d\sigma_{i+1}^e\, v_1\, \sigma_{i+1}^{-e}\sigma_i^d\sigma_{i+1}^e\, \dots\, \sigma_{i+1}^{-e}\sigma_i^d\sigma_{i+1}^e\, v_k : \tag{3.2.4}$$

using the transformations mentioned in the statement. Assume for instance $e = +1$ and $d = -1$. Then reduction can be done by moving the initial σ_i to the right. First transforming $\sigma_i v_0$ into $v_0\sigma_i$ can be made by a sequence of left reversings (in the case of negative letters) and positive equivalences (in the case of positive letters). Then we find the pattern $\sigma_i\sigma_{i+1}^{-1}$, which becomes $\sigma_{i+1}^{-1}\sigma_i^{-1}\sigma_{i+1}\sigma_i$ by a left reversing. So, at this point, we have transformed the initial word into

$$v_0\ \sigma_{i+1}^{-1}\sigma_i^{-1}\sigma_{i+1}\sigma_i\ v_1\ \sigma_{i+1}^{-1}\ v_2\ \dots\ v_{k-1}\ \sigma_{i+1}^{-1}\ v_k\ \sigma_i^{-1}. \tag{3.2.5}$$

After k such sequences of reductions, and a last left reversing to delete the final pattern $\sigma_i\sigma_i^{-1}$, we reach the form (3.2.4), as we wished. The argument is similar in the case $e = -1$, $d = 1$, with negative equivalences instead of positive equivalences, and right reversing instead of left reversing. In the case when the exponents e and d have the same sign, we use a similar procedure to move the final generator σ_i^{-e} to the left. □

By applying Proposition 3.2.13, we deduce the following upper complexity bound for those words obtained using handle reduction:

Proposition 3.2.15. — *Assume that the braid word w' is obtained from w using handle reduction. Then w' is traced in the Cayley graph of $|w|$ from $\overline{D(w)}$.*

3.2.2. The height of a braid word. — The previous boundedness result is not sufficient for proving that handle reduction always converges. In particular, it does not discard the possibility that loops occur. To go further, we need a new parameter, the height.

Definition 3.2.16. — Assume that w is a braid word. The *height* of w is defined to be the maximal number, over all i, of letters σ_i occurring in a word containing no σ_i^{-1} and traced in the Cayley graph of $|w|$.

Lemma 3.2.17. — *Assume Property* **A**, *and let w be a braid word of length ℓ and width n. Then the height of w is bounded above by $(n-1)^{\ell n(n-1)/2}$.*

Proof. — Assume that u is a braid word traced in the Cayley graph $\Gamma(|w|)$ of $|w|$ in which the letter σ_i^{-1} does not occur. We claim that the various letters σ_i that occur in u correspond to different edges in $\Gamma(|w|)$. Indeed, saying that the same edge is met twice would mean some closed path in this graph contains at least one σ_i-labelled edge and no σ_i^{-1}-labelled edge. In other words, some braid word containing at least one letter σ_i and no letter σ_i^{-1} would represent the unit braid 1. By Property $\mathbf{A}_i$ (Proposition 2.3.1)—which we have seen can be deduced from Property **A** directly—this is impossible. Thus the number of letters σ_i in u is bounded above by the total number of edges in $\Gamma(|w|)$.

We can evaluate the latter as follows. Up to a translation by a power of the shift mapping, we may assume that w is a word over $\sigma_1^{\pm 1}, \ldots, \sigma_{n-1}^{\pm 1}$. Let S_n denote the set of all positive n-strand braid words that represent divisors of the element Δ_n defined in (1.1.1). Garside proved that S_n is closed under left and right least common multiple in the monoid B_n^+ [**66**], and it follows that S_n is closed under left reversing, in the sense that, if u, v belong to S_n, then reversing uv^{-1} on the left ends up with a word $v'^{-1}u'$ such that u' and v' still belong to S_n [**38**]. Similarly, S_n is closed under right reversing. Then, an easy induction shows that, if there are q negative letters in w, then the word $D(w)$ can be decomposed into a product of q words in S_n. Similarly, if there are p positive letters in w, then the word $\widetilde{N}(w)$ can be decomposed into a product of p words in S_n. Hence the braid $|w|$ is the product of at most $p+q=\ell$ divisors of Δ_n. As the length of a divisor of Δ_n is bounded above by $n(n-1)/2$, the length of $|w|$ is at most $\ell n(n-1)/2$. Now, at most $n-1$ edges start from a given vertex in the Cayley graph of a word over $\sigma_1^{\pm 1}, \ldots, \sigma_{n-1}^{\pm 1}$, hence, in the Cayley graph of a word of length L, there are no more than $(n-1)^L$ edges. □

Let us now consider handle reduction. Assume that $w_0 = w, w_1, \ldots$ is a sequence of handle reductions from w. The first point is that the number of σ_1-handles in w_i is not larger than the number of σ_1-handles in w, as reducing one σ_1-handle lets at most one new σ_1-handle appear. We deduce a well-defined notion of inheriting between σ_1-handles such that each σ_1-handle in the initial word w possesses at most one heir in each word w_k. We shall assume that the number of σ_1-handles in every word w_k is the same as in w_0. If it is not the case, *i.e.*, if some σ_1-handle vanishes without heir, say from w_k to w_{k+1}, we cut the sequence at w_k and restart from w_{k+1}. In this way, the heir of the pth σ_1-handle of w_0 (when enumerated from the left) is the pth σ_1-handle of w_k. Let us define the pth *critical prefix* $\pi_p(w_k)$ of w_k to be the braid represented by the prefix of w_k ending at the first letter of the pth σ_1-handle. There are two cases according to whether the first letter of the handle is σ_1 or σ_1^{-1}. We shall assume here that this letter is σ_1, and briefly mention at the end of the argument the changes for the σ_1^{-1}-case. The key point is the following observation: for every p, we

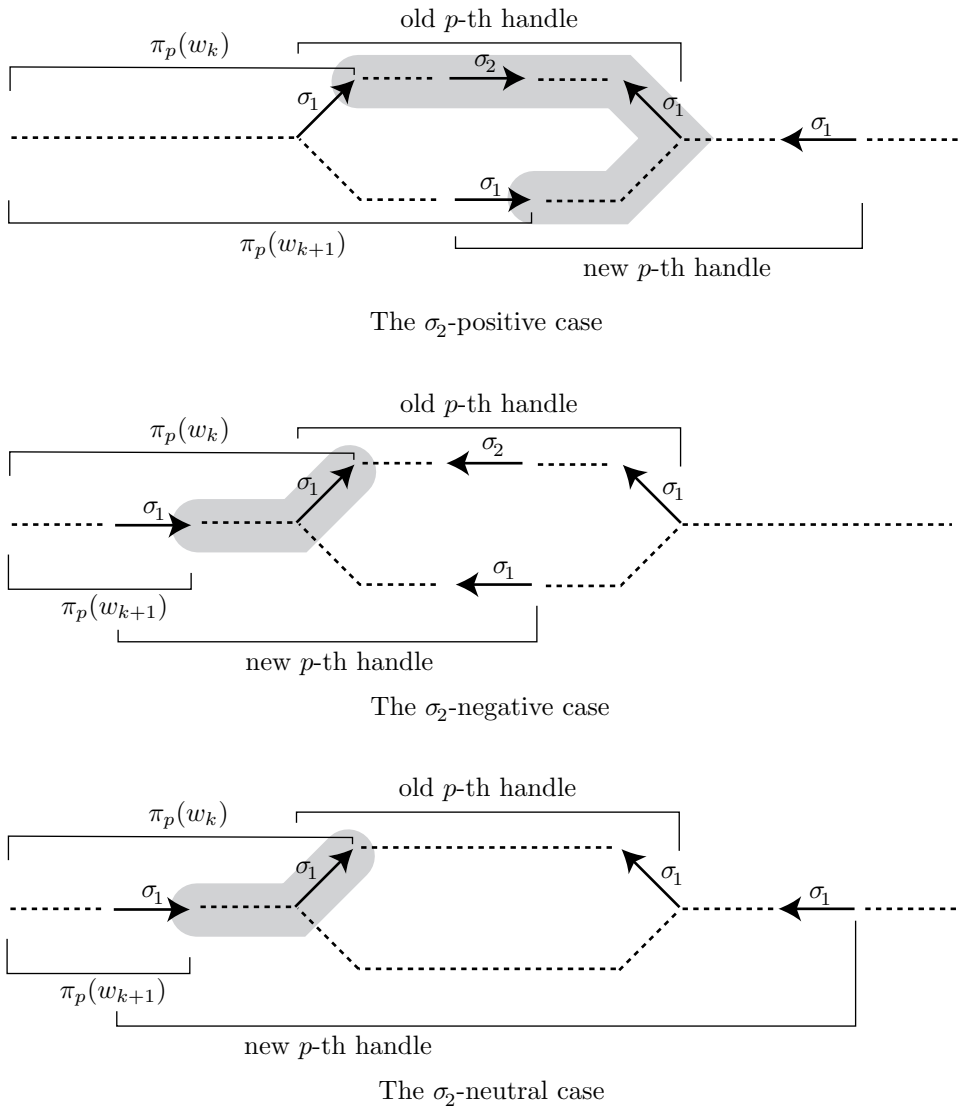


FIGURE 3.5. Critical prefixes

have (using the σ-ordering of braids)

$$\pi_p(w_0) \geqslant \pi_p(w_1) \geqslant \pi_p(w_2) \geqslant \dots, \tag{3.2.6}$$

and, at each step where the handle being reduced is the pth σ_1-handle, the inequality is strict. More precisely, we have the following key result.

Lemma 3.2.18. — *Assume that the pth σ_1-handle is reduced from w_k to w_{k+1}, and that the handle begins with σ_1^{+1}. Then some braid word $u_{p,k}$ traced in $\Gamma(|w|)$ from $\pi_p(w_k)$ to $\pi_p(w_{k+1})$ contains one σ_1^{-1} and no σ_1.*

Proof. — The result can be read on the diagrams of Figure 3.5, where we have represented the paths associated respectively with w_k (up) and w_{k+1} (down) in the Cayley graph of B_∞, assuming that the pth σ_1-handle has been reduced. The word $u_{p,k}$ appears in grey, and the point is that, in every case, *i.e.*, both if σ_2 appears positively or negatively (or not at all) in the handle, this word $u_{p,k}$ contains one letter σ_1^{-1} and no letter σ_1. □

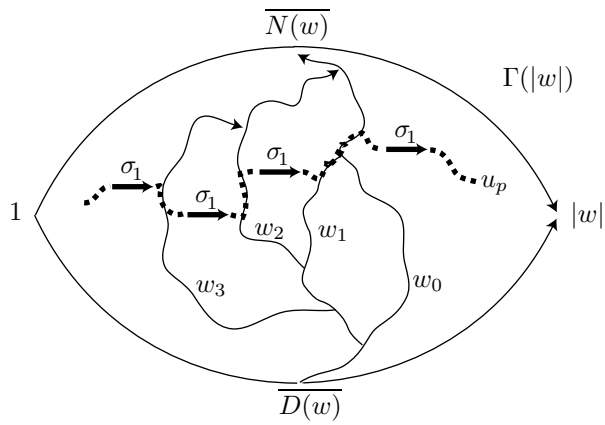


FIGURE 3.6. The witness word u_p contains one letter σ_1 for each reduction of the pth σ_1-handle—and no letter σ_1^{-1}.

If the handle reduction from w_k to w_{k+1} is not the pth σ_1-handle, several cases are possible. If the reduction involves a σ_i-handle with $i \geqslant 2$, or it involves the qth σ_1-handle with $q \neq p \pm 1$, then we have $\pi_p(w_k) \equiv \pi_p(w_{k+1})$, and we complete the definition with $u_{p,k} = \varepsilon$. If the reduction involves the $p\pm1$st σ_1-handle, the equivalence $\pi_p(w_k) \equiv \pi_p(w_{k+1})$ need not be true in general, but, as can be seen on Figure 3.5 again, some word $u_{p,k}$ containing neither σ_1 nor σ_1^{-1} goes from $\pi_p(w_k)$ to $\pi_p(w_{k+1})$ in $\Gamma(|w|)$. Now, by construction, the word $u_p = u_{p,0}u_{p,1}u_{p,2}\dots$ is traced in the Cayley graph of $|w|$, it is σ_1-negative, and the number N of steps in the sequence $(w_0, w_1, \dots)$ where the pth σ_1-handle has been reduced is equal to the number of letters σ_1^{-1} in u (see Figure 3.6). It follows that the number N is bounded above by the height of w, say h.

In the case of a pth handle beginning with σ_1^{-1}, the argument is similar, but Inequality (3.2.6) has to be reversed, and σ_1 and σ_1^{-1} have to be exchanged in Lemma 3.2.18.

Finally, in every case, the heirs of each σ_1-handle of the initial word w are involved in at most h reduction steps, and we can state:

Lemma 3.2.19. — *Assume that w is a braid word of height h containing c σ_1-handles. Then the number of σ_1-handle reductions in any sequence of handle reductions from w is bounded above by ch.*

Assuming again that $w_0 = w, w_1, \dots$ is a sequence of handle reductions from w, we can now iterate the result and consider the σ_2-handle reductions: the previous argument gives an upper bound for the number of σ_2-handle reductions between two successive σ_1-handle reductions, and, more generally, for the number of σ_{i+1}-handle reductions between two σ_i-handle reductions. Using a coarse upper bound on the lengths of the words w_i, one obtains the following generalization of Lemma 3.2.19:

Lemma 3.2.20. — *Assume that w is a braid word of length ℓ, width n, and height h. Then the number of handle reductions in any sequence of handle reductions from w is bounded above by $\ell(2h)^{2n-1}$.*

Proof (sketch). — There are two key points. Firstly, when handle reduction is performed, the height of the words never increases, so it remains bounded by h. Indeed, positive and negative equivalences preserve the absolute value, while left and right reversing preserve it or, possibly, replace it by a word that is a left or a right divisor of the previous absolute value. Secondly, reducing a σ_1-handle may create new σ_2-handles, but this number is bounded by the number of σ_2 (or σ_2^{-1}) that were present in the σ_1-handle that has been reduced. As, by hypothesis, a permitted σ_1-handle includes no σ_2-handle, the number of σ_2's in a permitted σ_1-handle is bounded above by the height h, and, therefore, reducing the σ_1-handles creates at most $h+1$ new σ_2-handles.

Let us consider an arbitrary (finite, or possibly infinite) sequence of reductions starting from w. Writing N_i for the number of σ_i-reductions in this sequence, and c_i for the initial number of σ_i-handles in w, we obtain $N_1 \leqslant c_1 h$ by Lemma 3.2.19, then $N_2 \leqslant (c_2 + N_1(h+1))h$, and, similarly, $N_{i+1} \leqslant (c_{i+1} + N_i(h+1))h$ for every i. Using the obvious bound $c_i \leqslant \ell$, we deduce $N_i \leqslant (2^i - 1)\ell h^{2i-1}$ for each i, and the coarse bound $\sum N_i \leqslant \ell(2h)^{2n-1}$ follows. □

Inserting the previous bound on the height given by Lemma 3.2.17, we deduce Proposition 3.2.2.

3.3. Implementations and variants

The complexity bound of Proposition 3.2.2 might suggest that handle reduction is not efficient in practice, in particular when compared with other solutions to the word problem of braids. This is not the case: we give below experimental data witnessing a much better behaviour.

3.3.1. Strategies. — As it stands, handle reduction is not an algorithm: a given braid word may contain several handles, and, in order to obtain a deterministic method, we have to fix a strategy for choosing which handle should be reduced first. Several choices are natural. We can systematically reduce the first (leftmost) handle (which is then necessarily permitted), or we can try to reduce only the unavoidable handles by considering the σ_1-handles only, together with the nested handles. In the first case, we finish with a word that contains no more handle; in the second case, we finish with a word that is σ-positive, σ-negative, or empty, but, for instance, if the word is σ_1-positive, it may still contain σ_2-handles: this of course is not a problem if we wish only to compare the considered braid with 1.

More efficient versions are obtained by using the classical divide-and-conquer strategy: in order to reduce a (long) word w, we decompose w into w_1w_2 where the length of w_1 and w_2 are approximately equal, we reduce w_1 and w_2 separately (using the same method iteratively), and, having found reduced words w_1', w_2' equivalent to w_1 and w_2, we finally reduce $w_1'w_2'$. If w_1' and w_2' happen to be both σ-positive, or σ-negative (thus with probability 1/2), the last step vanishes.

Handle reduction turns out to be a very efficient method in practice. Table 3.1 gives statistical data comparing the overall computation times needed to reduce random words and to compute their greedy normal form: the latter is the standard method derived from the automatic structure of B_n, and it is known to have a quadratic times complexity (for fixed n). The values given below correspond to the following implementations: in the case of handle reduction, the divide-and-conquer trick is applied until the length reaches 4 times their width (meaning that a 16-strand braid word of length 128 is cut, but one of length at most 64 is not); in the case of the normal form, the letters are entered one after the other and the normal form is computed at each time, which, owing to the key property that the cascades are short in general, increases the speed by diminishing the length of the intermediate words that have to be stored. As the reader can see, handle reduction is more efficient in practice: the average time for reducing a braid word of length 4,000 on any number of strands is less than one second, while computing the normal form is 10 times longer in the case of 4 or 16 strands, and much more when the number of strands increases.

Table 3.2 provides additional information about handle reduction, namely the length of the final (reduced) braid words obtained in the handle reduction algorithm. The reader can see that the final length is never much more than the initial length, which suggests that the exponential upper bound following from Proposition 3.2.2 is *very* far from optimal.

3.3.2. Generalized handles. — We have defined a σ_i-handle to be a braid word of the form $\sigma_i^{\pm 1} u \sigma_i^{\mp 1}$ where $\sigma_1^{\pm 1}, \ldots, \sigma_i^{\pm 1}$ do not occur in u. Let us define a *generalized* σ_i-handle to be a similar braid word $\sigma_i^{\pm 1} u \sigma_i^{\mp 1}$ where only $\sigma_{i-1}^{\pm 1}$ and $\sigma_i^{\pm 1}$ are forbiden in u.

	4 strands	16 strands	64 strands
64 crossings	0.20 *vs.* 5.36	0.03 *vs.* 8.65	0.016 *vs.* 23.1
256 crossings	2.71 *vs.* 77.4	0.45 *vs.* 105	0.14 *vs.* 194
$1,024$ crossings	54.5 *vs.* 1,526	10.2 *vs.* 1,378	1.56 *vs.* 1,899
$4,096$ crossings	1,560 *vs.* 29,900	1,635 *vs.* 21,990	33 *vs.* 23,640

TABLE 3.1. Handle reduction *vs.* normal form: comparison of average CPU times in millisec. on an AMD Duron processor at 750 MHz; samples of $1,000$ random braid words; C++ implementation by Hervé Sibert

	4 strands	16 strands	64 strands
64 crossings	73 (134)	62.6 (90)	63.6 (70)
256 crossings	308 (448)	258 (366)	254 (282)
$1,024$ crossings	1,257 (1,566)	1,126 (2,422)	1,023 (1,122)
$4,096$ crossings	5,034 (5614)	5,745(14,682)	4,169 (5,302)

TABLE 3.2. Length of the final word obtained in handle reduction: average case, and (bracketed) worst case; samples of $1,000$ random braid words; implementations by Hervé Sibert

Then the results for generalized handle reduction are the same as for handle reduction: in particular, the convergence result of Proposition 3.2.2 extends without change.

3.3.3. Coarse reduction. — Let us come back to standard handles. Instead of reducing a σ_i-handle by pushing the $(i+1)$st strand over the next crossings as above, we could also push it over the whole nested part, as displayed in Figure 3.7.

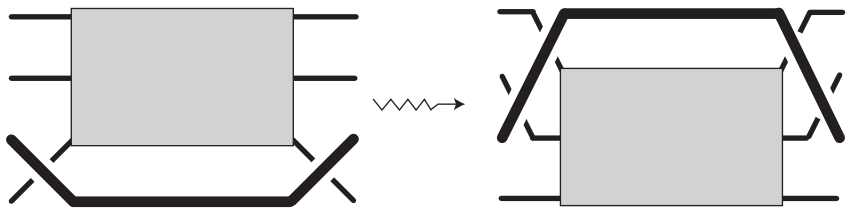


FIGURE 3.7. Coarse reduction

This transformation, which we shall call *coarse reduction*, amounts to replacing a handle of the form $\sigma_i^e \mathrm{sh}(u) \sigma_i^{-e}$ with $\sigma_{i+1}^{-e} \cdots \sigma_{n-1}^{-e}\, u\, \sigma_{n-1}^{e} \cdots \sigma_{i+1}^{e}$.

Convergence of coarse reduction is an open question: the argument of Subsection 3.2.2 is still working, but the one of Subsection 3.2.1 is not, as it need not be

true that every word obtained from a braid word w using coarse reduction must be traced in the Cayley graph of $|w|$. Experiments suggest that coarse reduction always converges, but the proof is still to be found.

CHAPTER 4

FINITE TREES

We describe in this chapter the combinatorial approach developed by Serge Burckel in [**14, 15, 16**] in order to give a constructive proof of the left-well-orderability of the monoid B_n^+ previously established by Richard Laver in [**94**] using the method described in Section 2.4. Burckel's method consists of encoding positive braid words by uniform finite trees (finite trees where all branches have the same length) and deducing a well-ordering $\prec$ on positive braid words from some natural ordering of the associated trees. The main result is then that, if β_1 and β_2 are positive braids, then $\beta_1 < \beta_2$ holds if and only if the $\prec$-least word representing β_1 is $\prec$-smaller than the $\prec$-least word representing β_2. One deduces that the ordered set $(B_n^+, <)$ is order-isomorphic to the ordinal $\omega^{\omega^{n-2}}$.

The tree approach developed here gives proofs of Properties **C** and **S** (assuming no previous result), but, in order to establish that the tree ordering $\prec$ actually induces the σ-ordering $<$, we need to assume Property **A**. We ask whether this approach can prove Property **A** as well, thereby giving a completely independent construction of the σ-ordering.

The techniques used in this chapter are purely combinatorial, and, especially in the general case (4 strands and more), they are not really well understood: there is no doubt that something deep is hidden here, but, frustratingly, the conceptual explanation is still missing.

Throughout the chapter, we denote by W_n the set of all positive n-strand braid words, *i.e.*, the set of all words over the alphabet $\{\sigma_1, \ldots, \sigma_{n-1}\}$. Let us mention that all constructions in Sections 4.1 and 4.2 can be stated in terms of words over an arbitrary alphabet, since the specific braid word equivalence relation $\equiv$ is used from Section 4.3 only.

4.1. Encoding positive braid words in trees

The method relies on coding positive braid words by finite trees. The general idea is to decompose a positive braid word involving n different generators into a product of braid words each of which involves at most $n-1$ different generators. Then, assuming that we have defined an ordering on the latter words, we shall define an ordering on words with n generators by using a lexicographical extension of the order on words on $n-1$ generators.

4.1.1. The case of three strands. — To make reading easier, we first give a separate description for the case of 3-strand braids. In this case, the involved trees simply are sequences of positive integers. Two generators are used, namely σ_1 and σ_2, and the decomposition consists in parsing the word into alternating blocks of σ_1's and σ_2's. Such a decomposition is unique if we require the last word to be a block of σ_2's and every block, except possibly the last one, to be nonempty.

Definition 4.1.1. — For u in W_3, the *code* of u is defined to be the unique sequence of positive integers $(e_1, \dots, e_p)$ satisfying $u = \sigma_1^{e_1}\sigma_2^{e_2} \cdots \sigma_1^{e_{p-1}}\sigma_2^{e_p-1}$ for p even, and $u = \sigma_2^{e_1}\sigma_1^{e_2} \cdots \sigma_1^{e_{p-1}}\sigma_2^{e_p-1}$ for p odd. The integer p is called the *breadth* of u.

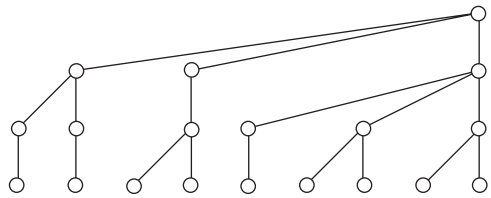
For instance, the code of $\sigma_2^2\sigma_1\sigma_2^3$ is $(2,1,4)$, and its breadth is 3, while the code of $\sigma_1^2\sigma_2\sigma_1^3$ is $(2,1,3,1)$, and its breadth is 4. It is clear that the code of a word determines that word completely, and that, conversely, each finite sequence of positive integers is the code of a unique positive 3-strand braid word.

4.1.2. The general case. — We iterate the coding of 3-strand braid words by sequences of integers to any number of strands, thus using sequences of sequences of integers for 4-strand braid words, sequences of sequences of sequences of integers for 5-strand braid words, *etc.* Such iterated sequences naturally appear as finite trees.

Definition 4.1.2. — An *n-uniform tree* is defined to be a positive integer for $n = 2$, and to be a finite sequence of $(n-1)$-uniform trees for $n \geqslant 3$.

We associate a graph with every uniform tree as follows: for $t = (t_1, \dots, t_p)$, the graph of t consists of a single root connected with the p graphs of $t_1, \dots, t_p$ enumerated from left to right; if t is the integer e, the graph of t is a root with e leaves below it. We use in the sequel the standard vocabulary for trees, thus speaking of root, leaves, nodes, successors of a node (here successor always means immediate successor), *etc.* Observe that every branch in an n-uniform tree, defined as a path from the root to a leaf, has length $n-1$, *i.e.*, it contains n nodes.

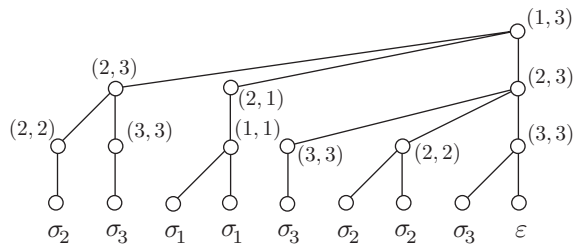
Example 4.1.3. — A typical 4-uniform tree is $((1,1),(2),(1,2,2))$. The associated graph is



In order to construct a coding of braid words by uniform trees, we first define the braid word associated with a uniform tree, *i.e.*, we begin with the decoding process.

Definition 4.1.4. — Assume that t is an n-uniform tree. We attribute *labels* to the inner nodes of t as follows. First, we attribute the label $(1, n-1)$ to the root. Then, if the label (i,j) with $i<j$ has been attributed to a node, we alternatively attribute the labels $(i+1,j)$ and $(j-1,i)$ to the successors of this node starting from the right. Symmetrically, if the label (i,j) with $i>j$ has been attributed to a node, we alternatively attribute the labels $(i-1,j)$ and $(j+1,i)$ to the successors of this node starting from the right. Finally, we attribute to each leaf the label σ_i where (i,i) is the label of its predecessor, except for the rightmost leaf, which receives an empty label ε (we recall that ε denotes the empty word). The word *coded by* t is the word obtained by concatenating the labels of the leaves of t enumerated from left to right.

Example 4.1.5. — For instance, labelling the tree considered in Example 4.1.3 gives



Thus the word coded by this tree is $\sigma_2\sigma_3\sigma_1^2\sigma_3\sigma_2^2\sigma_3$.

Lemma 4.1.6. — *Every positive n-strand braid word is coded by a unique n-uniform tree, i.e., coding establishes a bijective correspondence between n-uniform trees and positive n-strand braid words.*

Proof. — We describe a coding process that associates with each braid word u in W_n an n-uniform tree t. We prove that u is coded by t and t is the unique tree coding u. The inductive argument consists in defining the code of $\sigma_i u$ from the code of u.

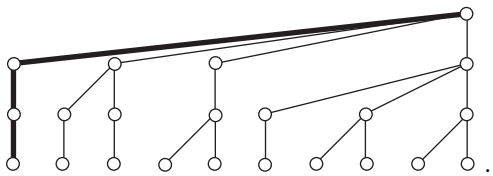
First, we define the code of the empty braid word to be the tree $(\dots(1)\dots)$, $n-2$ pairs of parentheses. The graph associated with this tree consists of a single branch, the label of its unique leaf is ε by definition, the word coded is ε, and ε is

coded by no other tree. Assume now that the code t of u has been constructed and it is unique. Let i be an integer between 1 and $n-1$. Let t' be the tree obtained by adding to t a new leaf on the left, and a branch that connects this leaf to the lowest left node N in t whose label is a pair (j,k) satisfying $j \leqslant i \leqslant k$ or $k \leqslant i \leqslant j$—in which case, we simply say that (j,k) *contains* i. Such a node exists, as the label of the root of t is $(1, n-1)$ by definition. We claim that the word coded by t' is $\sigma_i u$. Indeed, the hypothesis that we take the lowest possible left node N means that the label of the left successor of N in t does not contain i. This implies that the label of the left successor of N in t' has the form (j,i) for some j, and, therefore, the label of the leftmost leaf in t' is σ_i. Now we claim that t' is the only tree coding for $\sigma_i u$. Indeed, if t'' codes for $\sigma_i u$, then the tree obtained from t'' by removing the leftmost branch codes u, so, by induction hypothesis, it is equal to t. So t'' is obtained from t by adding one new branch. Let N' be the node of t where this additional branch is grafted. By construction, the label of N' must contain i, so N' cannot be below N. On the other hand, assume that N' lies strictly above N. Then i is one end of the label of N' and it belongs to the label of its left successor in t'', so, by construction, it cannot belong to the label of the second successor of N' in t'', which is the left successor of N' in t. Hence i cannot belong to the label of N in t, a contradiction. □

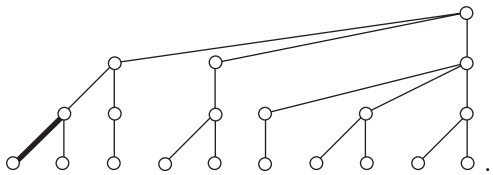
Definition 4.1.7. — For u in W_n, the *n-code* of u is defined to be the unique n-uniform tree t such that u is coded by t.

Example 4.1.8. — Let $u = \sigma_2\sigma_3\sigma_1^2\sigma_3\sigma_2^2\sigma_3$. We have seen that the 4-code of u is the 4-uniform tree $((1,1),(2),(1,2,2))$. For $n \geqslant 4$, the word u belongs to W_n, so it admits an n-code, which is obtained by adding $n-4$ pairs of parentheses around the 4-code of u, *i.e.*, by adding a single branch of length $n-4$ above the root of the 4-code of u.

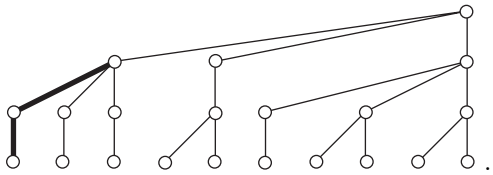
Let us look now for the code of $\sigma_i u$. This code is constructed by adding a new branch to the tree t that codes u. For $i = 1$, the lowest left node of t whose label contains 1 is the root of t. So the coding of $\sigma_1 u$ is



For $i = 2$, the lowest left node of t whose label contains 2 lies on level 2 from the root (just above the leaves). So the coding of $\sigma_2 u$ is



Finally, for $i = 3$, the lowest left node of t whose label contains 3 lies on level 1 from the root. So the coding of $\sigma_3 u$ is



Observe that, by construction, adding the $n-1$ possible generators to a word in W_n corresponds to adding a left branch in the $n-1$ possible ways to a uniform tree of height n.

In the case of 3-strand braid words, the current coding by trees coincides with our previous coding by sequences of positive integers. Coding a braid word by a tree amounts to gathering the letters into blocks of close letters, inductively: for instance, the bracketed expression of the word $\sigma_2\sigma_3\sigma_1^2\sigma_3\sigma_2^2\sigma_3$ is

$$\Big(\big((\sigma_2)(\sigma_3)\big)\ \big((\sigma_1\sigma_1)\big)\ \big((\sigma_3)(\sigma_2\sigma_2)(\sigma_3)\big)\Big).$$

In particular, there exists a distinguished decomposition that corresponds to isolating the successors of the root, for instance $(\sigma_2\sigma_3)\,(\sigma_1\sigma_1)\,(\sigma_3\sigma_2\sigma_2\sigma_3)$ in the previous case.

4.2. A well-ordering on positive braid words

We introduce now our main tool, namely a linear ordering of braid words.

4.2.1. The case of three strands. — Again, we begin with 3-strand braid words.

Definition 4.2.1. — (i) Assume that $<$ is an order on a set X. For $\vec{x}$, $\vec{y}$ finite sequences of elements of X, we say that $\vec{x} <^{\text{ShortLex}} \vec{y}$ is true if either the length of $\vec{x}$ is strictly smaller than the length of $\vec{y}$, or these lengths are equal and $\vec{x}$ precedes $\vec{y}$ with respect to the lexicographical extension of $<$, *i.e.*, for some i we have $x_i < y_i$ and $x_k = y_k$ for $k < i$.
(ii) For u, v in W_3, we say that $u \prec v$ holds if the code of u precedes the code of v with respect to the order $<^{\text{ShortLex}}$ deduced from the standard order of the integers.

The ShortLex-extension of a linear order is a linear order, hence, by construction, $\prec$ is a linear order on W_3. For instance, we have $\sigma_2^2\sigma_1\sigma_2^3 \prec \sigma_1^2\sigma_2\sigma_1^3$, as $(2,1,4) <^{\text{ShortLex}} (2,1,3,1)$ is true: here the breadth of the first word is smaller than the breadth of the second one. On the other hand, we have $\sigma_2^2\sigma_1\sigma_2^2 \prec \sigma_2^2\sigma_1\sigma_2^3$, as $(2,1,3) <^{\text{ShortLex}} (2,1,4)$ holds: here, the breadths are the same, the first two components of the parsings coincide, but not the third one.

Although easy, the following well-ordering property will be crucial in the sequel. We refer to any textbook, for instance [**95**], for basic notions about ordinal arithmetic.

Such details are not crucial here: the only point is that we have a well-ordered set, which enables us to perform inductive arguments.

Proposition 4.2.2. — (i) *The linear order $\prec$ on W_3 is compatible with multiplication on the left, and it includes the subword order: $u \prec \sigma_i u$ and $u \prec u\sigma_i$ hold for every u and every i. Every word u admits an immediate successor with respect to $\prec$, namely $u\sigma_2$.*
(ii) *The linear order $\prec$ on W_3 is a well-ordering of type ω^ω. For every u in W_3, the rank of u in $(W_3, \prec)$, i.e., the order type of the set of all predecessors of u, is*

$$\omega^{p-1} \cdot e_1 + \omega^{p-2} \cdot (e_2 - 1) + \cdots + \omega \cdot (e_{p-1} - 1) + (e_p - 1), \tag{4.2.1}$$

where $(e_1, \dots, e_p)$ is the code of u.

Proof. — (i) Assume $u \prec v$. We wish to show that $\sigma_i u \prec \sigma_i v$ holds for $i = 1, 2$. Let $(e_1, \dots, e_p)$ be the code of u, and $(e'_1, \dots, e'_q)$ be the code of v. By hypothesis, we have

$$(e_1, \dots, e_p) <^{\text{ShortLex}} (e'_1, \dots, e'_q). \tag{4.2.2}$$

If both p and q are even, we have seen above that the codes of $\sigma_1 u$ and $\sigma_1 v$ are respectively $(e_1 + 1, e_2, \dots, e_p)$ and $(e'_1 + 1, e'_2, \dots, e'_q)$, and it is clear that (4.2.2) implies

$$(e_1 + 1, e_2, \dots, e_p) <^{\text{ShortLex}} (e'_1 + 1, e'_2, \dots, e'_q),$$

i.e., $\sigma_1 u \prec \sigma_1 v$. Similarly, the codes of $\sigma_2 u$ and $\sigma_2 v$ are respectively $(1, e_1, \dots, e_p)$ and $(1, e'_1, \dots, e'_q)$, and (4.2.2) implies also

$$(1, e_1, \dots, e_p) <^{\text{ShortLex}} (1, e'_1, \dots, e'_q),$$

which implies $\sigma_2 u \prec \sigma_2 v$. The argument is symmetric if both p and q are odd. Assume now that p and q do not have the same parity. Necessarily $p < q$ holds. The inequalities we have to establish are now

$$(e_1 + 1, e_2, \dots, e_p) <^{\text{ShortLex}} (1, e'_1, e'_2, \dots, e'_q),$$
$$(1, e_1, e_2, \dots, e_p) <^{\text{ShortLex}} (e'_1 + 1, e'_2, \dots, e'_q).$$

The first follows from $p < q$. The second holds because $q \neq 0$ implies $e'_1 \geqslant 1$.

The relation $u \prec \sigma_i u$ holds for all u and i if and only if the inequalities

$$(e_1, e_2, \dots, e_p) <^{\text{ShortLex}} (1, e_1, e_2, \dots, e_p),$$
$$(e_1, e_2, \dots, e_p) <^{\text{ShortLex}} (e_1 + 1, e_2, \dots, e_p)$$

are true, which follows from the definition of $<^{\text{ShortLex}}$. The argument for $u \prec u\sigma_i$ is similar.

Finally, every sequence $(e_1, \dots, e_p)$ admits $(e_1, \dots, e_p{+}1)$ as an immediate successor with respect to $<^{\text{ShortLex}}$. If $(e_1, \dots, e_p)$ is the code of u, then $(e_1, \dots, e_p + 1)$ is the code of $u\sigma_2$. Hence $u\sigma_2$ is the immediate successor of u with respect to $\prec$.

(ii) For each p, the lexicographical order on the length p sequences of positive integers is a well-ordering of type ω^p. Hence, by definition of $<^{\mathrm{ShortLex}}$, the set of all codes is the disjoint sum of a well-ordering of type ω, followed by a well-ordering of type ω^2, followed by a well-ordering of type ω^3, *etc.* Therefore, the codes ordered by $<^{\mathrm{ShortLex}}$ and, equivalently, the words in W_3 ordered by $\prec$, make a well-ordering of type $\omega + \omega^2 + \omega^3 + \cdots$, *i.e.*, of type ω^ω.

Assume that u admits the code $(e_1, \ldots, e_p)$. The order type of the sequences of length at most $p-1$ is ω^{p-1}, so the rank of u in W_3 is $\omega^{p-1} + \alpha$, where α is the rank of the sequence $(e_1, \ldots, e_p)$ among all possible sequences of length exactly p ordered lexicographically. As we consider only sequences with positive entries, α is equal to

$$\omega^{p-1} \cdot (e_1 - 1) + \omega^{p-2} \cdot (e_2 - 1) + \cdots + \omega \cdot (e_{p-1} - 1) + (e_p - 1),$$

which gives (4.2.1). □

Example 4.2.3. — The rank of the word $\sigma_2^2\sigma_1\sigma_2^3$ is the ordinal $\omega^2 \cdot 2 + 3$, as its code is the sequence $(2, 1, 4)$, while the rank of $\sigma_1^2\sigma_2\sigma_1^3$ is $\omega^3 \cdot 2 + \omega \cdot 2$, as its code is $(2, 1, 3, 1)$. By definition of rank, $u \prec v$ holds if and only if the rank of u is less than the rank of v.

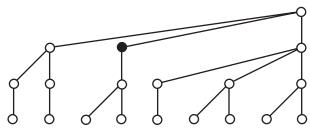
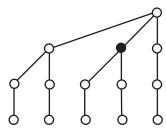
Observe that the well-ordering $\prec$ is *not* compatible with multiplication on the right. For instance, we have $\sigma_2 \prec \sigma_1$ and $\sigma_1^2 \prec \sigma_2\sigma_1$, since $(2) <^{\mathrm{ShortLex}} (1, 1)$ and $(2, 1) <^{\mathrm{ShortLex}} (1, 1, 1)$ hold.

4.2.2. The general case. — The principle for ordering uniform trees is similar.

Definition 4.2.4. — (i) Assume that t and t' are n-uniform trees. For $n = 2$, we say that $t <^{\mathrm{ShortLex}} t'$ holds if $t < t'$ holds (in this case, t and t' are positive integers). Otherwise, assume $t = (t_1, \ldots, t_p)$ and $t' = (t'_1, \ldots t'_{p'})$ where $t_1, \ldots, t'_{p'}$ are $(n-1)$-uniform trees. We say that $t <^{\mathrm{ShortLex}} t'$ holds if either $p < p'$ holds, or there exists i such that $t_k = t'_k$ holds for $k < i$ and $t_i <^{\mathrm{ShortLex}} t'_i$ holds.
(ii) Assume that u and v are positive n-strand braid words. We say that $u \prec v$ holds if the code of u is $<^{\mathrm{ShortLex}}$-smaller than the code of v.

Observe that, in the case of 3-strand braid words, Definition 4.2.4 subsumes Definition 4.2.1, as can be expected.

Example 4.2.5. — Let us consider the words $u = \sigma_2\sigma_3\sigma_1^2\sigma_3\sigma_2^2\sigma_3$ and $v = \sigma_2\sigma_3\sigma_2\sigma_1$. The codes of u and v are


and

respectively.

Then $u \prec v$ holds (although the tree of v appears as thinner than the tree of u):

in both cases the roots have three successors, and the left successors of the root are equal, but the second successor of the root in u has only one successor, while the one in v has two successors (the corresponding nodes appear in black).

Proposition 4.2.6. — (i) *The linear order $\prec$ on W_n is a well-ordering of type $\omega^{\omega^{n-2}}$. It is compatible with multiplication on the left, and it includes the subword order: $u \prec \sigma_i u$ and $u \prec u\sigma_i$ hold for every u and every i.*
(ii) *For u in W_n coded by t, the rank $\mathrm{rk}(u)$ of u in $(W_n, \prec)$ is $t-1$ for $n=2$, and, otherwise, it is*

(4.2.3)
$$\omega^{\omega^{n-3}\cdot(p-1)} \cdot (1+\mathrm{rk}(u_1)) + \omega^{\omega^{n-3}\cdot(p-2)} \cdot \mathrm{rk}(u_2) + \cdots + \omega^{\omega^{n-3}} \cdot \mathrm{rk}(u_{p-1}) + \mathrm{rk}(u_p),$$

where t is $(t_1, \ldots, t_p)$, and $u_1, \ldots, u_p$ are the words of W_{n-1} coded by $t_1, \ldots, t_p$ respectively.

Proof. — The argument is the same as in the case of three strands. An induction on n shows that the order type of $(W_n, \prec)$ is $\omega^{\omega^{n-2}}$, as the order on n-uniform trees is the ShortLex extension of the order on $(n-1)$-uniform trees. Similarly, Formula (4.2.3) asserts that the rank of u is the rank of $(u_1, \ldots, u_p)$ in the ShortLex extension of the order $\prec$ on W_{n-1}. Indeed, the order type of W_{n-1} is $\omega^{\omega^{n-3}}$, hence there are $\omega^{\omega^{n-3}\cdot(p-1)}$ words in W_n with breadth at most $p-1$, and the rank of u is the latter ordinal augmented by the rank of u among all words of breadth p, which is the rank of $(u_1, \ldots, u_p)$ in the lexicographical extension of $\prec$ on W_{n-1}. □

Example 4.2.7. — The rank of the word $u = \sigma_2\sigma_3\sigma_1^2\sigma_3\sigma_2^2\sigma_3$ (considered in the previous examples) is

$$\omega^{\omega\cdot 2} \cdot (1+\mathrm{rk}(u_1)) + \omega^{\omega} \cdot \mathrm{rk}(u_2) + \mathrm{rk}(u_3),$$

where u_1, u_2, u_3 are the 3-strand braid words with respective codes $(1,1)$, (2), and $(1,2,2)$. Applying the same argument to u_1, u_2, and u_3, which amounts to using (4.2.3), we obtain $\mathrm{rk}(u_1) = \omega$, $\mathrm{rk}(u_2) = 1$, and $\mathrm{rk}(u_3) = \omega^2 + \omega + 1$, so, finally, the rank of u is

$$\omega^{\omega\cdot 2} \cdot (1+\omega) + \omega^{\omega} \cdot 1 + \omega^2 + \omega + 1 = \omega^{\omega\cdot 2+1} + \omega^{\omega} + \omega^2 + \omega + 1.$$

Observe that, because of the empty contribution of the rightmost branch in a tree and of the shifting of the letters, the word u is *not* the product of u_1, u_2, u_3.

4.3. Reduction of positive braid words

4.3.1. A proof of Property C. — We shall partition positive braid words into two disjoint subsets, namely the so-called reducible and irreducible words. Simultaneously, we shall construct a reduction map denoted red, defined on reducible words, with the property that, if u is reducible, then $\mathrm{red}(u)$ is an equivalent braid word satisfying

$\mathrm{red}(u) \prec u$. As the relation $\prec$ well-orders braid words, every sequence of reductions leads to an irreducible word in a finite number of iterations, so, in particular, every braid word is equivalent to (at least) one irreducible word. Then the main result is

Proposition 4.3.1. — *Assume that u, v are positive braid words, v is irreducible, and $u \prec v$ holds. Then the word $u^{-1}v$ is equivalent to a σ-positive word, i.e., the braid $\overline{u}^{-1}\overline{v}$ admits a σ-positive representative.*

We deduce Property $\mathbf{C}^+$, which we have seen is equivalent to Property $\mathbf{C}$:

Corollary 4.3.2 **(Property $\mathbf{C}^+$)**. — *Assume that β_1, β_2 are positive braids. Then the braid $\beta_1^{-1}\beta_2$ admits a representative braid word that is σ-positive, σ-negative, or empty.*

Proof. — As the relation $\prec$ is a linear ordering and every positive braid can be represented by an irreducible word (which, as said above, will immediately follow from the definition of reduction), Proposition 4.3.1 implies that, for all positive braids β_1, β_2, the quotient $\beta_1^{-1}\beta_2$ admits a representative braid word that is σ-positive, σ-negative, or empty. □

The rest of the section is devoted to defining reduction and proving Proposition 4.3.1.

4.3.2. Word reduction, case of 3 strands. — As the construction of word reduction is a little intricate in the general case, we once again begin with the special case of three strands.

Definition 4.3.3. — Let u be a positive 3-strand braid word, and $(e_1, \dots, e_p)$ be its code. We say that u is *reducible* if we have $e_k = 1$ for some k with $2 \leqslant k \leqslant p-2$. In this case, we define $\mathrm{red}(u)$ to be the word obtained from u by replacing the rightmost pattern of the form $(x, 1, y, z)$ in the code of u with $(x-1, y, 1, z+1)$, and contracting for $x - 1 = 0$. So, assuming that k is the maximal index satisfying $e_k = 1$ and $k \leqslant p-2$, the word $\mathrm{red}(u)$ admits the code

$$(e_1, \dots, e_{k-2}, e_{k-1} - 1, e_{k+1}, 1, e_{k+2} + 1, e_{k+3}, \dots, e_p), \quad \text{for } e_{k-1} \geqslant 2, \tag{4.3.1}$$

$$(e_1, \dots, e_{k-2} + e_{k+1}, 1, e_{k+2} + 1, e_{k+3}, \dots, e_p), \quad \text{for } e_{k-1} = 1. \tag{4.3.2}$$

In (4.3.1) and (4.3.2), it should be understood that $e_{k+2}+1$, *i.e.*, e_p+1, is the last term in the case $k = p-2$. Observe that (4.3.1) and (4.3.2) are not really different: for $e_{k-1} = 1$, (4.3.1) gives the sequence $(e_1, \dots, e_{k-2}, 0, e_{k+1}, 1, e_{k+2}+1, e_{k+3}, \dots, e_p)$, which is not a code, but becomes (4.3.2) when we gather the two blocks that border the empty block.

Example 4.3.4. — Let $u = \sigma_1\sigma_2\sigma_1^2\sigma_2\sigma_1$, a braid word representing Δ_3^2. Its code is $(1, 1, 2, 1, 1, 1)$, and u is reducible, as the second and the fourth entry in the code is 1. For computing $\mathrm{red}(u)$, we consider the rightmost reducible pattern in the code

of u, here $(2,1,1,1)$, which we replace with $(1,1,1,2)$, obtaining $(1,1,1,1,1,2)$. Thus $\mathrm{red}(u)$ is $\sigma_1\sigma_2\sigma_1\sigma_2\sigma_1\sigma_2$.

The latter word is reducible as well, since the second, third, and fourth entries in its code are 1. Reduction leads to the code $(1,1,0,1,1,3)$, hence, by contraction, to $(1,2,1,3)$. We thus obtain $\mathrm{red}^2(u) = \sigma_1\sigma_2^2\sigma_1\sigma_2^2$. The latter word is an irreducible representative of Δ_3^2, *i.e.*, one no reduction may be applied to.

By definition, the word u is reducible if and only if it contains the pattern $\sigma_1\sigma_2\sigma_1$, or a pattern of the form $\sigma_2\sigma_1\sigma_2^e\sigma_1$ with $e \geqslant 1$. The technical interest of reduction appears in the following result.

Lemma 4.3.5. — *Assume that u is a reducible positive 3-strand braid word. Then we have $\mathrm{red}(u) \equiv u$ and $\mathrm{red}(u) \prec u$.*

Proof. — Constructing $\mathrm{red}(u)$ from u consists in replacing in u a subword of the form $\sigma_i^p\sigma_j\sigma_i^q\sigma_j^r$, with $p,q > 0$ and $\{i,j\} = \{1,2\}$, by the word $\sigma_i^{p-1}\sigma_j^q\sigma_i\sigma_j^{r+1}$. In order to prove $u \equiv \mathrm{red}(u)$, it suffices to check that the latter patterns are equivalent. Now we have

$$\begin{aligned}\sigma_i^p\sigma_j\sigma_i^q\sigma_j^r &= \sigma_i^{p-1}\sigma_i\sigma_j\sigma_i\sigma_i^{q-1}\sigma_j^r \equiv \sigma_i^{p-1}\sigma_j\sigma_i\sigma_j\sigma_i^{q-1}\sigma_j^r \\ &\equiv \sigma_i^{p-1}\sigma_j^2\sigma_i\sigma_j\sigma_i^{q-2}\sigma_j^r \equiv \cdots \equiv \sigma_i^{p-1}\sigma_j^q\sigma_i\sigma_j\sigma_j^r = \sigma_i^{p-1}\sigma_j^q\sigma_i\sigma_j^{r+1}.\end{aligned}$$

It remains to prove $\mathrm{red}(u) \prec u$. We keep the notation of Definition 4.3.3. In the case $e_{k-1} = 1$, the breadth of $\mathrm{red}(u)$ is strictly less than the breadth of u, so the result is true. Otherwise, the breadths are equal, the $k-2$ first entries of the codes coincide, and e_{k-1} in u is replaced with $e_{k-1} - 1$ in $\mathrm{red}(u)$. So $\mathrm{red}(u) \prec u$ holds. □

It follows that, if u is a reducible word, then there exists an integer m such that the word $\mathrm{red}^m(u)$ is irreducible. This word will be denoted $\mathrm{red}^*(u)$ in the sequel. For instance, we have seen in Example 4.3.4 that, for $u = \sigma_1\sigma_2\sigma_1^2\sigma_2\sigma_1$, the word $\mathrm{red}^2(u)$ is irreducible. So $\mathrm{red}^*(u)$ is defined to be the latter word.

4.3.3. The connection between $\prec$ and $<$ in the case of 3 strands. — We are now ready to prove Proposition 4.3.1 in the case of 3-strand braids. We prove the result using ordinal induction on the rank $\mathrm{rk}(v)$ of v in the well-ordered set $(W_3, \prec)$. For ρ an ordinal, we denote by $(\mathcal{S}_\rho)$ the statement:

> *If v is irreducible and $\mathrm{rk}(v) \leqslant \rho$ is true,*
> *then $u \prec v$ implies that $u^{-1}v$ is equivalent to a σ-positive word.*

Our aim is to prove $(\mathcal{S}_\rho)$ for every ordinal ρ with $\rho < \omega^\omega$. First, $(\mathcal{S}_0)$ is vacuously true, as, if v has rank 0, then v is the empty word, and $u \prec v$ never holds. So the point is to prove that, if $(\mathcal{S}_\xi)$ holds for $\xi < \rho$, which we shall denote by $(\mathcal{S}_{<\rho})$, then $(\mathcal{S}_\rho)$ holds as well. Let us assume that v is irreducible with rank ρ, and that $u \prec v$ holds. We wish to prove that $u^{-1}v$ is equivalent to a σ-positive word. It is not easy

to establish the result directly, for we have few hypotheses about u. The idea will be to consider only some special words u below v for which the result can be proved directly, and to show that this is sufficient to conclude in the general case. In the current case of 3-strand braid words, the task will be easy.

Definition 4.3.6. — Assume u, $v \in W_3$, and $p > 0$. Let q be the breadth of v. We say that u is a *p-companion of v* if u has breadth $q-1$ and its code begins with p.

By definition, if u is a companion of v, then $u \prec v$ holds. The converse need not be true, but the companions of v are almost cofinal in v, in the sense that every predecessor of v is a predecessor of some companion of v, or it has a special form.

Lemma 4.3.7. — *Assume that u, v are 3-strand positive braid words and $u \prec v$ holds. Then at least one of the following holds:*
(i) *We have $u = \sigma_i u_0$, $v = \sigma_i v_0$ and $u_0 \prec v_0$ for some i, u_0, v_0;*
(ii) *For p sufficiently large, we have $u \prec u' \prec v$ for every p-companion u' of v.*

Proof. — Let q be the breadth of v. Assume first that u is of breadth q. Then u and v must begin with the same letter, namely σ_1 if q is odd, and σ_2 if q is even. So we can write $u = \sigma_i u_0$, $v = \sigma_i v_0$. As $v_0 \prec u_0$ would imply $v \prec u$, the inequality $u_0 \prec v_0$ must be true.

Assume now that u is of breadth $q-1$. Then, by definition, $u \prec u'$ holds whenever u' is a p-companion of v with $p > e$, where e is the first entry in the code of u.

Finally, for u of breadth $\leqslant q-2$, we have $u \prec u'$ for every companion u' of v. □

We arrive at the key point of the argument.

Lemma 4.3.8. — *Assume that $(\mathcal{S}_{<\rho})$ is true and v is an irreducible 3-strand positive braid word with rank ρ and breadth at least 2. Then, for every p, there exists an irreducible p-companion u of v such that $u^{-1}v$ is equivalent to a σ-positive word.*

Proof. — Let $(e_1, \dots, e_q)$ be the code of v. We define u to admit the code $(p, 2, 2, \dots, 2, 1, 1)$ with length $q-1$ for $q \geqslant 5$, the code $(p, 1, 1)$ for $q = 4$, the code $(p, 1)$ for $q = 3$, and the code (p) for $q = 2$. By definition, u is a p-companion of v, and it is irreducible. We prove that $u^{-1}v$ is equivalent to a σ_1-positive word by considering the possible values of q. For $q = 2$, we have $v = \sigma_1^{e_1}\sigma_2^{e_2-1}$ and $u = \sigma_2^{p-1}$, so $u^{-1}v$ is a σ_1-positive braid word. For $q = 3$, we find $v = \sigma_2^{e_1}\sigma_1^{e_2}\sigma_2^{e_3-1}$, and $u = \sigma_1^p$, hence

$$u^{-1}v = \sigma_1^{-p}\sigma_2^{e_1}\sigma_1^{e_2}\sigma_2^{e_3-1} \equiv \sigma_2\sigma_1\sigma_2^{-(p-1)}\sigma_1^{e_1-1}\sigma_2^{-1}\sigma_1^{e_2-1}\sigma_2^{e_3-1},$$

and the latter word is σ_1-positive. For $q = 4$, we have $v = \sigma_1^{e_1}\sigma_2^{e_2}\sigma_1^{e_3}\sigma_2^{e_4-1}$ and $u = \sigma_2^p\sigma_1$, hence

$$u^{-1}v = \sigma_1^{-1}\sigma_2^{-p}\sigma_1^{e_1}\sigma_2^{e_2}\sigma_1^{e_3}\sigma_2^{e_4-1} \equiv \sigma_2\sigma_1^{-p}\sigma_2^{-1}\sigma_1^{e_1-1}\sigma_2^{e_2}\sigma_1^{e_3}\sigma_2^{e_4-1},$$

and the latter braid word is equivalent to the σ_1-positive word

$$\sigma_2^2\sigma_1\sigma_2^{-(p-2)}\sigma_1^{e_1-1}\sigma_2^{-1}\sigma_1^{e_2-2}\sigma_2^{-1}\sigma_1^{e_3-1}\sigma_2^{e_4-1}.$$

Assume $q \geqslant 5$ with q odd. Write $v = \sigma_2 v_0$, and $u = \sigma_1^p \sigma_2^2 u_0$. Then we have

$$\sigma_1 \sigma_2^{p+2} u_0 \prec v_0 \prec v.$$

Indeed, either v begins with at least two σ_2's, and v_0 has breadth q while $\sigma_1 \sigma_2^{p+2} u_0$ has breadth $q-1$, or v begins with only one σ_2, and both $\sigma_1 \sigma_2^{p+2} u_0$ and v_0 have breadth $q-1$, but the first entry in the code of $\sigma_1 \sigma_2^{p+2} u_0$ is 1, while the first entry in the code of v_0, which is e_2, is at least 2 by hypothesis. Now, v_0 is irreducible, hence the induction hypothesis implies that $(\sigma_1 \sigma_2^{p+2} u_0)^{-1} v_0$ is equivalent to some σ_1-positive word w, and we have then

$$w \equiv (\sigma_1 \sigma_2^{p+2} u_0)^{-1} v_0 \equiv (\sigma_2 \sigma_1 \sigma_2^{p+2} u_0)^{-1} v \equiv (\sigma_1^p \sigma_2 \sigma_1 \sigma_2^2 u_0)^{-1} v.$$

Consider the word $\sigma_1 \sigma_2^2 u_0$: it is irreducible by construction, and $\sigma_1 \sigma_2^2 u_0 \prec v$ holds. Now $\sigma_2 u_0 \prec \sigma_1 \sigma_2^2 u_0$ holds, so the hypothesis $(\mathcal{S}_{<\rho})$ implies that $(\sigma_2 u_0)^{-1} (\sigma_1 \sigma_2^2 u_0)$ is equivalent to some σ_1-positive word w'. We deduce

$$u^{-1} v \equiv (\sigma_2^2 u_0)^{-1} (\sigma_2 \sigma_1 \sigma_2^2 u_0)(\sigma_1^p \sigma_2 \sigma_1 \sigma_2^2 u_0)^{-1} v \equiv w' w,$$

hence $u^{-1} v$ is equivalent to the σ_1-positive word $w'w$. The argument is similar when q is even, exchanging σ_1 and σ_2, which possibly gives a σ_2-positive word. □

We are now ready to complete the argument.

Proof of Proposition 4.3.1, case of 3 strands. — Assume that $(\mathcal{S}_{<\rho})$ is true, and v is an irreducible word with rank ρ. Assume $u \prec v$. By Lemma 4.3.7, two cases are possible. Assume first that $u = \sigma_i u_0$ and $v = \sigma_i v_0$ hold for some i, u_0, v_0. Then, by definition, v_0 is irreducible, and $v_0 \prec v$ holds. Hence, by induction hypothesis, $u_0^{-1} v_0$ is equivalent to a σ_1-positive word, and so is $u^{-1} v$, which is equivalent. Note that this covers in particular the case when the breadth of v is 1.

Assume now that $u \prec u'$ holds for every p-companion u' of v. By Lemma 4.3.8, there exists such a companion u' that is irreducible and such that ${u'}^{-1} v$ is equivalent to a σ-positive word. The induction hypothesis implies that $u^{-1} u'$ is equivalent to a σ-positive word, as $u' \prec v$ holds. We deduce that $u^{-1} v$ is equivalent to a σ-positive word, and $(\mathcal{S}_\rho)$ is true. □

4.3.4. Reducible and irreducible braid words, general case. — Let us turn to the general case. The principle remains the same. As above, we introduce the notion of a reducible word, and define a reduction that maps every braid word that is not irreducible to a $\prec$-smaller equivalent braid word. In order to define reduction, we need a few geometric notions connected with trees. First, we attribute to every node in a uniform tree an *address* that describes the path from the root to that node. Addresses are finite sequences of nonnegative integers.

Definition 4.3.9. — Assume that t is a uniform tree. The empty address is attributed to the root of t, and, assuming that the address α has been given to a node, we attribute the addresses $\alpha 0$, $\alpha 1$, ... to its successors enumerated from right to left.

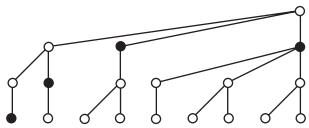
So, for instance, 0^{n-1} is always the address of the rightmost leaf in an n-uniform tree. The reader can check that the addresses of the leaves in the tree of Example 4.1.3 respectively are 210, 200, 101, 100, 020, 011, 010, 001, 000, when enumerated from left to right.

If N is a node in a uniform tree t, the word obtained by concatenating the labels of the leaves below N enumerated from left to right will be called the word *below* N. So, for instance, if t codes for u, then, by definition, u is the word below the root of t.

Definition 4.3.10. — (i) Assume that t is a uniform tree, and N is a node of t that does not lie on the rightmost branch. Let $\alpha(p+1)0^m$ be the address of N in t. We define the *right neighbour* of N in t to be the node with address αp in t.

(ii) Assume that u is a positive braid word, and t is the code of u. The *main decomposition* of u is the sequence $(\sigma_i, u_1, u_2, \dots)$, where σ_i is the first letter in u, *i.e.*, it is the word below the leftmost leaf in t, and u_j is, for every j, the word below N_j in t, where N_0 is the leftmost leaf in t and, for every j, the node N_{j+1} is the right neighbour of N_j in t.

Example 4.3.11. — The right neighbour of a node N is the closest node N_1 on the right of N: we reach N_1 from N by going up in the tree toward the root, and we go one step to the right as soon as possible. For instance, let us consider again the word $\sigma_2\sigma_3\sigma_1^2\sigma_3\sigma_2^2\sigma_3$ of Example 4.1.3. Starting from the left leaf in the associated tree, we find the successive right neighbours



Hence the main decomposition of u is $(\sigma_2, \sigma_3, \sigma_1^2, \sigma_3\sigma_2^2\sigma_3)$.

By construction, the main decomposition of a word u is a decomposition of u in the sense that, if it is $(\sigma_i, u_1, \dots, u_\ell)$, then $u = \sigma_i u_1 \cdots u_\ell$ holds.

For i, j positive integers, we shall denote by $W_{i,j}$ the set of all positive braid words that involve only letters from σ_i to σ_j; we denote by $W_{i,j}^\bullet$ the subset of $W_{i,j}$ consisting of all words in $W_{i,j}$ that finish with σ_j. With this notation, the set W_n of all positive n-strand braid words coincides with $W_{n-1,1}$ and $W_{1,n-1}$. An induction shows that the braid word that lies below a node labelled (i, j) in a uniform tree belongs to $W_{i,j}$, and even to $W_{i,j}^\bullet$, except possibly in the case of a node on the rightmost branch: because the last leaf is labelled ε, we cannot be sure in this case that the considered word finishes with σ_j. It is easy to verify the following result:

Lemma 4.3.12. — *Assume that u is a positive n-strand braid word and the main decomposition of u begins with (σ_i, u_1, u_2). Then exactly one of the following cases occurs:*

(i) *The word u begins with σ_i^2;*
(ii) *We have $u_1 \in W^\bullet_{i+1,j+1}$ and $u_2 \in W^\bullet_{j,i'}$ for some $j \geqslant i$ and $i' \leqslant i$;*
(iii) *We have $u_1 \in W^\bullet_{i-1,j-1}$ and $u_2 \in W_{j,i'}$ for some $j \leqslant i$ and $i' \geqslant i$; moreover, we have $u_2 \in W^\bullet_{j,i'}$ if $\sigma_i u_1 u_2$ is a proper prefix of u, i.e., if the main decomposition of u has length 4 at least.*

We are now ready to define reducible words and the associated reduction mapping. For i, j positive integers, we shall use $\sigma_{i,j}$ to denote the braid word $\sigma_i \cdots \sigma_j$. Then, for every braid word u involving the generators $\sigma_i^{\pm 1}, \ldots, \sigma_{j-1}^{\pm 1}$ only, we have

$$\sigma_{i,j}\, u \equiv \mathrm{sh}^e(u)\, \sigma_{i,j} \tag{4.3.3}$$

with $e = 1$ for $i < j$, and $e = -1$ for $i > j$.

Definition 4.3.13. — Assume that u is a positive n-strand braid word. We say that u is *reducible* if there exists a decomposition $u_0 \sigma_i u_1 u_2 u_3$ of u such that (σ_i, u_1, u_2) is the beginning of the main decomposition of $\sigma_i u_1 u_2 u_3$ and one of the following occurs:
(i) The word u_1 contains neither σ_i nor σ_{i-1} nor σ_{i+1}; then we put (Figure 4.1)

$$\mathrm{red}(u) = u_0 u_1 \sigma_i u_2 u_3;$$

(ii) The word u_1 is $\sigma_{i+1,j+1}$ for some $j \geqslant i$; then write $u_2 = u_2' u_2''$, with $u_2' = \varepsilon$ or $u_2' \in W^\bullet_{i,j}$ and u_2'' not containing σ_j, and put (Figure 4.2)

$$\mathrm{red}(u) = u_0\, \mathrm{sh}(u_2')\, \sigma_{i,j}\, u_2''\, \sigma_{j+1}\, u_3;$$

(iii) The word u_1 is $\sigma_{i-1,j-1}$ for some $j \leqslant i$; then write $u_2 = u_2' u_2''$, with $u_2' = \varepsilon$ or $u_2' \in W^\bullet_{i,j}$, u_2'' not containing σ_j, and $u_3 \neq \varepsilon$, and put (Figure 4.3)

$$\mathrm{red}(u) = u_0\, \mathrm{sh}^{-1}(u_2')\, \sigma_{i,j}\, u_2''\, \sigma_{j-1}\, u_3.$$

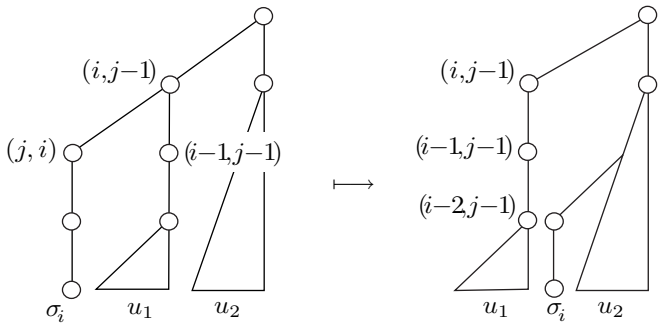


FIGURE 4.1. Reduction, case (i) with $j \leqslant i$ (and $u_0 = \varepsilon$)

The reader can check that the current notion of reduction generalizes the one considered previously in the case of 3 strands. Reduction consists in pushing to the right a generator or a sequence of consecutive generators through a block it commutes

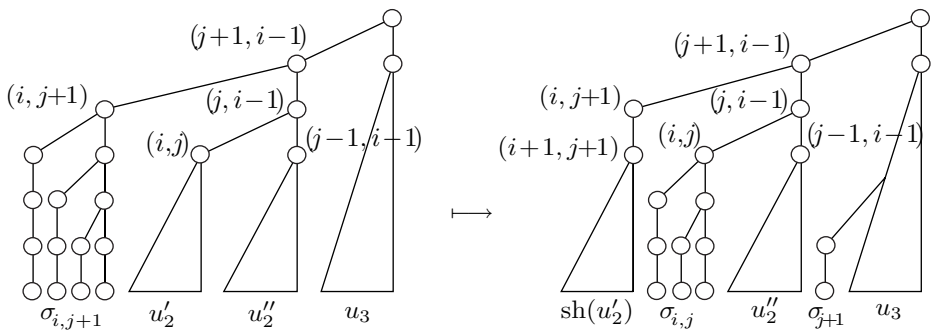


FIGURE 4.2. Reduction, case (ii) (with $u_0 = \varepsilon$)

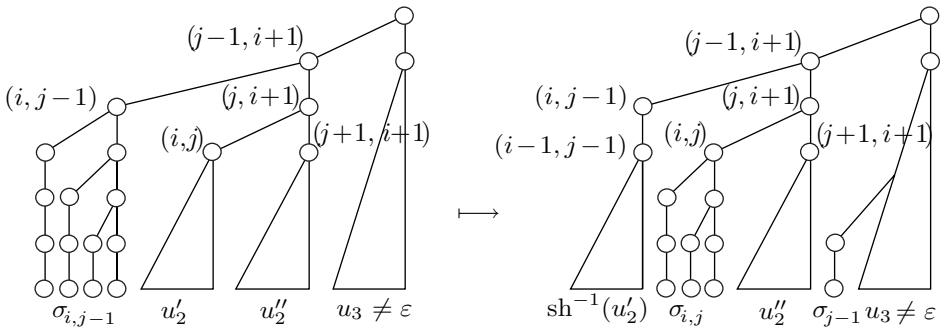


FIGURE 4.3. Reduction, case (iii) (with $u_0 = \varepsilon$)

with, possibly at the expense of shifting the indices. The problem is to recognize where we stop, and this can be decided only when the geometry of the tree associated with the word is known. In particular, in the case of a pattern $\sigma_1\sigma_3$, there is no uniform rule to decide whether the generator σ_1 has to be pushed through σ_3 or not: the word $\sigma_1\sigma_3$ is irreducible, but, as checked below, the word $\sigma_1\sigma_3\sigma_2\sigma_1$ is reducible to $\sigma_3\sigma_1\sigma_2\sigma_1$, where the letter σ_1 has crossed the letter σ_3. This is why the definition of a reducible word does not state something like "assume that u has a subword of the form ...", but involves the more complicated framework of the main decomposition.

Example 4.3.14. — Consider the word $u = \sigma_1\sigma_3\sigma_2\sigma_1$. The suffixes σ_1, $\sigma_2\sigma_1$, and $\sigma_3\sigma_2\sigma_1$ of u are irreducible, but u is reducible. Indeed, the main decomposition of u is $(\sigma_1)(\sigma_3)(\sigma_2\sigma_1)$, and u is reducible since "u_1", here σ_3, contains no σ_1 or σ_2. Applying case (i) of reduction yields the word $\sigma_3\sigma_1\sigma_2\sigma_1$. The suffixes σ_1 and $\sigma_2\sigma_1$ are irreducible, but $\sigma_1\sigma_2\sigma_1$ is reducible. Indeed, its main decomposition is $(\sigma_1)(\sigma_2)(\sigma_1)$, and "$u_1$", *i.e.*,

σ_2, is $\sigma_{2,2}$, while "u_2" is σ_1. Applying case (ii) of reduction yields the word $\sigma_3\sigma_2\sigma_1\sigma_2$, which is irreducible.

The extension of Lemma 4.3.5 is:

Lemma 4.3.15. — *Assume that u is a reducible positive braid word. Then we have* $\mathrm{red}(u) \equiv u$ *and* $\mathrm{red}(u) \prec u$.

We skip the proof, which consists in checking all the possible cases. (We suggest to the reader to draw the trees associated with the words of Example 4.3.14 and to check that they form a decreasing sequence with respect to the ShortLex ordering.) As in the case of 3-strand braid words, the fact that the order $\prec$ is a well-ordering implies that, for every braid word u, there exists a finite integer k such that the word $\mathrm{red}^k(u)$ is irreducible. The latter word will still be denoted by $\mathrm{red}^*(u)$.

4.3.5. The connection between $\prec$ and $<$. — As above, we denote by $(\mathcal{S}_\rho)$ the statement

If v is irreducible and $\mathrm{rk}(v) \leqslant \rho$ *is satisfied,*
then $u \prec v$ implies that $u^{-1}v$ is equivalent to a σ-positive word.

Our aim is to prove $(\mathcal{S}_\rho)$ inductively for every ordinal ρ. To this end, we again introduce some special words, called the companions of v, for which the result is easier, and we show that almost every predecessor of v is a predecessor of some companion of v.

Definition 4.3.16. — Assume that v is a nonempty positive n-strand braid word. Let $\alpha(q+1)0^m$ be the address of the leftmost leaf in the code of t. We say that u is a *p-companion* of v if the leftmost address in the code of u begins with αqp.

By definition, $u \prec v$ holds for every companion u of v: the main parts of the shapes of the codes of u and v coincide, but we removed the leftmost branch from u to v.

The convenient extension of Lemma 4.3.7 guaranteeing that there are enough companions below a given word is as follows:

Lemma 4.3.17. — *Assume that u, v are n-strand positive braid words and $u \prec v$ holds. Then at least one of the following holds:*
(i) *We have $u = \sigma_i u_0$ and $v = \sigma_i v_0$ with $u_0 \prec v_0$ for some i, u_0, v_0;*
(ii) *We have $v = \sigma_i^2 v_0$ and $u \prec \sigma_i v_0$ for some i, v_0;*
(iii) *The word v begins with no σ_i^2 and, for p sufficiently large, $u \prec u'$ holds for every p-companion u' of v.*

The argument consists in considering all possible shapes for the tree associated with v: only in some very special cases, typically when the main decompositions of u and v are too similar, we have enough space to find companions of v that are bigger than u.

The key technical result is then the following counterpart to Lemma 4.3.8.

Lemma 4.3.18. — *Assume that* $(\mathcal{S}_{<\rho})$ *is true and* v *is an irreducible* n*-strand positive braid word of rank* ρ *such that the main decomposition of* v *has length at least* 3*, and* v *begins with* σ_i^2 *for no* i*. Then, for every* p*, there exists an irreducible* $2p$*-companion* u *of* v *such that* $u^{-1}v$ *is equivalent to a* σ*-positive word.*

We skip the proof, which carefully takes all cases into account. We are now ready to complete the argument, *i.e.*, to establish Proposition 4.3.1.

Proof of Proposition 4.3.1, general case. — Our aim is to prove Property $(\mathcal{S}_\rho)$ for every ordinal ρ less than ω^{ω^ω}. First, as in the case of 3 strands, $(\mathcal{S}_0)$ is vacuously true, and the point is to prove that $(\mathcal{S}_{<\rho})$ implies $(\mathcal{S}_\rho)$. So, we assume that v is irreducible with rank ρ, and that $u \prec v$ holds.

By Lemma 4.3.17, three cases are possible. Assume first $u = \sigma_i u_0$, $v = \sigma_i v_0$ and $u_0 \prec v_0$. Then, by definition, v_0 is irreducible, and $v_0 \prec v$ holds. Hence, by induction hypothesis, $u_0^{-1}v_0$ is equivalent to a σ-positive word, and, therefore, so is $u^{-1}v$, which is equivalent.

Assume then $v = \sigma_i^2 v_0$ and $u \prec \sigma_i v_0$. Then, by hypothesis, $\sigma_i v_0$ is irreducible, and $\sigma_i v_0 \prec v$ holds. Hence, by induction hypothesis, $u^{-1}\sigma_i v_0$ is equivalent to a σ-positive word. For a similar reason, $v_0^{-1}\sigma_i v_0$ is equivalent to a σ-positive word, and we deduce that $u^{-1}v$, which is equivalent to the product of the latter words, is equivalent to a σ-positive word.

Assume now that $u \prec u'$ holds for every p-companion u' of v with p sufficiently large. By Lemma 4.3.18, and, because $(\mathcal{S}_{<\rho})$ is assumed to be true by induction hypothesis, there exists such an irreducible companion u' of v such that ${u'}^{-1}v$ is equivalent to a σ-positive word. As $u' \prec v$ holds, the induction hypothesis implies that $u^{-1}u'$ is also equivalent to a σ-positive word, and, again, we deduce that $u^{-1}v$ is equivalent to a σ-positive word as σ-positive braid words are closed under multiplication.

Finally, assume that the main decomposition of v has length at most 2, *i.e.*, v has the form $\sigma_i v_1$ with v_1 in $W_{i+1,n-1}$. Then $u \prec v$ implies either $u = \sigma_i u_1$ with $u_1 \prec v_1$, and this case has been treated above, or $u \in W_{i+1,n-1}$. In the latter case, σ_i occurs in $u^{-1}v$, but neither σ_i^{-1} nor any $\sigma_k^{\pm 1}$ with $k < i$ does, so $u^{-1}v$ is σ-positive. We conclude that $(\mathcal{S}_\rho)$ is true in each case. □

Thus our proof of Property $\mathbf{C}^+$ is complete.

4.4. Applications

4.4.1. A characterization of the braid ordering. — So far Property **A** did not enter the picture: we did not use it, nor did we prove it either. When we introduce

it, we obtain a new characterization of irreducible braid words, and we deduce a new form for the braid ordering.

Proposition 4.4.1. — *A positive braid word v is irreducible if and only if it is the $\prec$-least element of its $\equiv$-equivalence class.*

Proof. — By Lemma 4.3.15, if v is not irreducible, we have $\mathrm{red}(v) \prec v$, and, therefore, v is not the $\prec$-least element of its class. Conversely, assume that v is irreducible, and $u \prec v$ holds. By Proposition 4.3.1, $u^{-1}v$ is equivalent to a σ-positive word, so, by Property **A**, the braids $\overline{u}$ and $\overline{v}$ cannot be equal, *i.e.*, we have $u \not\equiv v$. Hence v is the $\prec$-least element in its $\equiv$-equivalence class. □

We then obtain the fourth equivalent definition of the braid ordering mentioned in Introduction:

Proposition 4.4.2. — *Assume that β_1, β_2 are positive braids. Then $\beta_1 < \beta_2$ holds if and only if we have $u_1 \prec u_2$, where u_i is the unique irreducible word representing β_i.*

(In the above "if" statement, it is sufficient to assume that u_2 is irreducible.) As an application, we deduce:

Proposition 4.4.3. — *For every integer n, the restriction of the braid order $<$ to B_n^+ is a well-ordering of type $\omega^{\omega^{n-2}}$.*

Proof. — Let W_n^{irred} denote the set of all irreducible words in W_n. By Proposition 4.4.2, $(B_n^+, <)$ is isomorphic to $(W_n^{irred}, \prec)$. As $\prec$ is a well-ordering on W_n, its restriction to W_n^{irred} is also a well-ordering, and the order type of $(W_n^{irred}, \prec)$ is at most the type of $(W_n, \prec)$, which is $\omega^{\omega^{n-2}}$. As some words in W_n are reducible, the previous inequality could be strict. This however does not happen. Indeed, let us denote by d the mapping of W_n into itself that maps every word u to the word u' such that the tree associated with u' is obtained from the tree associated with u by doubling the degree of every inner node. For instance, the tree associated with $\sigma_3\sigma_1$ is $((1),(1),(1))$, hence the tree associated with $d(u)$ is $((2,2),(2,2),(2,2),(2,2),(2,2),(2,2))$, and we have $d(u) = (\sigma_2^2\sigma_1^2\sigma_2^2\sigma_3^2)^2(\sigma_2^2\sigma_1^2\sigma_2^2\sigma_3)$. Then the image of d consists of irreducible words only, and d is $\prec$-increasing. Thus, with obvious notation, the order type of $(d(B_n^+), <)$ is $\omega^{\omega^{n-2}}$, and, therefore, the order type of $(B_n^+, <)$ is at least, hence is exactly, this ordinal. □

4.4.2. A proof of Property S. — We have seen in Chapter 1 that Property **S** implies that the restriction of the braid ordering to B_n^+ is a well-ordering, but we have established no converse implication. So the proof of the latter result given above gives no proof of Property **S**. However, we shall see now how to prove the latter directly.

Proposition 4.4.4. — *Assume that u is an irreducible n-strand positive braid word. Then, for each i, the word $u^{-1}\sigma_i u$ is equivalent to a σ-positive word.*

Proof. — We use induction on the rank of u in $(W_n, \prec)$. By definition, we have $u \prec \sigma_i u$, hence, if $\sigma_i u$ is irreducible, the result follows from Proposition 4.3.1. Assume now that $\sigma_i u$ is reducible, say $u = u_1 u_2 u_3$, where (σ_i, u_1, u_2) is the beginning of the main decomposition of $\sigma_i u$. Assume first that u_1 does not contain $\sigma_{i\pm 1}$. Then we know that the word $u_2 u_3$ is irreducible, and $u_2 u_3 \prec u$ holds. Hence, by induction hypothesis, the braid $u_3^{-1} u_2^{-1} \sigma_i u_2 u_3$ is equivalent to a σ-positive word, and it is equivalent to $u^{-1}\sigma_i u$ as $\overline{u_1}$ and σ_i commute in B_n^+.

Assume now $u_1 = \sigma_{i+1,j+1}$ and $u_2 = u_2' u_2''$ with $u_2' \in W_{i,j}^{\bullet} \cup \{\varepsilon\}$ and u_2'' not containing σ_j. As u is assumed to be irreducible, the letter σ_i occurs in u_2'. So let us write $u_2'' = v\sigma_i w$, where v does not contain σ_i. As $w\, u_2''\, u_3 \prec u_2\, u_3$ holds, we have $\sigma_{i+2,j+1}\, w\, u_2''\, u_3 \prec \sigma_{i+2,j+1}\, u_2\, u_3 \prec u$. Applying the induction hypothesis, we deduce that $(\sigma_{i+2,j+1} w u_2'' u_3)^{-1} (\sigma_{i+1,j+1} w u_2'' u_3)$ is equivalent to a σ-positive word. Now this word is also equivalent to

$$(v\sigma_{i+1}\sigma_i\sigma_{i+2,j+1} w u_2'' u_3)^{-1} (v\sigma_{i+1}\sigma_i\sigma_{i+1,j+1} w u_2'' u_3). \tag{4.4.1}$$

Applying braid relations, we see that the first factor of (4.4.1) is equivalent to the inverse of $u_1 u_2 u_3$, *i.e.*, of u, while the second one is equivalent to $\sigma_i u_1 u_2 u_3$, *i.e.*, to $\sigma_i u$. So $u^{-1}\sigma_i u$ is equivalent to a σ-positive word. □

We deduce Property $\mathbf{S}^+$, which we have seen in Remark 1.2.16 to be equivalent to Property $\mathbf{S}$ (the same statement with β an arbitrary braid):

***Corollary 4.4.5* (Property $\mathbf{S}^+$).** — *Every braid of the form $\beta^{-1}\sigma_i\beta$ with β a positive braid admits a σ-positive representative braid word.*

Remark 4.4.6. — In the case of B_3^+, an inductive argument is not necessary. Indeed, if $(e_1, \dots, e_p)$ is the code of an irreducible word u, then the code of $\mathrm{red}^*(\sigma_i u)$ is $(e_1 + 1, e_2, \dots, e_p)$ if $p - i$ is odd, $(1, e_1, \dots, e_p)$ if $p - i$ is even and $e_1 \geqslant 2$ or $p \leqslant 2$ holds, and $(e_2 - 1, e_3, \dots, e_{p-2}, e_{p-1} + 1, 1, e_p + 1)$ if $p - i$ is even, and $e_1 = 1$ and $p \geqslant 3$ hold. In all three cases, we obtain $u \prec \mathrm{red}^*(\sigma_i u)$, and, therefore, we deduce that $u^{-1}\mathrm{red}^*(\sigma_i u)$, which is equivalent to $u^{-1}\sigma_i u$, is equivalent to a σ-positive word using Proposition 4.3.1.

4.4.3. Tables of normal forms and ranks. — We give below a list of the first braids in the well-orderings $(B_3^+, <)$ and $(B_4^+, <)$, together with their code and their rank. Every braid appears there exactly once, and the given representative is the unique irreducible one. For instance, $\sigma_2\sigma_1\sigma_2$ appears in Table 1, but $\sigma_1\sigma_2\sigma_1$ does not appear, as it is not irreducible. Note that the rank mentioned here is the rank with respect to $<$, which is not the rank with respect to $\prec$, as reducible words are skipped: the braid $\sigma_1\sigma_2^2\sigma_1$ is the ω^3th element of $(B_3^+, <)$, while, by Proposition 4.2.2, the word $\sigma_1\sigma_2^2\sigma_1$ is the $(\omega^3 + \omega^2)$th element of $(W_3, \prec)$: the missing words are the words $\sigma_1\sigma_2\sigma_1^{p+1}\sigma_2^q$, which are reducible.

braid	1	σ_2	σ_2^2	$\dots$	σ_1	$\sigma_1\sigma_2$	$\sigma_1\sigma_2^2$	$\dots$	σ_1^2	$\sigma_1^2\sigma_2$	$\dots$
code	(1)	(2)	(3)	$\dots$	$(1,1)$	$(1,2)$	$(1,3)$	$\dots$	$(2,1)$	$(2,2)$	$\dots$
rank	0	1	2	$\dots$	ω	$\omega+1$	$\omega+2$	$\dots$	$\omega\cdot 2$	$\omega\cdot 2+1$	$\dots$

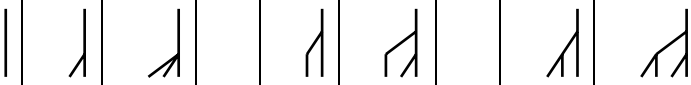
braid	$\sigma_2\sigma_1$	$\sigma_2\sigma_1\sigma_2$	$\dots$	$\sigma_2\sigma_1^2$	$\sigma_2\sigma_1^2\sigma_2$	$\dots$	$\sigma_2\sigma_1^3$	$\dots$
code	$(1,1,1)$	$(1,1,2)$	$\dots$	$(1,2,1)$	$(1,2,2)$	$\dots$	$(1,3,1)$	$\dots$
rank	ω^2	ω^2+1	$\dots$	$\omega^2+\omega$	$\omega^2+\omega+1$	$\dots$	$\omega^2+\omega\cdot 2$	$\dots$

braid	$\sigma_2^2\sigma_1$	$\dots$	$\sigma_2^2\sigma_1^2$	$\dots$	$\sigma_1\sigma_2^2\sigma_1$	$\dots$	$\sigma_1\sigma_2^2\sigma_1^2$	$\dots$	$\sigma_1\sigma_2^3\sigma_1$
code	$(2,1,1)$	$\dots$	$(2,2,1)$	$\dots$	$(1,2,1,1)$	$\dots$	$(1,2,2,1)$	$\dots$	$(1,3,1,1)$
rank	$\omega^2\cdot 2$	$\dots$	$\omega^2\cdot 2+\omega$	$\dots$	ω^3	$\dots$	$\omega^3+\omega$	$\dots$	$\omega^3+\omega^2$

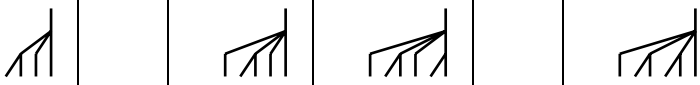
braid	$\dots$	$\sigma_1^2\sigma_2^2\sigma_1$	$\dots$	$\sigma_2\sigma_1^2\sigma_2^2\sigma_1$	$\dots$	$\sigma_1\sigma_2^2\sigma_1^2\sigma_2^2\sigma_1$	$\dots$	$\dots$
code	$\dots$	$(2,2,1,1)$	$\dots$	$(1,2,2,1,1)$	$\dots$	$(1,2,2,2,1,1)$	$\dots$	$\dots$
rank	$\dots$	$\omega^3\cdot 2$	$\dots$	ω^4	$\dots$	ω^5	$\dots$	$\dots$

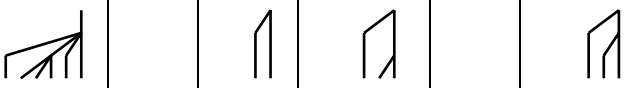
TABLE 4.1. Normal form and rank in $(B_3^+, <)$

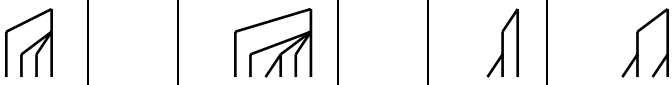
Let us mention that computing the rank of an irreducible braid word is in general a complicated process, and refer to [**16**] for more details.

braid	1	σ_3	σ_3^2	$\dots$	σ_2	$\sigma_2\sigma_3$	$\dots$	σ_2^2	$\sigma_2^2\sigma_3$
code									
rank	0	1	2	$\dots$	ω	$\omega+1$	$\dots$	$\omega\cdot 2$	$\omega\cdot 2+1$

braid	$\dots$	σ_2^3	$\dots$	$\sigma_3\sigma_2$	$\dots$	$\sigma_3\sigma_2^2$	$\dots$	$\sigma_3\sigma_2^3$
code								
rank	$\dots$	$\omega\cdot 3$	$\dots$	ω^2	$\dots$	$\omega^2+\omega$	$\dots$	$\omega^2+\omega\cdot 2$

braid	$\dots$	$\sigma_3^2\sigma_2$	$\dots$	$\sigma_2\sigma_3^2\sigma_2$	$\sigma_2\sigma_3^2\sigma_2\sigma_3$	$\dots$	$\sigma_2\sigma_3^2\sigma_2^2$
code							
rank	$\dots$	$\omega^2\cdot 2$	$\dots$	ω^3	ω^3+1	$\dots$	$\omega^3+\omega$

braid	$\dots$	$\sigma_2\sigma_3^3\sigma_2$	$\dots$	σ_1	$\sigma_1\sigma_3$	$\dots$	$\sigma_1\sigma_2$	$\dots$
code								
rank	$\dots$	$\omega^3+\omega^2$	$\dots$	ω^ω	$\omega^\omega+1$	$\dots$	$\omega^\omega+\omega$	$\dots$

braid	$\dots$	$\sigma_1\sigma_3\sigma_2$	$\dots$	$\sigma_1\sigma_2\sigma_3^2\sigma_2$	$\dots$	σ_1^2	$\sigma_1^2\sigma_3$
code							
rank	$\dots$	$\omega^\omega+\omega^2$	$\dots$	$\omega^\omega+\omega^3$	$\dots$	$\omega^\omega\cdot 2$	$\omega^\omega\cdot 2+1$

braid	$\dots$	$\sigma_1^2\sigma_2$	$\dots$	$\sigma_2\sigma_1$	$\dots$	$\sigma_2\sigma_1^2$	$\dots$
code							
rank	$\dots$	$\omega^\omega\cdot 2+\omega$	$\dots$	$\omega^{\omega+1}$	$\dots$	$\omega^{\omega+1}+\omega^\omega$	$\dots$

braid	$\sigma_2^2\sigma_1$	$\dots$	$\sigma_1\sigma_2\sigma_1^2$	$\dots$	$\sigma_3\sigma_2\sigma_1$	$\dots$
code						
rank	$\omega^{\omega+1}\cdot 2$	$\dots$	$\omega^{\omega+2}$	$\dots$	$\omega^{\omega\cdot 2}$	$\dots$

TABLE 4.2. Tree normal form and rank in $(B_4^+, <)$

CHAPTER 5

AUTOMORPHISMS OF A FREE GROUP

This short chapter is a transition between the combinatorial and the topological approaches, and it contains an account of the work developed by David Larue in [**89, 88**]. One obtains in this way a very short proof of Property **A**, and a (not so simple) proof of Property **C**—together with a new very simple proof of the fact that B_n embeds into $\mathrm{Aut}(F_n)$. Some arguments will simply be sketched, as they can be more naturally viewed from a topological perspective, as will be done in Chapters 6 and 7.

5.1. Keeping track of the letters

It has been known since Artin that the braid groups embed in groups of automorphisms of free groups, and it is natural to look at the images of those braids that admit σ-positive representatives under such embeddings.

5.1.1. Embedding B_n into $\mathrm{Aut}(F_n)$. — For $1 \leqslant n < \infty$, we denote by F_n the rank n free group based on $\{x_1, \dots, x_n\}$, and we denote by F_∞ the free group based on $\{x_i\,;\, 1 \leqslant i < \infty\}$.

Definition 5.1.1. — For $i < n$, we denote by $\widehat{\sigma}_i$ the automorphism of F_n defined by

$$\widehat{\sigma}_i(x_k) = \begin{cases} x_i x_{i+1} x_i^{-1} & \text{for } k = i, \\ x_i & \text{for } k = i+1, \\ x_k & \text{for } k \neq i, i+1. \end{cases} \tag{5.1.1}$$

Lemma 5.1.2. — *For $1 \leqslant n \leqslant \infty$, the mapping $\sigma_i \mapsto \widehat{\sigma}_i$ extends to a homomorphism of B_n into $\mathrm{Aut}(F_n)$.*

Proof. — The automorphisms $\widehat{\sigma}_i$ satisfy the braid relations. Alternatively, we can observe that the action of $\widehat{\sigma}_i$ on $(F_n)^n$ is the right action of conjugacy considered in

Section 2.1, and its compatibility with braid relations follows from conjugacy being a left self-distributive operation. □

For each braid β, we shall denote by $\widehat{\beta}$ the associated automorphism. For w a braid word, we denote by $\widehat{w}$ the automorphism associated with the braid represented by w.

The embedding of B_n into B_{n+1} induced by identity on σ_i's is compatible with the embedding of $\mathrm{Aut}(F_n)$ into $\mathrm{Aut}(F_{n+1})$ induced by the identity on x_i's, so there is no need to specify n here; equivalently, we may consider that we work with B_∞ and $\mathrm{Aut}(F_\infty)$.

5.1.2. Images of σ-positive braid words. — We shall prove that, if the braid β admits at least one σ-positive representative braid word, then the automorphism $\widehat{\beta}$ has some specific properties that can be read on the words $\widehat{\beta}(x_i)$.

In the sequel, we identify F_∞ with the set of all freely reduced words on $x_1^{\pm1}$, $x_2^{\pm1}$, $\ldots$, where we say that u is *freely reduced* if it contains no pattern of the form xx^{-1} or $x^{-1}x$. For u an arbitrary word on $x_1^{\pm1}$, $x_2^{\pm1}$, $\ldots$, we denote by $\mathrm{red}(u)$ the unique reduced word obtained from u by iteratively deleting all patterns xx^{-1} and $x^{-1}x$.

Definition 5.1.3. — For x a letter x_i or x_i^{-1}, we denote by $S(x)$ the subset of F_∞ consisting of all freely reduced words that end with x.

We shall investigate the image of the set $S(x_1^{-1})$ under the automorphism $\widehat{\sigma}_i^{\pm1}$. We denote by sh the (shift) endomorphism of the free monoid generated by $x_1^{\pm1}$, $x_2^{\pm1}$, $\ldots$ that maps $x_k^{\pm1}$ to $x_{k+1}^{\pm1}$ for every k. For f in $\mathrm{Aut}(F_\infty)$, we denote by $\mathrm{sh}(f)$ the automorphism of F_∞ defined by $\mathrm{sh}(f)(x_1) = x_1$, and $\mathrm{sh}(f)(x_{k+1}) = \mathrm{sh}(f(x_k))$.

Lemma 5.1.4. — *Every automorphism* $\mathrm{sh}(f)$ *maps* $S(x_1^{-1})$ *into itself.*

Proof. — Let us consider an arbitrary element of $S(x_1^{-1})$, say ux_1^{-1}, with $u \notin S(x_1)$. By construction, we have $\mathrm{sh}(f)(ux_1^{-1}) = \mathrm{red}\big(\mathrm{sh}(f)(u)\, x_1^{-1}\big)$. Assume $\mathrm{sh}(f)(ux_1^{-1}) \notin S(x_1^{-1})$. Then the final letter x_1^{-1} in $\mathrm{sh}(f)(u)x_1^{-1}$ is cancelled by some letter x_1 occurring in $\mathrm{sh}(f)(u)$. Such a letter x_1 in $\mathrm{sh}(f)(u)$ must come from a letter x_1 in u. So there exists a decomposition $u = u_1x_1u_2$ satisfying $\mathrm{sh}(f)(u_2) = 1$. As $\mathrm{sh}(f)$ is injective, the latter condition implies $u_2 = 1$, hence $u \in S(x_1)$, contradicting the hypothesis. □

Lemma 5.1.5. — *The automorphism* $\widehat{\sigma}_i$ *maps both* $S(x_i)$ *and* $S(x_i^{-1})$ *into* $S(x_i^{-1})$.

Proof. — Let us consider an arbitrary element of $S(x_i) \cup S(x_i^{-1})$, say ux_i^e with $e = \pm1$ and $u \notin S(x_i^{-e})$. Then we have $\widehat{\sigma}_i(ux_i^e) = \mathrm{red}\big(\widehat{\sigma}_i(u)x_ix_{i+1}^ex_i^{-1}\big)$. Assume $\widehat{\sigma}_i(ux_i^e) \notin S(x_i^{-1})$. This means that the final x_i^{-1} in $\widehat{\sigma}_i(ux_i^e)$ is cancelled by some letter x_i in $\widehat{\sigma}_i(u)$. This letter comes either from some x_{i+1} or from some $x_i^{e'}$ in u.

In the first case, we display the letter x_{i+1} involved in the cancellation by writing $u = u_1 x_{i+1} u_2$, where u_2 is a reduced word. We find

$$\widehat{\sigma}_i(ux_i^e) = \mathrm{red}\big(\widehat{\sigma}_i(u_1)x_i\widehat{\sigma}_i(u_2)x_i x_{i+1}^e x_i^{-1}\big),$$

and the hypothesis $\mathrm{red}\big(\widehat{\sigma}_i(u_2)x_i x_{i+1}^e\big) = \varepsilon$ implies $\widehat{\sigma}_i(u_2) = x_{i+1}^{-e}x_i^{-1} = \widehat{\sigma}_i(x_{i+1}^{-1}x_i^{-e})$. We deduce $u_2 = x_{i+1}^{-1}x_i^{-e}$, contradicting $u \notin S(x_i^{-e})$.

In the second case, we write similarly $u = u_1 x_i^{e'} u_2$ with $e' = \pm 1$. So we have

$$\widehat{\sigma}_i(ux_i^e) = \mathrm{red}\big(\widehat{\sigma}_i(u_1)x_i x_{i+1}^{e'} x_i^{-1}\widehat{\sigma}_i(u_2)x_i x_{i+1}^e x_i^{-1}\big),$$

and the hypothesis is $\mathrm{red}\big(x_{i+1}^{e'}x_i^{-1}\widehat{\sigma}_i(u_2)x_i x_{i+1}^e\big) = \varepsilon$. This implies $\mathrm{red}\big(\widehat{\sigma}_i(u_2)\big) = x_i x_{i+1}^{-e-e'}x_i^{-1} = \widehat{\sigma}_i(x_i^{-e-e'})$, hence $u_2 = x_i^{-e-e'}$. For $e = +1$, we obtain either $u_2 = x_i^{-2}$ (for $e' = +1$) or $u_2 = \varepsilon$ (for $e' = -1$), and, in both cases, $u \in S(x_i^{-e})$, a contradiction. Similarly, for $e = -1$, we obtain either $u_2 = \varepsilon$ (for $e' = +1$) or $u_2 = x_i^2$ (for $e' = -1$), and, in both cases, $u \in S(x_i^{-e})$, again a contradiction. □

We deduce the following implication:

Proposition 5.1.6. — *Assume that the braid β admits at least one σ_1-positive representative braid word. Then the word $\widehat{\beta}(x_1)$ ends with x_1^{-1}.*

Proof. — Our hypothesis implies that the automorphism $\widehat{\beta}$ admits a decomposition of the form

$$\widehat{\beta} = \mathrm{sh}(f_0) \circ \widehat{\sigma}_1 \circ \mathrm{sh}(f_1) \circ \cdots \circ \widehat{\sigma}_1 \circ \mathrm{sh}(f_p).$$

Then we have $\mathrm{sh}(f_p)(x_1) = x_1$, and $\widehat{\sigma}_1(x_1) = x_1 x_2 x_1^{-1}$, an element of $S(x_1^{-1})$. By Lemmas 5.1.4 and 5.1.5, every subsequent factor $\mathrm{sh}(f_k)$ and $\widehat{\sigma}_1$ maps $S(x_1^{-1})$ into $S(x_1^{-1})$. □

***Corollary 5.1.7* (Property A)**. — *A braid that admits at least one σ_1-positive representative braid word is not trivial.*

Proof. — If β admits a σ_1-positive representative, then, by Proposition 5.1.6, the word $\widehat{\beta}(x_1)$ is not equal to x_1, so $\widehat{\beta}$ is not the identity, and, therefore, β cannot be trivial. □

5.1.3. Injectivity results. — Assuming Property **C**, or simply the weaker form $\mathbf{C}_\infty$, *i.e.*, assuming that every braid in B_∞ admits at least one representative braid word that is σ-positive, σ-negative, or empty, we obtain a new proof of the injectivity of the map $\beta \to \widehat{\beta}$.

Proposition 5.1.8. — *The homomorphism $\beta \to \widehat{\beta}$ of B_n to $\mathrm{Aut}(F_n)$ is an embedding.*

Proof. — Assume that β is a nontrivial braid. We claim that $\widehat{\beta}$ is not the identity. By Proposition 5.1.6, this is the case when β admits a σ_1-positive representative, and, more generally, when β admits a σ-positive representative (by injectivity of the shift mapping). By applying the result to β^{-1}, we obtain similarly that $\widehat{\beta}$ is not the identity when β admits a σ-negative representative. Using Property $\mathbf{C}_\infty$, we conclude that $\beta = 1$ is the only case that has not been considered. □

The embedding $\beta \to \widehat{\beta}$ is not the only homomorphism of B_n to $\mathrm{Aut}(F_n)$. In [**140**], Wada describes several other possibilities. In [**132**], V. Shpilrain shows, using the above mentioned method, that some of these homomorphisms are faithful:

Proposition 5.1.9. — *The following maps induce embeddings of B_n into $\mathrm{Aut}(F_n)$:*
(i) $\widehat{\sigma}_i(x_i) = x_i^p x_{i+1} x_i^{-p}$, $\widehat{\sigma}_i(x_{i+1}) = x_i$, $\widehat{\sigma}_i(x_k) = x_k$ *for* $k \neq i, i+1$, *with fixed* $p \neq 0$;
(ii) $\widehat{\sigma}_i(x_i) = x_i x_{i+1}^{-1} x_i^{-1}$, $\widehat{\sigma}_i(x_{i+1}) = x_i$, $\widehat{\sigma}_i(x_k) = x_k$ *for* $k \neq i, i+1$;
(iii) $\widehat{\sigma}_i(x_i) = x_i^2 x_{i+1}$, $\widehat{\sigma}_i(x_{i+1}) = x_{i+1}^{-1} x_i^{-1} x_{i+1}$, $\widehat{\sigma}_i(x_k) = x_k$ *for* $k \neq i, i+1$.

For instance, one can check that, in case (iii), the image of a braid β admitting at least one σ-positive representative is an automorphism $\widehat{\beta}$ such that $\widehat{\beta}(x_1)$ begins with x_1^2. Notice that each embedding argument gives a new proof of Property **A**. In particular, the argument in case (iii) gives an especially short proof of this property.

5.1.4. Characterization of the braid ordering. — We proved above an implication, namely that β admitting a σ_1-positive representative causes $\widehat{\beta}(x_1)$ to end with x_1^{-1}. We shall see now that the converse implication is also true (provided Property $\mathbf{C}_\infty$ is known).

Lemma 5.1.10. — *For $k \neq i, i+1$ and $e = \pm 1$, the automorphism $\widehat{\sigma}_i$ maps $S(x_k^e)$ into itself.*

The easy proof is analogous to that of Lemma 5.1.6.

Lemma 5.1.11. — *The automorphism $\widehat{\sigma}_i$ maps $S(x_{i+1})$ into $S(x_i) \cup S(x_{i+1}) \cup S(x_{i+1}^{-1})$, and $S(x_{i+1}^{-1})$ into $S(x_i^{-1})$.*

Proof. — Let us consider an arbitrary element of $S(x_{i+1})$, say ux_{i+1} with $u \notin S(x_{i+1}^{-1})$. Then we have $\widehat{\sigma}_i(ux_{i+1}) = \mathrm{red}\big(\widehat{\sigma}_i(u)x_i\big)$. Assume $\widehat{\sigma}_i(ux_{i+1}) \notin S(x_i)$. This means that the final letter x_i is cancelled by some letter x_i^{-1} in $\widehat{\sigma}_i(u)$. This letter comes either from some x_{i+1}^{-1} or from some $x_i^{\pm 1}$ in $\widehat{\sigma}_i(u)$. As previously, we consider the possible cases.

The first case is $u = u_1 x_{i+1}^{-1} u_2$. We have now

$$\widehat{\sigma}_i(ux_{i+1}) = \mathrm{red}\big(\widehat{\sigma}_i(u_1) x_i^{-1} \widehat{\sigma}_i(u_2) x_i\big),$$

and the hypothesis is $\mathrm{red}(\widehat{\sigma}_i(u_2)) = \varepsilon$. This implies $u_2 = \varepsilon$, and, therefore, $u \in S(x_{i+1}^{-1})$, a contradiction.

The second case is $u = u_1 x_i^e u_2$ with $e = \pm 1$ and $u_1 \notin S(x_i^{-e})$. We find

$$\widehat{\sigma}_i(ux_{i+1}) = \mathrm{red}\big(\widehat{\sigma}_i(u_1)x_i x_{i+1}^e x_i^{-1}\widehat{\sigma}_i(u_2)x_i\big),$$

and the hypothesis is $\mathrm{red}(\widehat{\sigma}_i(u_2)) = \varepsilon$. This implies $u_2 = \varepsilon$, hence $u = u_1 x_i^e$. Thus we have

$$\widehat{\sigma}_i(ux_{i+1}) = \mathrm{red}\big(\widehat{\sigma}_i(u_1)x_i x_{i+1}^e\big).$$

Assume that the final letter x_{i+1}^e vanishes in the reduction. Then x_{i+1}^e vanishes with some letter x_{i+1}^{-e} that necessarily comes from some letter x_i^{-e} in $\widehat{\sigma}_i(u_1)$. So there must exist a decomposition $u_1 = u_1' x_i^{-e} u_1''$, giving

$$\widehat{\sigma}_i(ux_{i+1}) = \mathrm{red}\big(\widehat{\sigma}_i(u_1')x_i x_{i+1}^{-e} x_i^{-1}\widehat{\sigma}_i(u_1'')x_i\big)$$

with $\mathrm{red}\big(x_i^{-1}\widehat{\sigma}_i(u_1'')x_i\big) = \varepsilon$. As above, this implies $u_1'' = \varepsilon$, and, therefore, $u_1 \in S(x_i^{-e})$, a contradiction. So $\widehat{\sigma}_i(u) \in S(x_{i+1}^e)$ is the only possibility.

The argument for the image of $S(x_{i+1}^{-1})$ is similar. □

Finally, using the fact that the sets $S(x_i^{\pm 1})$ form a partition of $F_\infty \setminus \{1\}$ and applying the previous lemmas, we see that the only possibilities for the images under the inverse automorphisms $\widehat{\sigma}_i^{-1}$ are as follows:

Lemma 5.1.12. — *The automorphism $\widehat{\sigma}_i^{-1}$ maps $S(x_k^e)$ into itself for $k \neq i$, $i+1$ and $e = \pm 1$; it maps $S(x_i)$ to $S(x_{i+1})$, $S(x_i^{-1})$ to $S(x_i) \cup S(x_i^{-1}) \cup S(x_{i+1}^{-1})$, and both $S(x_{i+1})$ and $S(x_{i+1}^{-1})$ to $S(x_{i+1})$.*

Then gathering the results, we obtain:

Proposition 5.1.13. — *Let β be an arbitrary braid.*
(i) *If β admits a σ_1-positive representative braid word, then the word $\widehat{\beta}(x_1)$—a freely reduced word by definition—ends with x_1^{-1};*
(ii) *If β admits a σ_1-free representative braid word, then the word $\widehat{\beta}(x_1)$ is x_1;*
(iii) *If β admits a σ_1-negative representative braid word, then the word $\widehat{\beta}(x_1)$ ends with $x_k^{\pm 1}$ for some k with $k \geqslant 2$.*

Proof. — By construction, $\widehat{\beta}(x_1) = x_1$ is true if β admits a a σ_1-free representative. If β admits a σ_1-positive representative, we have seen in Proposition 5.1.6 that $\widehat{\beta}(x_1)$ lies in $S(x_1^{-1})$. Assume finally that β admits a σ_1-negative representative. Thus β can be expressed as $\beta_1\sigma_1^{-1}\mathrm{sh}(\beta_2)$, where β_1 admits a representative containing no σ_1. Then $\mathrm{sh}(\widehat{\beta_2})$ maps x_1 to itself, hence $(\sigma_1^{-1}\mathrm{sh}(\beta_2))\widehat{\ }$ maps x_1 to $x_2^{-1}x_1x_2$, an element of $S(x_2)$, hence of $\bigcup_{k\geqslant 2} S(x_k^{\pm 1})$. Then $\widehat{\sigma}_1^{-1}$ and all $\widehat{\sigma}_k^{\pm 1}$ with $k \geqslant 2$ map $\bigcup_{k\geqslant 2} S(x_k^{\pm 1})$ into itself: indeed, $\widehat{\sigma}_1$ is the only automorphism in the considered family that possibly maps an element of $S(x_k^{\pm 1})$ with $k \geqslant 2$ into $S(x_1^{\pm 1})$. □

Applying the shift operation, we obtain similarly the following more general result:

Proposition 5.1.14. — *Let β be an arbitrary braid.*
(i) *If β admits a σ-positive braid representative, then there exists i such that $\widehat{\beta}(x_j)$ is equal to x_j for $j < i$, and $\widehat{\beta}(x_i)$ ends with x_i^{-1};*
(ii) *If β admits a σ-negative braid representative, then there exists i such that $\widehat{\beta}(x_j)$ is equal to x_j for $j < i$, and $\widehat{\beta}(x_i)$ ends with $x_k^{\pm 1}$ for some k with $k \geqslant i+1$.*

Assuming Property $\mathbf{C}_\infty$, every braid admits a representative that is σ_1-positive, σ_1-negative, or σ_1-free, and we deduce that the previous implications actually are equivalences.

Corollary 5.1.15. — *Let β be an arbitrary braid.*
(i) *The braid β admits a σ_1-positive representative braid word if and only if the word $\widehat{\beta}(x_1)$ ends with x_1^{-1};*
(ii) *The braid β admits a σ_1-free representative braid word if and only if the word $\widehat{\beta}(x_1)$ is x_1;*
(iii) *The braid β admits a σ_1-negative representative braid word if and only if the word $\widehat{\beta}(x_1)$ ends with $x_k^{\pm 1}$ for some k with $k \geqslant 2$;*
(iv) *The braid β admits a σ-positive braid representative if and only if there exists i such that $\widehat{\beta}(x_j)$ is equal to x_j for $j < i$, and $\widehat{\beta}(x_i)$ ends with x_i^{-1};*
(v) *The braid β admits a σ-negative braid representative if and only if there exists i such that $\widehat{\beta}(x_j)$ is equal to x_j for $j < i$, and $\widehat{\beta}(x_i)$ ends with $x_k^{\pm 1}$ for some k with $k \geqslant i+1$.*

We obtain in this way the fifth characterization of the braid ordering $<$ mentioned in Introduction:

Corollary 5.1.16. — *Let β_1, β_2 be any braids. Then $\beta_1 < \beta_2$ is true if and only if, for some i, the automorphism associated with $\beta_1^{-1}\beta_2$ maps x_j to x_j for $j < i$, and it maps x_i to a word that ends with x_i^{-1}.*

We also obtain a new braid comparison algorithm with exponential complexity: if order to decide whether $\beta > 1$ is true, we compute the reduced word $\widehat{\beta}(x_1)$: if it ends with x_1^{-1}, we deduce $1 < \beta$; if it ends with $x_k^{\pm}$ for some $k \geqslant 2$, we deduce $1 > \beta$. Otherwise, we know that $\widehat{\beta}(x_1)$ must be x_1. In this case, we compute $\widehat{\beta}(x_2)$: if the latter word ends with x_2^{-1}, we deduce $1 < \beta$; if it ends with $x_k^{\pm}$ for some $k \geqslant 3$, we deduce $1 > \beta$. Otherwise, $\widehat{\beta}(x_2)$ must be x_2, and we continue similarly with $\widehat{\beta}(x_3)$. By construction, if β is specified using some braid word of length ℓ, the lengths of the words $\widehat{\beta}(x_k)$ are bounded by 3^ℓ, and they can be computed iteratively in a number of steps of the same order.

5.1.5. Extension to general Artin groups. — In [**89**], Larue extends the argument given above to prove not only Property $\mathbf{A}$, but, more generally, Property $\mathbf{A}_i$, *i.e.*, that a braid word that contains at least one letter σ_i and no letter σ_i^{-1} is never

trivial—an extension which we have seen in Chapter 2 can also be deduced from Property **A** directly.

In [**133**], Hervé Sibert uses a similar argument to prove the counterpart of Property **A** in every Artin group (Subsection 1.1.4):

Proposition 5.1.17. — *Assume that G is an Artin group, and s is any one of the standard generators of G (those mentioned in the standard presentation). Then an element of G that can be representated by a word containing at least one letter s and no letter s^{-1} cannot be* 1.

In the case of type A, *i.e.*, of braid groups, and of type B, and, more generally, of products of such groups, Property **C** is true, and we can obtain a linear ordering. In all other cases, the counterpart of Property **C** is false, and we obtain partial orderings that are not linear.

5.2. From an automorphism back to a braid

In [**88**], Larue gave the first published proof of Property **C** in its full strength (not just Property $\mathbf{C}_\infty$), taking $\mathbf{C}_\infty$ as an hypothesis. His proof is very similar to the topological argument given three years later in the paper [**60**], whose authors were unaware of his work. Larue's argument, which takes place entirely in the group-theoretic setting, and uses almost no plane topology, is quite intricate; we shall only sketch it here, and refer the reader to Chapter 6 for a somewhat more detailed account of the argument using curve diagrams (which, in addition, does not require to take Property $\mathbf{C}_\infty$ as an hypothesis). The core of the problem is to be able to recover a braid from its image in $\mathrm{Aut}(F_n)$. The main result is the following:

Proposition 5.2.1. — *Assume that w is an n-strand braid word of length ℓ such that $\widehat{w}(x_1)$ ends with x_1^{-1}. Then w is equivalent to a σ_1-positive n-strand braid word w' of length at most $\ell + n^2 3^\ell/4$.*

In order to prove Proposition 5.2.1, we shall use, for convenience, a slightly different generating set for the free group F_n, namely y_i with $i = 1, \dots, n$, where y_i is defined by $y_i = x_i^{-1} \cdots x_1^{-1}$. We have the following relation between a reduced word u with letters $x_i^{\pm 1}$ and a reduced word u' with letters $y_i^{\pm 1}$ representing the same element of F_n: u ends with x_i if and only if u' ends in y_i^{-1} (for $i = 1, \dots, n$), and u equals x_1^{-1} if and only if u' equals y_1^{-1}. Moreover, the action by the braid σ_i sends the generator y_i to $y_{i-1} y_i^{-1} y_{i+1}$, and leaves all other generators fixed (with the convention $y_0 = 1$).

Geometrically, we can identify F_n with the fundamental group of the n times punctured disk D_n included in $\mathbb{C}$ whose base point is the point -1 of $\mathbb{C}$, and whose punctures are contained in the real line (see the beginning of Chapter 1). Under this identification, the generators y_i ($i = 1, \dots, n$) can be represented by n simple loops in D_n which are disjoint except in the basepoint, and where the curve corresponding

to y_i winds once around the i leftmost punctures in an anticlockwise sense (as shown in Figure 1.3). Moreover, if an element y of $\pi_1(D_n)$ is given by a reduced word of length ℓ in the letters $y_1, \ldots, y_n$, then y can be represented by a path in D_n which has exactly ℓ intersections with the horizontal axis $\mathbb{R}$ to the right of the first puncture, and this is the minimum intersection number among all paths representing y.

It is easy to prove inductively that for any braid word w of length ℓ, the element $\widehat{w}(y_1)$ of F_n is given by a word of length at most 3^ℓ in the generators $y_1, \ldots, y_n$.

Our aim is to write the braid represented by w^{-1} as a σ_1-negative braid word, and more precisely as a product of braid words of a particular form. For $1 \leqslant r < s < t \leqslant n$, we define

$$w_{r,s,t} = (\sigma_s \sigma_{s-1} \ldots \sigma_{r+1})(\sigma_{s+1} \ldots \sigma_{r+2}) \ \ldots \ (\sigma_{t-1} \ldots \sigma_{t-s+r}),$$

and, for $0 \leqslant r < s < t \leqslant n$, we define

$$w'_{r,s,t} = (\sigma_s^{-1} \sigma_{s-1}^{-1} \ldots \sigma_{r+1}^{-1})(\sigma_{s+1}^{-1} \ldots \sigma_{r+2}^{-1}) \ \ldots \ (\sigma_{t-1}^{-1} \ldots \sigma_{t-s+r}^{-1}).$$

We observe that the braid words $w_{r,s,t}$ and $w'_{r,s,t}$ are σ_1-free for $r \geqslant 1$, and that $w'_{r,s,t}$ is σ_1-negative for $r = 0$. Note that the length of $w_{r,s,t}$ and of $w'_{r,s,t}$ is at most $n^2/4$. We intend to find a product of such words equivalent to w^{-1}.

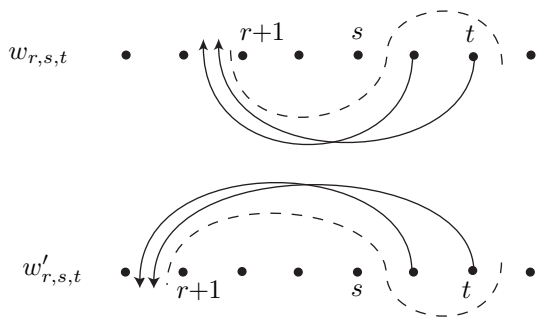


FIGURE 5.1. The braid words $w_{r,s,t}$ and $w'_{r,s,t}$

In Figure 5.1 the braids represented by $w_{r,s,t}$ and $w'_{r,s,t}$ are sketched. The meaning of the dashed arcs will be explained shortly. The key lemma is now

Lemma 5.2.2. — *If w is a braid word such that $\widehat{w}(y_1)$ ends with y_i^{-1} for some i in $\{1, \ldots, n\}$, then there exists a braid word v equal to some $w_{r,s,t}$ or $w'_{r,s,t}$ as above such that*

(i) *the length of the reduced form of $\widehat{v}(\widehat{w}(y_1))$ is at least 2 less than that of $\widehat{w}(y_1)$,*

(ii) *the reduced form of $\widehat{v}(\widehat{w}(y_1))$ is either y_1, or it still ends with y_i^{-1} for some i in $\{1, \ldots, n\}$.*

Proof. — We define an $(r \downarrow s \uparrow t)$-arc (with $1 \leqslant r < s < t \leqslant n$) and a $(r \uparrow s \downarrow t)$-arc (with $0 \leqslant r < s < t \leqslant n$) of $\widehat{w}(y_1)$ to be a (non-oriented) subarc of the loop $\widehat{w}(y_1)$ whose extremal points lie between the rth and the $(r+1)$-st, and between the tth and $(t+1)$st puncture, respectively, in the horizontal axis, and which intersect the axis precisely once more, namely between the sth and $(s+1)$st puncure, as indicated in Figure 5.1.

One can prove that if $\widehat{w}(y_1)$ ends with y_i^{-1}, then $\widehat{w}(y_1)$ contains an $(r \downarrow s \uparrow t)$-arc or an $(r \uparrow s \downarrow t)$-arc, for some r, s, t in the legal range, as a subarc.

Now it is not too hard to see—and should be intuitively clear from Figure 5.1—that applying the braid (represented by) $w_{r,s,t}$ or $w'_{r,s,t}$, respectively, to such a loop $\widehat{w}(y_1)$ reduces the number of intersections of the loop with the horizontal line by at least two. In other words, with $v = w_{r,s,t}$ or $v = w'_{r,s,t}$ we have $|\widehat{v}(\widehat{w}(y_1))| \leqslant |\widehat{w}(y_1)| - 2$.

Moreover, one can show that the resulting element $\widehat{v}(\widehat{w}(y_1))$ of F_n still ends with y_i^{-1}, unless $\widehat{v}\widehat{w}$ is a σ_1-free braid, in which case we have $\widehat{v}(\widehat{w}(y_1)) = y_1$. □

Now Proposition 5.2.1 can be deduced from Lemma 5.2.2, by the following induction argument.

Proof of Proposition 5.2.1. — For the given braid word w, the element $\widehat{w}(y_1)$ of F_n has length at most 3^ℓ. Now we can apply a sequence of braids, each with at most $n^2/4$ crossings, to this element, reducing its length by at least two in each step. This process yields a σ_1-negative braid word w_1 of length at most $n^2 3^\ell/8$ such that the braid represented by $w_1 w$ is σ_1-free. To be precise: the word $w_1 w$ is of length at most $|w_1| + \ell$, which in turn is less than or equal to $\ell + n^2 3^\ell/8$; it may well contain letters $\sigma_1^{\pm 1}$, but it represents a σ_1-free braid. It follows that there exists another braid word w_2, say, equivalent to $w_1 w$, whose length is also bounded by $\ell + n^2 3^\ell/8$, but which does not contain any letter $\sigma_1^{\pm 1}$ (this can be seen e.g. by removing the first string). Now w and $w_1^{-1} w_2$ represent the same braid, and the latter word is σ_1-positive and of length at most $\ell + n^2 3^\ell/4$. □

Applying Proposition 5.2.1, we deduce, as in [**89**]

Corollary 5.2.3. — *Property* $\mathbf{C}_\infty$ *implies Property* $\mathbf{C}$*, i.e., assuming that every* n*-strand braid word* w *is equivalent to some braid word* w' *that is* σ_1*-positive,* σ_1*-negative, or* σ_1*-free, we can require in addition that* w' *is an* n*-strand braid word as is* w*.*

Proof. — Assume that w is an n-strand braid word. By Property $\mathbf{C}_\infty$, there exists some braid word w', equivalent to w, that is σ_1-positive, σ_1-negative, or σ_1-free. In the first case, Proposition 5.1.6 implies that the automorphism $\widehat{\beta}$ maps x_1 to some word that ends with x_1^{-1}, and, then, Proposition 5.2.1 implies that there exists another n-strand braid word w'' equivalent to w and w' that is σ_1-positive. Applying the same argument to w^{-1} gives the result when w' is σ_1-negative. Finally, if w' is σ_1-free, then we directly obtain that the braid represented by w belong to the image of the shift

mapping, and we obtain a σ_1-free n-strand representative for it by removing the first strand in w, in an obvious sense. □

The previous result is a little frustrating, as we would like to obtain a self-contained proof of Property **C**, one that does not take Property $\mathbf{C}_\infty$ as an hypothesis. Using the geometrical interpretation of the action of B_n on F_n, we can argue as follows.

***Corollary 5.2.4* (Property C).** — *Every braid in B_n admits a representative n-strand braid word that is σ_1-positive, σ_1-negative, or σ_1-free*

Proof. — Let β be a braid, and let w be a representative of β. If $\widehat{w}(x_1)$ ends with x_1^{-1}, then, by Proposition 5.2.1, w is equivalent to a σ_1-positive n-strand braid word. A second possibility is $\widehat{w}(x_1) = x_1$—we shall deal with this case later. If neither of these two possibilies is satisfied then the image of the curve $\widehat{w}(x_1)$ under the reflection of D_n in the horizontal axis ends with x_1^{-1}. This means that the image of the braid β under the reflection–automorphism of B_n, which sends every generator to its inverse, has a σ_1-positive representative. This implies that β itself has a σ_1-negative representative.

Finally, for $\widehat{w}(x_1) = x_1$, we consider the n-strand geometric braid obtained by removing the first strand of the geometric braid specified by w and replacing it by a strand which stays to the left of the other strands all along its length. The new braid is described by some σ_1-free braid word w', and, as it has the same curve diagram as w (see Chapter 6 below), the braid words w and w' are equivalent. This proves that β has a σ_1-free representative. □

We end this chapter by touching on one more application of the action of the braid group on the free group, which will be discussed in more detail in Chapter 7. One can equip the free group with a linear ordering which is not invariant under the action of F_n on itself by left multiplication, but which *is* invariant under the action of B_n on F_n. Thus any element of F_n on which B_n acts freely gives rise to a linear left-invariant ordering of B_n, by pulling back the ordering on the orbit. In this way we obtain many different orderings of B_n. This approach, which was studied in detail by Jonathon Funk in [**65**], is by and large equivalent to the Nielsen-Thurston type approach described in Chapter 7; however, it has the advantage of being very explicit, and since it avoids any geometrical tools, it generalizes neatly to braid groups with a countable infinity of strands. This approach brings into focus the connections between braid orderings and notions from topos theory.

CHAPTER 6

CURVE DIAGRAMS

We turn to a very different construction of the braid ordering based on a topological approach. The braid group B_n is isomorphic to the mapping class group, *i.e.*, the group of isotopy classes of self-homeomorphisms, of a disk with n punctures (Proposition 1.1.3). Looking at the images of the main diameter of such a disk under self-homeomorphisms naturally leads to a linear ordering of the mapping class group, which, as we shall see, coincides with the σ-ordering. This construction works not only for the braid groups, but more generally for mapping class groups of compact surfaces with nonempty boundary. The approach described in this chapter was developed in [**60**] and [**129**], and, like the self-distributivity approach, it is complete in the sense that it leads to proofs of all of Properties **A**, **C**, and **S**.

6.1. Mapping class groups and curve diagrams

Throughout this section, we shall denote D^2 the unit disk in $\mathbb{C}$ with centre 0, and D_n the same disk with n uniformly spaced points in the real axis $\mathbb{R} \cap D^2$ marked as distinguished points. These points will be called the *punctures*, and denoted $P_1, \dots, P_n$, from left to right. We also introduce notation for some diagrams in D_n: we write $e_0, \dots, e_n$ for the $n+1$ horizontal open line segments $(-1, P_1)$, (P_1, P_2), $\dots$, (P_{n-1}, P_n), $(P_n, +1)$, respectively, and E for their union (see Figure 6.1(a)).

The aim of this section is to formalize the following idea which was hinted at in the introduction: in order to specify the isotopy class of a homeomorphism $\varphi\colon D_n \to D_n$ which permutes the punctures, it suffices to specify the image under φ of the horizontal line E in D_n. This reduces the problem of understanding isotopy classes of homeomorphisms of a surface to the more intuitively accessible one of understanding isotopy classes of curves on the surface.

Definition 6.1.1. — The *curve diagram* of a homeomorphism $\varphi\colon D_n \to D_n$ is the image of E under φ. Two curve diagrams are *isotopic* if there exists an isotopy of D_n

which deforms one diagram into the other and leaves $P_1, \dots P_n$ and the boundary of D^2 fixed during the isotopy.

Proposition 6.1.2. — *Two homeomorphisms $\varphi, \psi\colon D_n \to D_n$ are isotopic if and only if their associated curve diagrams are isotopic (isotopies in both instances are to be fixed on the punctures and the boundary).*

We skip the proof, which is standard. Proposition 6.1.2 says that there is a natural correspondence between elements of $\mathcal{MCG}(D_n)$ (which in turn is naturally isomorphic to the braid group B_n) and isotopy classes of curve diagrams in D_n. Some examples are given in Figure 6.1.

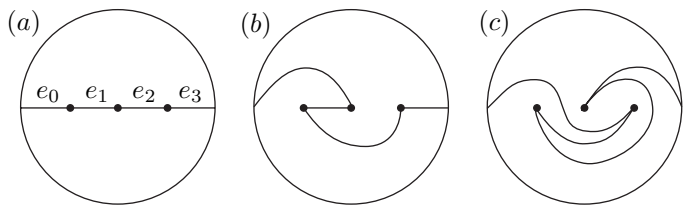


FIGURE 6.1. Examples of curve diagrams: the curve diagrams of the trivial braid, of σ_1 and of $\sigma_2^{-1}\sigma_1$.

We shall now give a very informal explanation how geometric braids give rise to curve diagrams, and conversely, given the curve diagram of a braid β in B_n, how to recover a geometric braid representing β.

What follows is a recapitulation of the proof of Proposition 1.1.3. Let β be a geometric n-braid, sitting in the cylinder $[0,1] \times D^2$, whose n strands are starting at the puncture points of $\{0\} \times D^2$ and ending at the puncture points of $\{1\} \times D^2$. A good way to imagine the curve diagram of a homeomorphism φ associated with β is as follows. We think of the diagram E in D_n as consisting of $n+1$ segments of rubber band in D_n. If we now slide the diagram E from the $\{1\}$-level in the cylinder back to the 0-level, then the braid β corresponds to a dance of the n puncture points in the disk D^2. During the dance, the rubber bands get stretched and deformed, and the resulting picture in $\{0\} \times D^2$ is the curve diagram of the homeomorphism φ.

Conversely, given the curve diagram of some braid β in B_n, *i.e.*, in $\mathcal{MCG}(D_n)$, but not β itself, we want to get a geometric n-braid representing β. Our procedure for doing this will play a key role in what follows. We can place the curve diagram in the $\{0\} \times D^2$-level of the cylinder. If we now authorize the puncture points to move, then due to the elastic force of the rubber bands the diagram will untangle and become the straight horizontal line in D_n. While this is happening, we can slide D_n along the cylinder into the $\{1\} \times D^2$-level. Then the punctures will trace out the geometric braid β, which we have thus recovered from the curve diagram.

6.2. A braid ordering using curve diagrams

The aim of this section is to construct an ordering $<_{\mathrm{CD}}$ of B_n using curve diagrams. We shall see that this ordering coincides with the ordering which is the subject of this book.

6.2.1. Definition of the order. — Given two braids, or equivalently two elements of $\mathcal{MCG}(D_n)$ represented by two homeomorphisms φ and ψ, we want to define which one is larger. The first step is to superimpose the curve diagrams of φ and ψ in D_n. A priori, these diagrams may intersect each other unnecessarily; for instance, φ-curves and ψ-curves may have points of tangency, or may intersect in infinitely many isolated points, or may simply enclose bigons—see Figure 6.2.

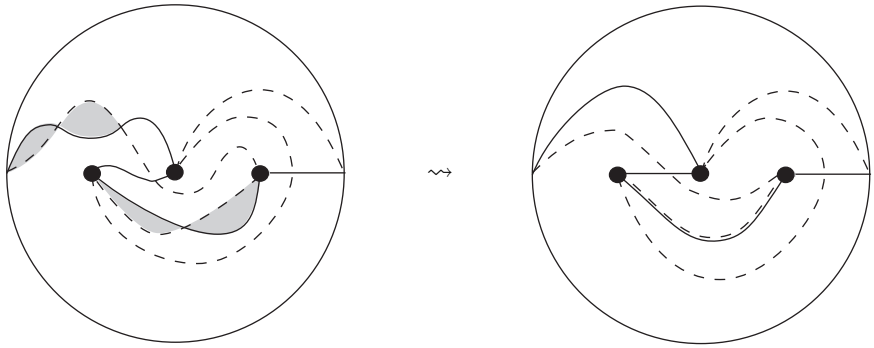


FIGURE 6.2. Removing four bigons in the curve diagrams for σ_1 and $\sigma_2^{-1}\sigma_1$. The diagrams in the resulting picture are tight with respect to each other

So the second step is to pull the curve diagrams tight. That means, we want to isotope the two curve diagrams independently until they are in a position relative to each other where the number of intersection points is as small as possible. The only exception is: if one of the $n+1$ φ-curves coincides exactly with one of the ψ-curves, we count this as -1 intersection point.

One possible way to find such an tight positioning of the curves is to start with two diagrams with finite intersection number, and successively decrease the intersection number by removing bigons until no more bigons are left. It is a nonobvious fact [**60**] that pairs of curve diagrams without bigons are always tight (a similar but more general statement will be proved in Section 8.1 below). This opens up a theoretically neat (although not practical) method for tightening curve diagrams: one equips D_n with a hyperbolic metric, in which the n puncture points are cusps, and lets the $2(n+1)$ curves of the diagrams flow into geodesics. Since geodesics never form bigons, this yields a tight pair of diagrams. This observation is the foundation of the techniques which will be introduced in Chapter 7. The crucial fact for our purposes is that such a tight positioning of two curve diagrams is essentially unique:

Proposition 6.2.1. — *Suppose that C and C' are two curve diagrams, which are isotopic, and are both in tight position with respect to another curve diagram C''. Then there exists an isotopy of the disk which fixes the diagram C'' setwise, fixes the punctures and the boundary of the disk, and transforms C' into C''.*

A proof, which is elementary, is given in [**60**]; very similar results were, however, well-known before this paper. We refer the reader to Section 8.1.2 for an argument for this Proposition, but in a slightly different setting.

As a third step, we want to define a relation between curve diagrams. Suppose that C_1 and C_2 are two non-isotopic curve diagrams. Let us imagine that we sit down at the point -1 of D_n, and walk along the curves of C_1. For some initial period of time the curves of C_2 may exactly coincide with those of C_1. At some moment, however, either at -1 or at one of the puncture points, the diagrams C_1 and C_2 will diverge (otherwise we would have $\beta_1 = \beta_2$). At this divergence point, C_2 will either set out into to upper component of $D_n \setminus C_1$ (the one containing the point $(0,1)$), or into the lower component (which contains $(0,-1)$). In the first case we say β_1 *goes more to the left than* β_2, and in the second that β_1 *goes more to the right than* β_2.

Definition 6.2.2. — A relation $<_{\text{CD}}$ on B_n is defined as follows. If β_1 and β_2 are two distinct braids, then we superimpose their curve diagrams, C_1 and C_2 say, pull them tight, and define that $\beta_1 >_{\text{CD}} \beta_2$ is true if C_1 goes more to the left than C_2 and $\beta_1 <_{\text{CD}} \beta_2$ is true if C_2 goes more to the left than C_1.

Lemma 6.2.3. — *The relation $<_{\text{CD}}$ is a (strict) linear ordering on B_n, and it is left-invariant.*

Proof. — In order to see that $<_{\text{CD}}$ is transitive, we need to know that any three curve diagrams can be drawn in D_n in such a way that they are mutually tight. In [**60**], this is proved using the so-called triple reduction lemma. Alternatively, we can use hyperbolic geometry: if all the curves of all three curve diagrams consist of geodesics of the same hyperbolic structure on D_n, then the three diagrams will be mutually tight.

Seeing that the ordering $<_{\text{CD}}$ is left-invariant is easy: suppose that $\varphi_1, \varphi_2, \varphi$ are homeomorphisms of D_n representing three braids β_1, β_2, β, and that the φ_1-diagram goes more to the left than the φ_2-diagram. If we apply φ to the curve diagrams of φ_1 and φ_2, we obtain curve diagrams for $\varphi\varphi_1$ and $\varphi\varphi_2$, respectively. Since we have applied the same homeomorphism φ to the two diagrams, their relative position is unchanged: the diagram of $\varphi\varphi_1$ will still be to the left of the diagram of $\varphi\varphi_2$. This proves that $\beta_1 >_{\text{CD}} \beta_2$ implies $\beta\beta_1 >_{\text{CD}} \beta\beta_2$. □

6.2.2. A proof of Property A. — The main result of this section is

Proposition 6.2.4. — *Suppose that the braid β admits at least one σ-positive representative braid word. Then we have $\beta >_{\mathrm{CD}} 1$. Similarly, a braid β with a σ-negative representative satisfies $\beta <_{\mathrm{CD}} 1$.*

Proof. — Let w be a σ-positive braid word representing β, with curve diagram C. For simplicity, let us first assume that w is in fact σ_1-positive; then w is of the form $w_0\sigma_1 w_1\sigma_1 w_2 \dots \sigma_1 w_k$, where the words w_i contain no letter $\sigma_1^{\pm 1}$. We shall now consider the curve diagrams of various braids, with particular attention to their first curves, *i.e.*, those starting at -1. Our aim is to prove that an initial segment of the first curve of C lies in the upper half of D_n, which implies $\beta >_{\mathrm{CD}} 1$.

The first curve of the curve diagram of the braid represented by w_k coincides with the one of the trivial braid: it's just a horizontal line segment, because the first strand does not cross any other strands. Acting on this curve diagram by σ_1 yields the curve diagram of $\sigma_1 w_k$; its first curve, *i.e.*, the image of the arc e_0 under $\sigma_1 w_k$, is an arc in the upper half of D_n connecting -1 to the second-leftmost puncture P_2. In particular, this first curve has the following property: its first intersection with the vertical line through the leftmost puncture P_1 lies in the upper half of D_n. Now we observe that successively applying any sequence of braids $\sigma_2^{\pm 1}, \dots, \sigma_{n-1}^{\pm 1}$ and σ_1 (but not σ_1^{-1}) cannot change this property. In particular, the curve diagram C of the braid word w has this property, which implies that an initial segment of C lies in the upper half of D_n.

This completes the proof for σ_1-positive braid words. For braids that admit a representative braid word containing at least one σ_i, and no σ_i^{-1} and no $\sigma_j^{\pm 1}$ with $j < i$, the argument is similar, guaranteeing that the first $i-1$ curves of the curve diagram will be horizontal line segments, and the ith will set off into the upper half of D_n. □

The curve diagram of a σ-positive word has to diverge (to the left, in fact) from the trivial curve diagram, so we deduce:

***Corollary 6.2.5* (Property A)**. — *A braid that admits at least one σ_1-positive representative braid word is not trivial.*

6.2.3. A proof of Property C. — The aim of this section is to use curve-diagram techniques in order to (re)prove the comparison property (Property **C**).

***Proposition 6.2.6* (Property C)**. — *Every braid in B_n admits a representative braid word that is σ-positive, σ-negative, or empty,*

Proof. — (Figure 6.3) This is a sketch—for details see [**60**]. The strategy is to prove that any braid β with $\beta >_{\mathrm{CD}} 1$, *i.e.*, whose curve diagram first diverges from the trivial curve diagram into the upper half of D_n, has a σ-positive representative.

For simplicity we will assume the curve diagram of β diverges *immediately* to the left, and conclude that β admits a σ-positive representative. The other possibility

is that the diagram coincides with e_0 and only diverges later. In this case β admits σ_1-free representatives and a similar argument (which we leave to the reader) will show that at least one such representatives is σ-positive.

We recall the technique which was introduced in the proof of Proposition 1.1.3: we can reconstruct a braid from its curve diagram by placing the curve diagram in the 0-level of a cylinder $[0,1] \times D^2$, untangling the diagram while sliding the disk into the 1-level, and observing the trace of the puncture points under this movement. We shall apply this technique, but very carefully avoid using the letter σ_1^{-1}.

So let C be a curve diagram which is tight with respect to the trivial curve diagram (the straight horizontal line), and whose first curve sets out immediately into the upper half of D_n. Our aim is to isotope C into the trivial diagram. At every moment in time we can imagine a vertical line through the leftmost puncture. What we have to avoid during the isotopy is that any puncture ever hits this line *below* the leftmost puncture, in order to turn itself into the new leftmost puncture—punctures are only allowed to travel into the leftmost position by hitting the *upper* half of the vertical line through the current leftmost puncture (this condition will turn to be equivalent to the requirement that the resulting geometric braid is described by a σ_1-positive word).

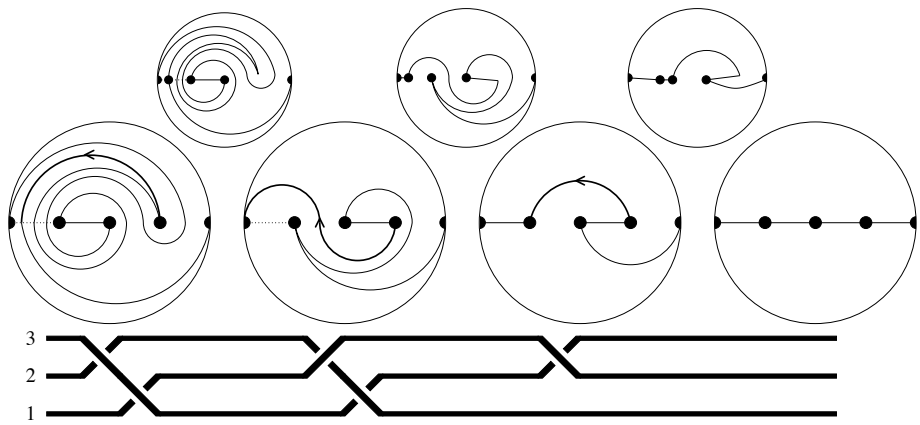


FIGURE 6.3. How to find a σ-positive representative of a braid word, given the curve diagram. To untangle it, we have to alternately slide punctures along useful arcs and tighten the diagram. The useful arcs are drawn in bold line.

The idea for actually finding such an untangling-movement is as follows: we recall that e_0 denotes the horizontal line segment from -1 to the leftmost puncture. The curve diagram C intersects e_0 in a finite number of points (one of which is the point -1). We define a *useful arc* to be a subarc of one of the curves of C which

starts at one of the intersection points with e_0, has an initial segment which lies in the upper half of D_n, does not intersect e_0 again, and terminates at any puncture (P say) other than the leftmost one. A simple geometric argument guarantees that in any curve diagram whose first curve sets off into the upper half of D_n such a useful arc can indeed be found. Now we are ready to start untangling the diagram C: we slide the puncture P back along the useful arc, all the way to its starting point in e_0. Note that during this movement the puncture P will slide exactly once (near the end of its voyage) *over* the current leftmost puncture to achieve itself the new leftmost position, but it never slides *under* the leftmost puncture. This completes the first stage of our untangling process. The curve diagram thus deformed may well fail to be tight with respect to the trivial curve diagram, so we pull it tight, and obtain a new curve diagram C'.

Now we claim that the first curve of C' still sets off into the upper half of D_n, or it might possibly coincide with e_0. This can be proved by another easy geometric argument (and it is actually very plausible if one thinks of the curve diagram as being realised by rubber bands). In the first case, we can iterate our construction (that is, find a useful arc in C', etc.). In the second case we can simply untangle the curve diagram C' without any special care.

This process has to terminate with the trivial curve diagram—indeed, with respect to a suitably defined notion of complexity of curve diagrams, C' is simpler than C, and the complexity decreases further in each successive iteration. □

We remark that the method introduced in this section for finding a σ-positive representative of a positive braid (by first constructing its curve diagram and then carefully undoing this diagram again) is algorithmically very inefficient: the length of the resulting σ-positive braid word depends, in general, exponentially on the length of the input braid word. For more comments on the algorithmic aspect see Section 6.3.2.

Proposition 6.2.4 tells us that the σ-ordering of Chapter 1 is included in the ordering $<_{\mathrm{CD}}$. Since both $<_C$ D and the σ-ordering are linear orderings, they must therefore coincide. In other words, we have established the sixth equivalence mentioned in the Introduction:

Proposition 6.2.7. — *For all braids β_1, β_2, the relation $\beta_1 < \beta_2$ is true if and only if we have $\beta_1 <_{\mathrm{CD}} \beta_2$, i.e., the standardized curve diagram associated with β_1 first diverges from that associated with β_2 towards the left.*

6.2.4. A proof of Property S. — It is possible to give a proof of Property **S** using the machinery of this chapter, and this was done in [**142**]. We shall do it here by using the notion of a Dehn half-twist.

Definition 6.2.8. — Let α be a simple arc connecting two distinct points, say P and Q, in the plane (or, more generally, in an arbitrary oriented surface). A *Dehn half-twist* around α is a homeomoprhism φ of the plane such that φ is the identity outside a small neighbourhood of α, it flips α, and it screws clockwise a small neighbourhood U of α as shown in Figure 6.4.

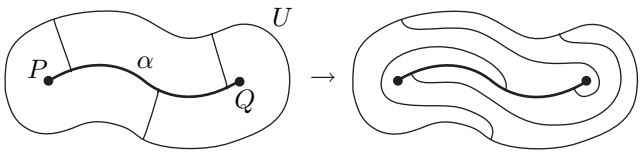


FIGURE 6.4. Dehn half-twist around the arc α

For instance, the element of $\mathcal{MCG}(D_n)$ defined by a Dehn half-twist around the segment e_i is the image of the braid σ_i under the isomorphism of B_n to $\mathcal{MCG}(D_n)$.

Let φ be a Dehn half-twist around an arc α, and γ be a curve intersecting α transversally at some point, say R. The image of γ under φ behaves as follows. When x goes along γ, the image $\varphi(x)$ also goes along γ until coming close enough to R, then it turns to the left and goes to the endpoint of α, keeping close to α, then it turns clockwise around that point, goes to the other endpont of α, crossing α at a point "symmetric" to R, turns counterclockwise around it, then it returns into a small neighbourhood of R keeping close to α, and, finally, it continues moving along γ.

With the previous notion at hand, we can prove:

Proposition 6.2.9. — *For every braid β, we have $1 <_{\mathrm{CD}} \beta\sigma_i\beta^{-1}$.*

Proof. — Let ψ be the homeomorphism of D_n corresponding to β, and let α be the arc $\psi(e_i)$. Then a Dehn half-twist φ around α represents the braid $\beta\sigma_i\beta^{-1}$. So, by definition of $<_{\mathrm{CD}}$, the inequality $1 <_{\mathrm{CD}} \beta\sigma_i\beta^{-1}$ follows from the above observation that $\varphi(E)$ diverges from E to the left provided we make sure that φ can be chosen so that $\varphi(E)$ is tight with respect to E.

Now, we can assume without loss of generality that α is tight with respect to the main diameter E. Then we claim that the Dehn half-twist φ can be chosen so that there is no bigon bounded by E and $\varphi(E)$. Indeed, in addition to coinciding parts of E and $\varphi(E)$, there are only unavoidable crossing points of E and $\varphi(E)$ that appear near crossing points of E and α. Between two such points the curve $\varphi(E)$ goes parallel to α. Therefore, assuming the existence of a bigon bounded by E and $\varphi(E)$ would imply the existence of a bigon bounded by E and $\varphi(\alpha)$. It is not hard to see now that only a small perturbation is needed for $\varphi(E)$ to become tight with respect to E. So the proof is complete. □

By Proposition 6.2.7, the σ-ordering of the braids coincides with the $<_{\mathrm{CD}}$-ordering, so we immediately deduce

***Corollary 6.2.10* (Property S).** — *Every braid of the form $\beta\sigma_i\beta^{-1}$ admits a σ-positive representative braid word.*

Let us still mention that more general results in the above direction can be proved using the refinements of the curve-diagram technique introduced in Chapter 7 (see Section 7.3).

6.3. Generalizations

In this section, we want to briefly review some further developments concerning two topics: firstly, attempts to find left-invariant orderings of other types of mapping class groups and braid groups, and secondly the efficiency of various algorithms for deciding which of two elements in a braid group is the larger.

6.3.1. Other mapping class groups and braid groups. — In Section 1.1 we defined the mapping class group $\mathcal{MCG}(\mathcal{S},\mathcal{P})$ of any compact surface $\mathcal{S}$ relative to a finite set of punctures $\mathcal{P}$. Closely related is the *n-strand braid group* $B_n(\mathcal{S})$ of a surface $\mathcal{S}$. It can be defined as the fundamental group of the configuration space of n unlabelled points in $\mathcal{S}$. More geometrically, we can fix arbitrarily n distinguished points $P_1,\dots,P_n$ in the interior of $\mathcal{S}$. Then $B_n(\mathcal{S})$ is the group of isotopy classes of braids in $[0,1]\times\mathcal{S}$, where each strand starts at one of the points $\{0\}\times P_i$ and ends at one of the points $\{1\}\times P_j$. For instance, we have $B_1(\mathcal{S})=\pi_1(\mathcal{S})$ for every surface $\mathcal{S}$. It is a simple fact [**7**] that for all compact surfaces except S^2, S^2 with one puncture, S^2 with two punctures, the torus, and the Klein bottle, $B_n(\mathcal{S})$ is in a natural way a subgroup of $\mathcal{MCG}(\mathcal{S},\{P_1,\dots,P_n\})$.

Proposition 6.3.1. — *Let $\mathcal{S}$ be any compact surface with or without punctures, orientable or nonorientable, but necessarily with $\partial\mathcal{S}\neq\emptyset$. Then $\mathcal{MCG}(\mathcal{S})$ is left-orderable.*

A proof of this fact appeared in [**129**]. It uses a simple generalisation of the curve diagram construction from Section 6.2. Since subgroups of left-orderable groups are also left-orderable, we have in particular that the braid groups in such surfaces are left-orderable. However, no interesting analogue of the notion of σ-positivity is known in this case.

The situation is much more subtle if $\mathcal{S}$ is a compact surface without boundary. The mapping class groups of such surfaces have torsion, and are consequently not left-orderable. However, it is still unknown whether or not their n-strand braid groups ($n\geqslant 2$) are left-orderable. The *pure* braid groups in $\mathcal{S}$ are bi-orderable by [**70**].

6.3.2. Relaxation algorithms. — As mentioned there, the algorithm given in Section 6.2 for finding a σ-positive representative for a braid β with $\beta > 1$ is very inefficient, in that its computational complexity depends exponentially on the length of the input braid word. We describe now a potentially much more efficient method, called the relaxation algorithm. This method, which has probably been rediscovered independently several times, is studied in some detail in [**143**]. It has two versions, called the standard- and the σ-consistent version, both of which take as input an arbitrary braid word w in the letters $\sigma_i^{\pm 1}$, and output a canonical word w' representing the same element of B_n as w. The only difference between the two algorithms is that, contrary to the standard version, the σ-consistent version outputs only σ-positive and σ-negative words. To explain the algorithm, we have to consider a larger set of generators of the braid group, called *semicircular braids.* These are braids which can be represented by a word of the form $\sigma_i^e \sigma_{i+1}^e \dots \sigma_{j-1}^e$ (with $1 \leqslant i < j \leqslant n$ and $e = \pm 1$), or of the form $\sigma_{i-1}^e \dots \sigma_{j+1}^e \sigma_j^e$ (with $n \geqslant i > j \geqslant 1$). Geometrically, this corresponds to a movement of the ith strand of the braid, by a semicircular movement, into the position between the jth and the $(j+1)$st strand. Now the standard relaxation algorithm works just like the algorithm given in the proof of Proposition 6.2.6: one calculates the curve diagram C of w, finds a braid word w'' whose action untangles the curve diagram, and the inverse of w'' is then our output braid word w'. However, whereas the braid word w'' in the proof of Proposition 6.2.6 was obtained as a product of slides along useful arcs, our word w'' will now be a product of semicircular braids. In each iteration, the algorithm considers all possible semicircular braids, and compares the results of their actions on C. It then chooses the semicircular braid β that reduces a (suitably defined) complexity of the curve diagram by as much as possible. Finally, the diagram C is replaced by the simplified curve diagram $\beta \cdot C$, and the next iteration can start. The algorithm stops when the trivial curve diagram E is reached.

The σ-consistent algorithm works in the same way, except that in all iterations the choice is made not among *all* semicircular braids, but a subset, so as to avoid using the letter σ_1 in the case of a braid β satisfying $\beta > 1$, and the letter σ_1^{-1} in the case of a braid β satisfying $\beta < 1$

This idea can be turned into a computer program [**143**] which appears to work extremely efficiently—the computation time required seems to grow quadratically with the length of the input braid, and this not only on average, but even in all worst-case scenarios. There is currently no explanation for this surprising behaviour. More details about this open problem and related algorithmic questions can be found in Chapter 10.

CHAPTER 7

HYPERBOLIC GEOMETRY

This chapter contains a summary of the approach developed in [**131**], and of Jonathon Funk's interpretation from [**65**]. It is based on an observation by William Thurston that, using hyperbolic geometry, one can let the group B_n act on the real line by order-preserving automorphisms. This approach arises very naturally from Nielsen–Thurston theory, and has the advantage of providing many possible left-invariant orderings of B_n.

In Section 7.1, we construct a natural action of the braid group on the real line, and explain how this action gives rise to many different orderings of B_n. These orderings are classified in Section 7.2. In Section 7.3 we prove that all orderings of B_n arising from our action on $\mathbb{R}$ have Property **S**. Finally in Section 7.4 we show how the methods in this chapter can be made effective through the use of cutting sequences.

Throughout this chapter, groups act on the left.

7.1. Uncountably many orderings of the braid group

In this section we first explain the general correspondence between group actions on the real line and orderings of groups. Then we construct a very natural B_n-action on $\mathbb{R}$. Finally we look at some examples of orderings arising from this construction. These examples introduce all the essential ideas for the systematic classification of orderings in the next section.

7.1.1. Orderability and group actions on $\mathbb{R}$. — It can be shown that a non-trivial group is left-orderable if and only if it acts on some linearly ordered set by order-preserving bijections in such a way that only 1 acts trivially. For our purposes, actions on the real line suffice, according to the following "folk-theorem", whose proof appears in [**68**].

Proposition 7.1.1. — *A countable group G is left-orderable if and only if G acts on $\mathbb{R}$ by orientation-preserving homeomorphisms in such a way that only 1_G acts by the identity map, i.e., if and only if there exists a monomorphism from G to* $\mathrm{Homeo}_+(\mathbb{R})$.

Proof. — (Sketch) Given an action of G on $\mathbb{R}$, we define an ordering $<_G$ as follows. We fix an enumeration $q_1, q_2, \dots$ of the rationals. Consider distinct elements g_1, g_2 in G. We let i be the smallest integer satisfying $g_1(q_i) \neq g_2(q_i)$. Now we define $g_1 <_G g_2$ if $g_1(q_i) < g_2(q_i)$ and $g_1 >_G g_2$ if $g_1(q_i) > g_2(q_i)$. This is a *linear* ordering: If we have $g_1(q_i) = g_2(q_i)$ for all i in $\mathbb{N}$, then g_1 and g_2 act by the same homeomorphism, because $\mathbb{Q}$ is dense in $\mathbb{R}$. The ordering is left-invariant, for $g_1(q_i) > g_2(q_i)$ implies $h \circ g_1(q_i) > h \circ g_2(q_i)$ for every h in G, because h acts orientation preservingly and i is also the least integer satisfying $h \circ g_1(q_i) \neq h \circ g_2(q_i)$.

Conversely, if $<$ is a left-invariant ordering of G, then one can construct an action of G on $\mathbb{R}$ as follows: one can construct an order-preserving injection $I\colon G \hookrightarrow \mathbb{R}$ such that all the image points are isolated. This uses the existence of a countable order-dense *discrete* subset of $\mathbb{R}$ with no first or last element; think of the midpoints of the deleted intervals in construction of the Cantor set, for example. Since G acts on itself in an order-preserving way by right multiplication, this yields an order-preserving action of G on the image of I. Finally, we interpolate, in order to extend this to an action on all of $\mathbb{R}$. □

We shall be primarily interested in a special case of the above result:

Definition 7.1.2. — Suppose G acts on $\mathbb{R}$, and suppose in addition that there exists a point x of $\mathbb{R}$ with $\mathrm{Stab}_G(x) = \{1\}$, *i.e.*, the points in the orbit of x under the G-action are in one-to-one correspondence with the elements of G. Then we can define a left-invariant ordering $<_x$ of G so that $g_1 <_x g_2$ holds if and only if we have $g_1(x) < g_2(x)$.

Note that for a different point y in $\mathbb{R}$, the same construction method and the same G-action may yield a very different ordering; we shall see many instances of this behaviour.

We need to set up some more notation. Two left-invariant orderings $<_1$ and $<_2$ of a group G are said to be *conjugate* if there exists an h in G such that $g_1 <_1 g_2$ is equivalent to $g_1h <_2 g_2h$. For instance, if a group G acts on $\mathbb{R}$, and some x in $\mathbb{R}$ satisfies $\mathrm{Stab}_G(x) = \{1\}$, then, for any h in G, the orderings of G induced by x and $h(x)$ are conjugate.

We recall that an ordering $<$ of G is said to be *dense* if, for any two elements g_1, g_2 of G satisfying $g_1 < g_2$, there exist infinitely many elements h in G with $g_1 < h < g_2$. By contrast, an ordering is *discrete* if every element has a well-defined predecessor and successor, *i.e.*, if for $g \in G$ there exists a largest g_1 and a smallest g_2 with $g_1 < g < g_2$. It is an easy exercise to show that any left-ordering of a group is either

dense or discrete. For example, the σ-ordering of B_n considered throughout this text is discrete: the predecessor of the braid β is $\beta\sigma_{n-1}^{-1}$, and its successor is $\beta\sigma_{n-1}$.

A *convex subgroup* of an ordered group $(G, <)$ is a subgroup H of G such that, for h_1, h_2 in H, and g in G, the relation $h_1 < g < h_2$ implies $g \in H$. For instance, the subgroup of B_n generated by $\sigma_2, \dots, \sigma_{n-1}$ is convex in the σ-ordering.

7.1.2. An action of B_n on $\mathbb{R}$. — We shall think of the braid group B_n as the mapping class group of a disk with n punctures: $B_n = \mathcal{MCG}(D_n)$. It seems certain that all the results in this chapter can be generalized to mapping class groups of any compact surface with nonempty boundary, with or without punctures. However, this has never been done explicitly.

The following result is essentially due to Nielsen [**112, 113**]. Together with Proposition 7.1.1, it implies that B_n is left-orderable.

Proposition 7.1.3. — *There is a natural action by orientation-preserving homeomorphisms of the braid group B_n on a topological space which is homeomorphic to the real line. Moreover, only the trivial braid acts as the identity homeomorphism.*

Proof. — The n-punctured disk D_n can be equipped with a hyperbolic metric. (Actually, there is a choice of a whole $\mathbb{R}^{2n-3}$-family—the Teichmüller space of D_n—of isotopy classes of hyperbolic metrics, but the choice is irrelevant for the $\mathbb{R}$-action we are aiming for.) The metric on D_n lifts to a metric on the universal cover $\widetilde{D_n}$ of D_n, and then $\widetilde{D_n}$ can be isometrically embedded in the hyperbolic plane $\mathbb{H}^2$. We can compactify $\mathbb{H}^2$ by adding a circle at infinity $S^1_\infty = \partial\mathbb{H}^2$. We can then go on to compactify $\widetilde{D_n}$ by attaching its limit points on S^1_∞. The resulting space is homeomorphic to a closed disk, and by abuse of notation we shall still denote it $\widetilde{D_n}$.

The circle $\partial\widetilde{D_n}$ has two types of points: firstly the points at infinity, which form a Cantor set in the circle, and secondly their complement, $\partial\widetilde{D_n} \cap \mathbb{H}^2 = p^{-1}(\partial D_n)$ (where p denotes the covering projection), which consists of a countable number of open arcs (see Figure 7.1). We now choose, once and for all, a basepoint $*$ in one of these arcs. We are finally ready to define the real line on which B_n shall act: it is $\partial\widetilde{D_n} \setminus \{*\}$.

Next we consider an element β of B_n, represented by some homeomorphism φ from D_n to D_n. There are infinitely many ways in which φ can be lifted to a homeomorphism $\widetilde{\varphi}$ of the universal cover (as many as there are elements of $\pi_1(D_n)$), but we have one preferred choice: we demand that $\widetilde{\varphi}$ fix our basepoint $*$. Thus we have constructed an action of B_n on the real line $\partial\widetilde{D_n} \setminus \{*\}$.

To see that this action is well-defined, we suppose that ψ is a different representative of the same element β of B_n. Then ψ is related to φ by a homotopy which is fixed on ∂D_n. This homotopy lifts to a homotopy between $\widetilde{\varphi}$ and $\widetilde{\psi}$ which is fixed on the arcs $\partial\widetilde{D_n} \cap \mathbb{H}^2$. Since these arcs are dense in $\partial\widetilde{D_n}$, the homeomorphisms φ and ψ lift to exactly the same action on $\partial\widetilde{D_n}$.

Finally, in order to prove the second statement, we suppose that a homeomorphism φ acts trivially on $\partial\widetilde{D_n}$. Then, in particular, this homeomorphism fixes all liftings of the basepoint $*$ of D_n, and thus induces the identity-homeomorphism on $\pi_1(D_n)$. This implies that φ represents the trivial element of $\mathcal{MCG}(D_n)$. $\square$

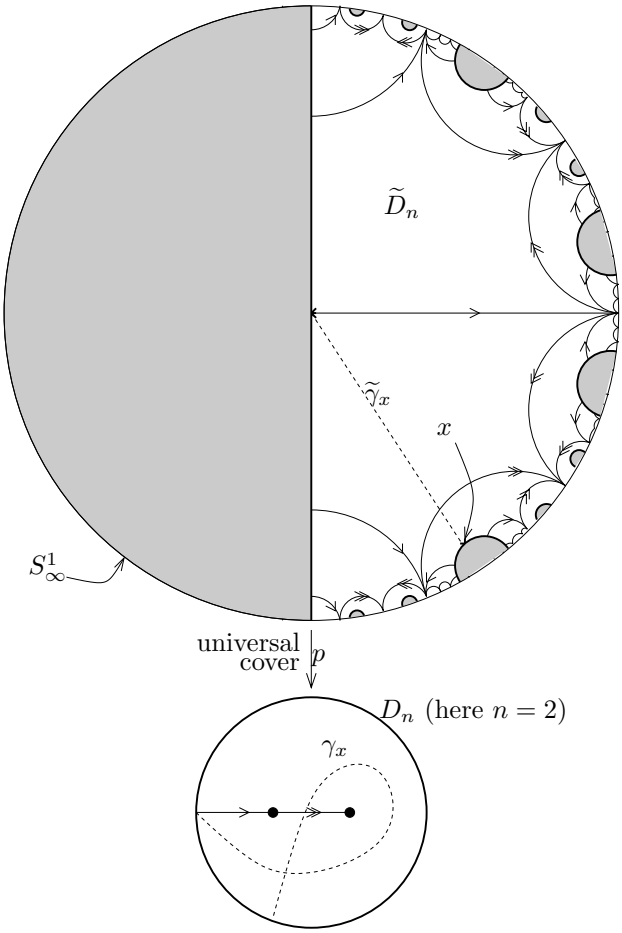


FIGURE 7.1. The disk D_n, and its universal cover $\widetilde{D_n}$ embedded in $\mathbb{H}^2$ (indicated as the unshaded part)

Our definition of the action may look rather abstract. We shall actually think of it in the following way: if x is a point in $\partial\widetilde{D_n} \setminus \{*\}$, then a path $\widetilde{\gamma}_x$ (typically without self-intersections) in $\widetilde{D_n}$ from the basepoint $*$ to x projects to a path γ_x (typically with many self-intersections) in D_n which starts at $p(*)$ in ∂D_n. Now a

homeomorphism φ acts on this path γ_x (just like it acted on curve diagrams), and the lifting of $\varphi(\gamma_x)$ is a path in $\widetilde{D_n}$ from $*$ to some point in $\partial\widetilde{D_n}$. This is the point that we define to be $\varphi(x)$. Note that, if x was a point in S^1_∞, then all the paths mentioned were necessarily infinite.

7.1.3. Examples of orderings induced by the action on $\mathbb{R}$. — Our aim now is to classify the order types of B_n arising from points x with $\mathrm{Stab}(x) = \{1\}$ in the above action of B_n on $\partial\widetilde{D_n} \setminus \{*\}$, *i.e.*, on $\mathbb{R}$. For instance, we shall prove that for $n \geqslant 3$ there are uncountably many conjugacy classes of dense orderings of B_n, but only a finite number (for which a formula will be given) of conjugacy classes of discrete orders arising from our current approach. However, the main classification results 7.2.5, and 7.2.8 are quite technical; in order to get some intuition for the most important phenomena that can occur, we start, in this section, by giving some examples.

The first set of examples is meant to illustrate how some fundamentally different order types can indeed result from our construction. We shall see examples of both dense and discrete orderings, which are all mutually non-conjugate. The second set of examples has the opposite aim: showing that many very different-looking geodesics give rise to the same orderings of B_n—thus the amount of different order types arising from our construction is quite limited, and this makes a classification possible.

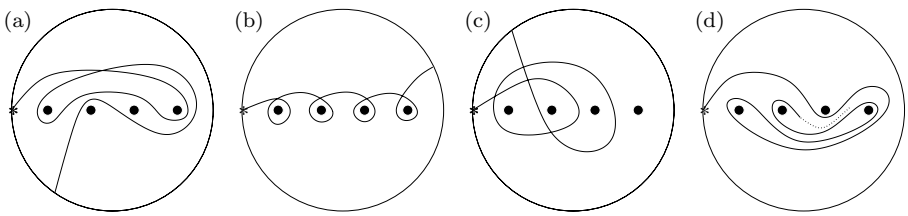


FIGURE 7.2. Four curves $\gamma_a, \gamma_b, \gamma_c$, and γ_d in D_4, for four points in $\partial\widetilde{D_4} \setminus \{*\}$. Figure (d) is meant to represent an infinite geodesic ray without self-intersections, whose closure would look like a generic geodesic lamination of D_4.

The first set of examples begins with the four curves shown in Figure 7.2. As explained above, each of the curves in this figure represents a point in $\partial\widetilde{D_4} \setminus \{*\}$, namely the endpoint of the lifting of the curve to $\widetilde{D_4}$ which starts at the point $*$ in $\partial\widetilde{D_4}$. We shall denote these four points a, b, c and d.

Lemma 7.1.4. — *The point a in $\partial\widetilde{D_4} \setminus \{*\}$ has a nontrivial stabilizer under the B_4-action, and thus does not induce a linear ordering of B_4.*

Proof. — We observe that the first and the fourth puncture of D_4 are in the same path component of $D_4 \setminus \gamma_a$. The element $\sigma_1^{-1}\sigma_2^{-1}\sigma_3\sigma_2\sigma_1$ of B_4 has support in this path

component and interchanges the first and fourth puncture. It is thus a nontrivial element of B_4 which acts trivially on the point a. □

More generally, we observe that a necessary condition for a geodesic to induce a linear ordering of B_4 is that it separates the punctures of D_n, in the sense that no two punctures of D_n should be in the same connected component of D_n with the geodesic removed. We shall see later on that this necessary condition is almost sufficient (a geodesic which separates the punctures may fail to induce a total ordering if it lies in a complementary component of the stable geodesic lamination of some pseudo-Anosov homeomorphism, but apart from this class of exceptions the necessary condition is sufficient).

From now on, we shall only be interested in geodesics which separate the punctures, like for instance the geodesics γ_b, γ_c and γ_d. It is a fact which for the moment we shall not prove, that the three points b, c and d do indeed have trivial stabilizer (with the proviso that for γ_c we need an extra technical hypothesis, which is generically satisfied).

Thus, using Definition 7.1.2, each of these three points b, c and d in $\partial\widetilde{D_4} \setminus \{*\}$ induce a linear ordering of B_4, denoted $<_b$, $<_c$ and $<_d$, respectively. For instance, saying that a homeomorphism φ of D_4 represents an element of $\mathcal{MCG}(D_4)$ which is larger than 1 in the $<_b$–ordering means that the endpoint of the lifting of $\varphi(\gamma_b)$ is to the left of the endpoint of the lifting of γ_b, as seen from the basepoint of $\widetilde{D_4}$.

We now claim that the orderings $<_b$, $<_c$ and $<_d$ of B_4 are fundamentally different from each other:

Proposition 7.1.5. — (i) *The orderings $<_b$ and $<_c$ are discrete, whereas $<_d$ is dense. In particular, $<_d$ is not conjugate to either $<_b$ or $<_c$.*

(ii) *The ordering $<_b$ has a convex subgroup $\langle \sigma_2, \sigma_3 \rangle$ isomorphic to B_3. By contrast, $<_c$ has a convex subgroup $\langle \sigma_1, \sigma_3 \rangle$ isomorphic to $\mathbb{Z}^2$. The orderings $<_b$ and $<_c$ are not conjugate.*

Proof. — (Sketch) We shall only give some plausibility arguments. The general philosophy is the following: if for some path γ and for some homeomorphism φ of D_n the support of φ is disjoint from a fairly long initial segment of γ, then the element of B_n represented by φ should be fairly close to 1_{B_n} in the ordering induced by γ.

For instance, in example (c) we consider the liftings to $\widetilde{D_n}$ with initial point $*$ of the curves γ_c and $\sigma_1^k(\gamma_c)$ (for any k in $\mathbb{Z} \setminus \{0\}$). These liftings coincide for quite a long initial segment of γ_c, namely along the segment drawn in thin line in Figure 7.3(b); by contrast, only a relatively short tail consisting of the lifting of the latter parts of γ_c, drawn bold in Figure 7.3(b), is at all affected by the σ_1-action. This means that the endpoint c of $\partial\widetilde{D_n} \setminus \{*\}$ of the lifting of γ_c is very close to the endpoint $\sigma_1^k(c)$ of the lifting of $\sigma_1^k(\gamma_c)$. By contrast, under the action of σ_3, the immobile initial segment of γ_c is much shorter, and the moving tail, bold in Figure 7.3(c), is much longer.

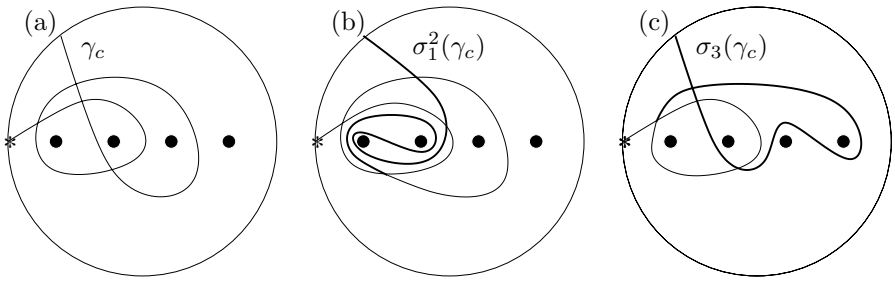


FIGURE 7.3. We have $1 < \sigma_1^2 < \sigma_3$ in the ordering associated to the geodesic γ_c

Indeed, it is easy to see that $\sigma_2(c)$ is further removed from c than $\sigma_3^k(c)$ for every k in $\mathbb{N}$. This is saying that, in $<_c$, we have $\sigma_1^k < \sigma_3$ for all k in $\mathbb{N}$ (and in a similar fashion one can see that $\sigma_3^\ell < \sigma_2$ holds for all ℓ in $\mathbb{N}$).

For $<_b$, by contrast, we have that $\sigma_3^k < \sigma_2^\ell < \sigma_1$ for all k, ℓ in $\mathbb{N}$. Indeed, the action of σ_3^k can be said to affect only the tail of γ_c after the second self-intersection, whereas the effect of a σ_2^ℓ-action is already noticable after the first self-intersection, and σ_1^m right from the start, for k, ℓ, m in $\mathbb{Z}$.

Finally for $<_d$, one can find homeomorphisms of D_n representing nontrivial elements of B_n which leave arbitrarily long initial segments of γ_d untouched, which implies that the order $<_d$ is dense.

The only part of the proposition that remains to be proven is that $<_b$ and $<_c$ are not conjugate. Let us suppose, for a contradiction, that they are. Since the convex subgroups of a group are linearly ordered by inclusion, this would imply that there exists a subgroup Γ of B_4 which is conjugate in B_4 to the subgroup $\langle \sigma_1, \sigma_3 \rangle$, such that $\langle \sigma_2, \sigma_3 \rangle$ either is included in Γ or includes Γ. The first case is impossible since an abelian group cannot contain a nonabelian one. The second case is impossible because the subgroup $\langle \sigma_2, \sigma_3 \rangle$ has support in a subdisk of D_4 which contains only three of the punctures, whereas any subset of D_4 supporting the subgroup $\langle \sigma_1, \sigma_3 \rangle$ (or any subgroup conjugate to this one) necessarily contains all four punctures. $\square$

To summarize, the ordering depends critically on how the geodesic cuts up the disk: if two of the punctures are first separated from the two others, and then each of the pairs is split, we obtain an ordering which is qualitatively different from the ordering induced by a geodesic which splits off one puncture at a time.

There is, however, one more effect that can lead to two different geodesics inducing different orderings, even though the two sequences of more and more finely chopped subdisks of D_n induced by cutting along the geodesics are topologically indistinguishable. This effect is illustrated in Figure 7.4.

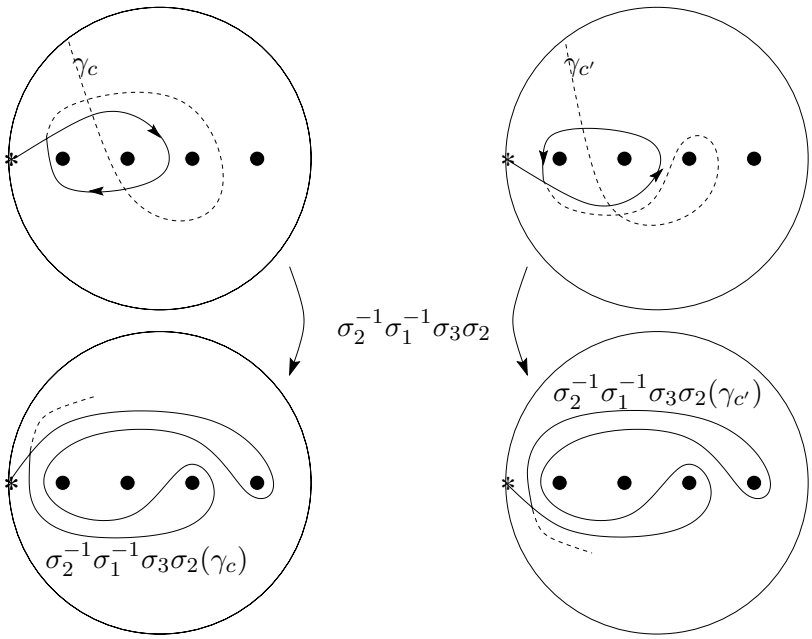


FIGURE 7.4. The geodesics γ_c and $\gamma_{c'}$ cut the disk D_n into pieces in very similar ways, yet $\sigma_2^{-1}\sigma_1^{-1}\sigma_3\sigma_2 >_c 1$ whereas $\sigma_2^{-1}\sigma_1^{-1}\sigma_3\sigma_2 <_{c'} 1$

There are initial parts of the geodesics γ_c and $\gamma_{c'}$ (drawn in solid line) which separate the first two punctures from the third and the fourth, and the pieces that result from the two cuts are topologically indistinguishable. The only difference between the curves is that the *direction* of the cut is opposite. This difference, however, suffices to make the orderings different: we can observe immediately in the picture that $\sigma_2^{-1}\sigma_1^{-1}\sigma_3\sigma_2 >_c 1$ holds, whereas $\sigma_2^{-1}\sigma_1^{-1}\sigma_3\sigma_2 <_{c'} 1$ does.

Indeed, using results that we shall present later on, one can prove that the geodesics γ_c and $\gamma_{c'}$ (including the dotted parts) induce orderings which are not even conjugate.

This completes our first set of examples: we have seen why different geodesics can induce different orderings of B_n. We are now ready for the second set of examples. We shall see that geodesics which on superficial inspection appear to have no resemblance can give rise to the exact same orderings. This happens if they cut the disk D_n into pieces in essentially the same way.

At the same time we shall see that the σ-ordering considered throughout the book is a special case of an ordering arising from our Nielsen–Thurston type construction. This yields the seventh equivalent definition of the σ-positive ordering mentioned in

the introduction: we consider the ordering associated with the geodesic γ_b, and its obvious generalization to disks with *any* finite number of punctures. We saw earlier that $\sigma_3^k <_b \sigma_2^\ell <_b \sigma_1$ holds for all natural numbers k and ℓ. In fact, we have the following stronger result:

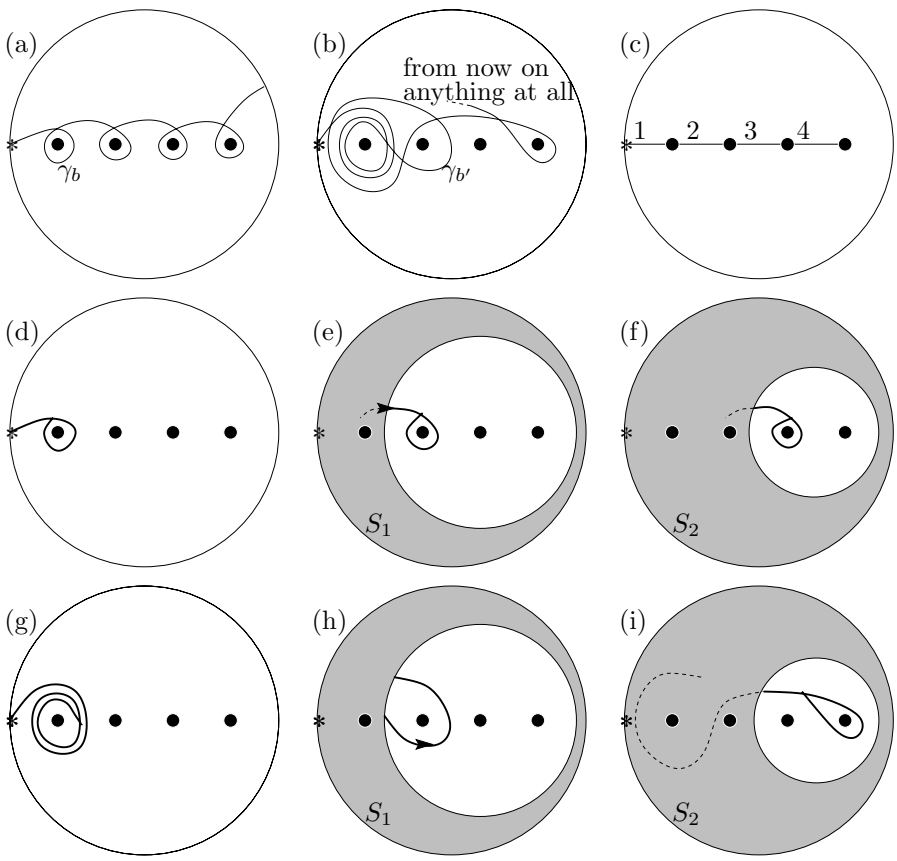


FIGURE 7.5. Two different geodesics γ_b and $\gamma_{b'}$ giving rise to the same ordering—in fact, this is exactly the σ-ordering considered thoughout this book. Figures (d)–(f) analyze γ_b, figures (g)–(i) analyze $\gamma_{b'}$

Proposition 7.1.6. — (i) *Let γ_b be the geodesics in D_n indicated in Figure 7.5(a). Then the ordering $<_b$ coincides precisely with the σ-ordering, i.e., for two braids β_1 and β_2 in B_n we have $\beta_1 < \beta_2$ if and only if the endpoint of the lifting of $\beta_1(\gamma_b)$ in is larger (as a real number) than the endpoint of the lifting of $\beta_2(\gamma_b)$.*

(ii) *The same statement holds for the geodesic $\gamma_{b'}$ in Figure 7.5(b).*

Proof. — Throughout the proof we shall refer to Figure 7.5. The proof is exactly the same for the two geodesics γ_b and $\gamma_{b'}$. In both cases, the geodesic has an initial segment up to the first self-intersection which separates the leftmost puncture from the other three. Moreover, if we pull tight the loop around the first puncture, we obtain exactly the first curve in the trivial curve diagram of Figure 7.6(c).

Considering the action of a braid β on these initial segments, we observe that the braid sends the segments further to the left if and only if it sends the first curve of the curve diagram in figure (c) to the left. In other words, the relation $\beta >_b 1$ or $\beta >_{b'} 1$ holds if and only if β admits a σ_1-positive braid word representative. (Here we are using the characterisation of the σ-ordering given in Chapter 6.) In the same way, for braids β which possess a σ_1-negative representative word, we have $\beta <_b 1$ and $\beta <_{b'} 1$. Finally, the inital segments of γ_b and $\gamma_{b'}$ are fixed up to isotopy if and only if the braid admits a σ_1-free representative word, *i.e.*, one without a letter σ_1 or σ_1^{-1}. In summary, the action on the initial segments in figures (d) or (g) already tell us whether a braid admits a σ_1-positive or a σ_1-negative representative word, and we only need to consider the later parts of the curves if the braid admits a σ_1-free representative word.

So from now on we suppose that the braid β admits a σ_1-free representative. Then the braid β is guaranteed to have a representative in the mapping class group whose support is disjoint from the region with geodesic boundary drawn shaded in figures (e) and (h). Thus it is completely irrelevant for the induced ordering what the geodesic does between the first self-intersection and the moment when it leaves the shaded region.

When the geodesic enters the region which contains the three rightmost punctures, we consider its segment up to the next self-intersection or the moment it reenters the shaded region (whichever happens first). These segments for γ_b and $\gamma_{b'}$, drawn with bold line in figures (e) and (h), both separate the second puncture from the third and fourth, and they do so in two ways which are essentially equivalent to each other. The precise sense in which the bold lines in figures (e) and (h) are equivalent is the following: if in each of the figures we add a neighbourhood of the bold line to the shaded region and then delete the bold lines themselves, we obtain two isotopic figures.

Since the braid β admits a σ_1-free representative braid word, it induces a braid on three strings, which has a representative that acts only on the non-shaded disk containing punctures number 2, 3 and 4. We can now use exactly the same argument as in the first step. The bold line is sent further to the left or right respectively, if and only if curve number 2 in the trivial curve diagram (c) is sent to the left or right, respectively. We remark that the cut is clockwise around the second puncture in diagram (e) but counterclockwise in diagram (h)—however, we observe that this

difference in orientation makes no difference for the induced ordering if only one puncture is being cut off.

If the bold lines in figures (e) and (h) as well are fixed up to isotopy, *i.e.*, if the braid β admits a representative braid word which is σ_2-free as well, then we consider the only region of D_n which can still support a nontrivial element of B_n, namely the disk with geodesic boundary containing punctures number 3 and 4, as shown in figure (f) and (i). Again, it is irrelevant for the induced ordering what the homotopy class of the geodesic is between the last point of the bold arc in the previous step and the first intersection point with the disk around the last two punctures. For instance, the part of the curve $\gamma_{b'}$ which is drawn as a dotted arc in figure (i) may be replaced by an arbitrarily complicated arc in the shaded region of figure (i) without any effect on the induced ordering.

Finally, the bold arcs in figures (f) and (i) separate the remaining two punctures. We recall that we are now considering the case where the braid β can be represented by a word σ_3^k with $k \in \mathbb{Z}$. We observe that the arcs drawn in bold face in figure (f) and (i), as well as the arc number (3) in the curve diagram (c), are sent further to the left in the case $k > 0$, are stabilized in the case $k = 0$, and are sent further to the right in the case $k < 0$. □

Remark 7.1.7. — There are orderings of B_n which cannot be obtained by our Nielsen–Thurston type construction. For instance, one can consider the exponent sum homomorphism $e\colon B_n \to \mathbb{Z}$, and define the following left-invariant order: a braid β is said larger than 1 if we have either $e(\beta) > 0$, or both $e(\beta) = 0$ and $\beta > 1$ in the σ-ordering. One can prove that this ordering is not conjugate to any ordering considered in this chapter.

7.2. The classification of orderings induced by the action on $\mathbb{R}$

In this section we state the main classification theorem for orderings arising from our action of B_n on the real line, and outline the proof. The explanation given here, using certain sequences of subsurfaces of D_n, is not quite the same as the one in [**131**], which uses curve diagrams. However, the two approaches are essentially equivalent.

Definition 7.2.1. — A geodesic γ_x in D_n *fills* the disk D_n, or is *filling*, if $D_n \setminus \gamma_x$ has no path-connected component that contains two or more of the punctures of D_n.

For instance, the geodesics in Figure 7.2(b), (c), and (d) are filling whereas the one in (a) is not. We shall only be interested in geodesics filling D_n, because we know from Lemma 7.1.4 that those that are not filling do not give rise to linear orderings.

Definition 7.2.2. — A filling geodesic γ is of *finite type* if one of the following two conditions is satisfied: either a finite initial segment of the geodesic already separates the punctures, or the geodesic *falls into a puncture*, in the sense that one of the

punctures has the property that all but a finite initial segement of γ lies in its cusp neighbourhood. A filling geodesic which is not of finite type is of *infinite type*.

For instance, the geodesics in Figure 7.2(b) and (c) are of finite type, and the one in (d) of infinite type. We stress that a finite type geodesic is not necessarily finite in length; for instance, a geodesic that falls into a puncture is infinite. Also, the geodesic $\gamma_{b'}$ in Figure 7.5(b) is of finite type, regardless of whether it terminates after finite time on ∂D_n or continues forever. The orderings induced by filling geodesics of finite and infinite type have very different properties, and we shall treat the two cases separately.

7.2.1. Finite type geodesics. — We start by looking at finite type geodesics. We are going to construct a geometrical invariant of such geodesics which contains just enough information about the geodesic in order to specify the induced ordering, but no more.

Definition 7.2.3. — A *subsurface sequence* is a finite sequence $S_0, S_1, \dots, S_{n-1}$ of open connected submanifolds of D_n satisfying $S_i \subsetneqq S_{i+1}$ for $i = 0, \dots, n-2$. Moreover, we require that S_0 is a regular neighbourhood of ∂D_n, and that S_{n-1} is D_n. We also require that, for $i > 0$, all components of $\partial \overline{S}_i$ are simple closed geodesics, one of which is ∂D_n.

We shall not indicate the submanifold S_0, which is just an annular neighbourhood of ∂D_n, in our pictures.

In the above situation, the surface S_i must be homeomorphic to a disk with $i+1$ holes (here the punctures are considered to be holes)—in particular S_{i+1} is obtained from S_i by adding one boundary component of $\overline{S}_i$ and, along this boundary component, an open surface homeomorphic to a disk with two holes. Thus if we compare the number of boundary components of $\overline{S}_i$ to those of $\overline{S}_{i+1}$, we can see three possible effects:

- (i) either $\overline{S}_{i+1}$ has one less boundary component than $\overline{S}_i$, as in the transition from (c) to (d) in Figure 7.6,
- (ii) or the number of boundary components remains constant, as in the transition from (e) to (f) in Figure 7.5,
- (iii) or it may increase by one, similar to the transition from (b) to (c) in Figure 7.6.

In case (iii), *i.e.*, when $\overline{S}_{i+1}$ has two boundary components which were not present in $\overline{S}_i$ while one boundary component of $\overline{S}_i$ has disappeared, there must be one more piece of information present, namely a transverse orientation to a geodesic segment that connects the two new boundary components of $\overline{S}_{i+1}$ inside $S_{i+1} \setminus S_i$.

Examples of subsurface sequences can be found in Figures 7.5(d)–(f), (g)–(i) and 7.6(b)–(d).

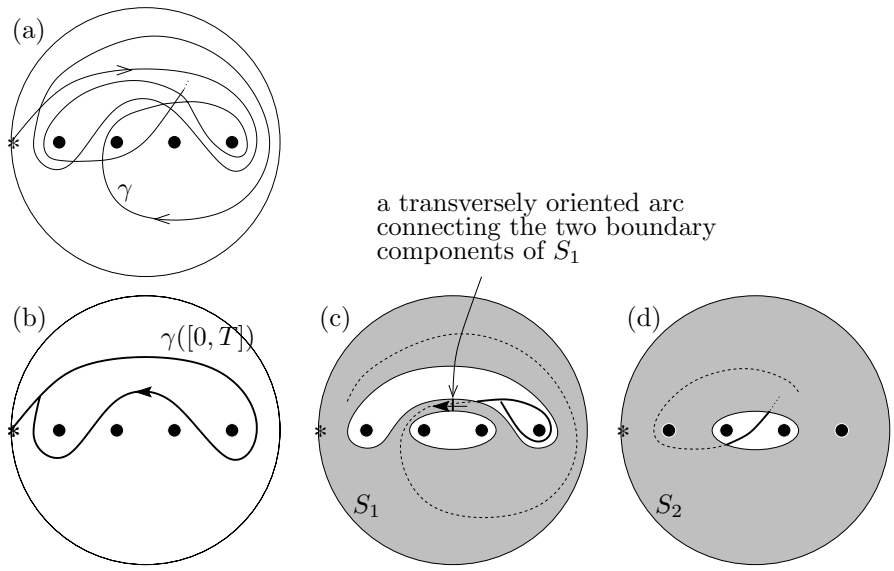


FIGURE 7.6. An example of the subsurface sequence associated to a geodesic γ

Definition 7.2.4. — Two subsurface sequences $S_1, \dots, S_{n-1}$ and $S'_1, \dots, S'_{n-1}$ are *conjugate* if there exists a homeomorphism ϕ of D_n such that $\phi(S_i)$ is isotopic to S'_i for $i = 1, \dots, n-1$.Moreover, if S_i comes equipped with a transversely oriented geodesic arc, then the same must be true for S'_i, and ϕ must carry (up to isotopy) the transversely oriented arc of S_i to the one of S'_i, preserving the transverse orientation.

Next we explain using Figure 7.6 how a finite or infinite geodesic $\gamma\colon [0, M] \to D_n$ or $\gamma\colon [0, \infty) \to D_n$ gives rise to a subsurface sequence. Let T be the smallest real number so that the initial segment $\gamma([0, T])$ of the geodesic γ has a self-intersection. Then $D_n \setminus \gamma([0, T])$ has two path components, each of which contains at least one of the punctures of D_n. We take an open regular neighbourhood $\widetilde{S}_1$ of $\partial D_n \cup \gamma([0, T])$, and define the surface S_1 to be the unique subsurface of D_n which is homotopic to $\widetilde{S}_1$ and whose closure has geodesic boundary components. Note that, if $D_n \setminus \gamma([0, T])$ has a component which contains only one puncture, then S_1 contains that entire component: see for instance Figure 7.5(e) and (h), where the boundary components of $\widetilde{S}_1$ that encircled the leftmost puncture has disappeared in S_1. Thus, if both components of $D_n \setminus \gamma([0, T])$ contain only one puncture, then we are in case (i) above, if one component contains one puncture and the other contains more, then we are in case (ii), and if both of components of $D_n \setminus \gamma([0, T])$ contain two or more punctures of D_n, then we are in case (iii). In the last case, S_1 has two boundary components,

which we connect by a geodesic segment in S_1—this segment is unique up to isotopy. We equip this geodesic segment with the transverse orientation induced from the orientation of $\gamma([0,T])$. For instance, in Figure 7.6 the orientation of the curve in (b) induces the transverse orientation on the geodesic arc in (c). This completes the construction of S_1.

The construction of the whole sequence of subsurfaces is now inductive. Let T' be the smallest positive real number with the property that $\gamma(T') \notin S_1$. The segment $\gamma((T,T'))$ of γ can be completely ignored. Furthermore, the point $\gamma(T')$ lies in the boundary of one of the components of $D_n \setminus S_1$, which are punctured disks. We look at the shortest initial segment of $\gamma([T',\infty))$ which either shows a self-intersection or intersects $\overline{S}_1$ again. We add an open neighbourhood of this segment to S_1, and homotope the resulting subsurface so that its closure has geodesic boundary. This homotopy may kill some boundary components of $\overline{S}_1$, like in the previous paragraph. This procedure yields the surface S_2, *etc.*

Now, we have the following classification result for the finite type orderings:

Theorem 7.2.5. — (i) *If γ_x is a filling geodesic of finite type, then it induces a linear ordering, i.e.,* $\mathrm{Stab}(x)$ *is trivial. Two filling geodesics of finite type γ_x and γ_y induce the same ordering if and only if they give rise to the same subsurface sequence. Moreover, γ_x and γ_y induce conjugate orderings if and only if their subsurface sequences are conjugate.*

(ii) *All orderings arising from filling geodesics of finite type are discrete.*

Proof. — (Sketch, for full details see [**131**]) The essential observation is that one can reconstruct the ordering $<_x$ from the subsurface sequence associated to γ_x. Indeed, suppose that β is an element of B_n, and we want to decide which of $\beta >_x 1$, $\beta = 1$, or $\beta <_x 1$ is true. We start by choosing the maximal integer i such that β has a representative homeomorphism with support disjoint from S_i. Then we let δ be an embedded geodesic segment in S_{i+1} with endpoints in $\partial\overline{S}_i$, which cannot be homotoped into $\overline{S}_i$: up to a movement of the endpoints in $\partial\overline{S}_i$, there is only one such segment. Now we restrict our attention to the component D' of $D_n \setminus S_i$ which contains $S_{i+1} \setminus S_i$. The segment δ cuts the punctured disk D' into two components. If one of these components contains only one of the punctures of D_n, then we choose an orientation for δ arbitrarily. If both components contain at least two punctures, then our geodesic segment δ intersects exactly once the transversely oriented geodesic segment in S_{i+1} that came with the subsurface sequence. This transverse orientation induces an orientation of δ. Now we recall that a representative of β acts on the punctured disk D' in a boundary-fixing way. We compare the geodesic δ with a geodesic representative of $\beta(\delta)$ in D', and more precisely we compare their initial segments with respect to the chosen orientation on δ. The key observation, which is proved in [**131**], is that $\beta >_x 1$ is true if and only if $\beta(\delta)$ branches off to the left of δ, and $\beta <_x 1$ is true if and only if $\beta(\delta)$ branches off to the right of δ. The case $\beta(\delta) = \delta$ cannot occur, because,

if it did, β would have a representative homeomorphism with support disjoint from all of S_{i+1}.

In particular, if we denote by B_S the subgroup of B_n consisting of elements which have representative homeomorphisms with support in a submanifold S of D_n, then we have a hierarchy of convex subgroups

$$\{1\} = B_{D_n \setminus S_{n-1}} \subseteq B_{D_n \setminus S_{n-2}} \subseteq \ldots \subseteq B_{D_n \setminus S_1} \subseteq B_{D_n} = B_n.$$

The order we have just constructed is linear. With a little more work one can prove [**131**] that two different subsurface sequences give rise to different orderings. Thus there is indeed a one-to-one correspondence between subsurface sequences and order types. This completes the proof of (a).

In order to see that the ordering induced by a finite type geodesic γ_x is discrete, as claimed in (ii), we observe that $D_n \setminus S_{n-2}$, the complement of the last proper subsurface in the sequence, consists exactly of one twice-punctured disk. We consider the element β of B_n which can be represented by an homeomorphism with support in $D_n \setminus S_{n-2}$, which exchanges these two punctures, namely the Dehn twist along an arc in $D_n \setminus S_{n-2}$ connecting the punctures (see Figure 6.4). This is the smallest element of B_n satisfying $\beta >_x 1$. □

We deduce the number of conjugacy classes of orderings of finite type on B_n:

Proposition 7.2.6. — *For $n \geqslant 2$, the number N_n of conjugacy classes of orderings of finite type of B_n is given by the recursive formula*

$$N_1 = 1, \quad N_2 = 1, \quad \textit{and} \quad N_n = \sum_{k=1}^{n-2} \binom{n-2}{k-1} N_k \, N_{n-k}. \tag{7.2.1}$$

Proof. — It is easy to show constructively that every subsurface sequence can be obtained as the subsurface sequence associated to some finite type geodesic. Therefore it suffices to prove that the number of conjugacy classes of subsurface sequences is given by (7.2.1). For our proof, it is more convenient to rewrite the formula as

$$N_2 = 1, \quad \text{and} \quad N_n = N_{n-1} + \sum_{k=2}^{n-2} \binom{n-2}{k-1} N_k \, N_{n-k}. \tag{7.2.2}$$

It is this formula that we shall prove.

The proof is by induction. For $n = 2$, (7.2.2) is obvious, since all our orderings must satisfy $\sigma_1 > 1$, and there is only one ordering of $B_2 \cong \langle \sigma_1 \rangle \cong \mathbb{Z}$ satisfying this condition.

Assuming that (7.2.2) holds for fewer than n punctures, we shall try to count conjugacy classes of subsurface sequences in D_n. There are two possibilities to be considered: either S_1 has one boundary component in the interior of D_n or two, corresponding to cases (ii) and (iii) in Definition 7.2.3. The two cases will correspond also to the two summands in (7.2.2).

In the first case we can, after a suitable conjugation, suppose that S_1 surrounds the leftmost puncture, and $\partial\overline{S}_1$ contains no point left of the leftmost puncture. Then there are, up to conjugacy, N_{n-1} ways left to complete the subsurface sequence in the remaining disk $D_n \setminus S_1$, which contains punctures number $2, \dots, n$.

In the second case, we can by a suitable conjugation achieve that S_1 is isotopic to a neighbourhood of ∂D_n, together with a neighbourhood of a vertical line between punctures number k and $k+1$ for $2 \leqslant k \leqslant n-2$, and moreover that the transverse orientation on the horizontal geodesic segment that connects the two boundary components of $\overline{S}_1$ in the interior of D_n points upwards. The number k is uniquely determined by these requirements. Thus we have found a way to conjugate a given subsurface sequence such that S_1 is of some canonical type, and we have to classify the possible ways of continuing the subsurface sequence. A subsurface sequence must contain $n-2$ more elements $S_2, \dots, S_{n-1}$. Among the pieces $S_i \setminus S_{i-1}$, exactly $k-1$ must lie in the left half, *i.e.*, in the component of $D_n \setminus S_1$ which contains punctures number $1, \dots, k$, and $n-k-1$ in the right half, *i.e.*, in the component of $D_n \setminus S_1$ which contains punctures number $k+1, \dots, n$. There are $\binom{n-2}{k-1}$ ways to distribute the $k-1$ steps in the left half over the $n-2$ steps that are left to be made. Moreover, in the left half there are N_k, and in the right half N_{n-k} different subsurface sequences. Thus in the second case there are $\binom{n-2}{k-1}\ N_k\ N_{n-k}$ different subsurface sequences once the choice of S_1 has been made. □

For instance, we get $N_2 = 1, N_3 = 1$, and $N_4 = 3$. The three different conjugacy classes of orderings of B_4 are represented by the geodesics γ_b of Figure 7.5 (the associated ordering is the σ-ordering), and γ_c and $\gamma_{c'}$ of Figure 7.4. This finishes our discussion of finite type geodesics.

7.2.2. Infinite type geodesics. — We now turn our attention to filling geodesics of infinite type. Such geodesics are necessarily infinite, because they fill the disk D_n while no finite initial segment does. Contrary to what may be suggested by Figure 7.2(d), such geodesics may have some self-intersections, but only finitely many. More precisely, if $\gamma\colon [0,\infty) \to D_n$ is a filling geodesic of infinite type, then there exists a T in $\mathbb{R}_+$ such that all the self-intersections of γ occur in the initial segment $\gamma([0,T])$. Cutting the disk D_n along this initial segment one obtains some simply-connected pieces, some pieces containing exactly one puncture, and exactly one piece that contains two or more punctures—in fact we shall see that this last component necessarily contains at least three punctures. The geodesic $\gamma([T,\infty))$ separates the punctures in this last component, but no finite initial segment does; actually, it looks like an infinite type geodesic without self-intersections. Therefore we can restrict our attention to such geodesics without self-intersection.

There is a large body of literature that helps us to understand what a filling infinite type geodesic γ without self-intersections must look like. Indeed, let us consider the

bi-infinite, oriented path in D_n obtained by running along the geodesic γ in the opposite direction (terminating on ∂D_n), followed by one turn around the circle ∂D_n, followed by the path γ again, this time in the same direction as γ. The closure of the unique geodesic isotopic to this path is a *geodesic lamination* of D_n in the sense of Nielsen–Thurston theory [**22**].

Definition 7.2.7. — A *subsurface sequence of infinite type* is a finite sequence S_1, $\dots$, S_k, with $k < n-1$, of open connected submanifolds of D_n such that $S_1, \dots, S_{k-1}$ satisfy the same conditions as the elements of a subsurface sequence according to Definition 7.2.3. Moreover, the surface $D_n \setminus \overline{S}_{k-1}$ must be connected, *i.e.*, homeomorphic to a disk with $n-k+1$ punctures. The surface S_k must include S_{k-1}, and $S_k \setminus S_{k-1}$ must be one complementary region of a geodesic lamination in the punctured disk $D_n \setminus S_{k-1}$. In particular, the frontier of S_k must form a geodesic lamination in $D_n \setminus S_{k-1}$.

From any filling geodesic of infinite type one can construct a subsurface sequence of infinite type: the construction procedure is entirely analogous to the finite type case.

A lot is known about the nature of geodesic laminations, and their behaviour under the action of B_n [**22, 119**]. All we need to know for our purposes are the following facts: there are uncountably many geodesic laminations for D_n for $n \geqslant 3$, whereas for $n = 2$ there are only the simple closed curves. For any element β of B_n, the action of β on D_n stabilizes either zero or two geodesic laminations (and if there are two, then β is said to be pseudo-Anosov, and the two laminations are called the stable and unstable laminations of β).

Then we have the following classification result for infinite type orderings:

Theorem 7.2.8. — (i) *All orderings arising from infinite type geodesics are dense.*

(ii) *All but countably many of the uncountably many geodesics of infinite type induce linear orderings of B_n, i.e., all but countably many infinite-type geodesics γ have the property that* $\mathrm{Stab}(\gamma)$ *is trivial.*

(iii) *Two geodesics γ_x, γ_y of infinite type induce the same ordering if and only if they give rise to the same subsurface sequence. Two geodesics γ_x, γ_y of infinite type induce conjugate orderings if and only if their subsurface sequences are conjugate.*

(iv) *There exist uncountably many different orderings of B_n which arise from infinite-type geodesics, and also uncountably many conjugacy classes of such orderings.*

Proof. — (Sketch) Let γ_x be a filling geodesic of infinite type. For the proof of (i) we notice that for an arbitrarily long initial segment of γ_x one can find a geodesic arc connecting two punctures which is disjoint from the initial segment of γ_x. So in order to find an element β of B_n such that $\beta(x)$ is arbitrarily close to x, it suffices

to take a homeomorphism representing a nontrivial element of B_n with support in a neighborhood of the geodesic arc, for instance a Dehn twist along that arc.

For the proof of (ii) we recall that the group B_n is countable, and that each element of B_n stabilizes at most two geodesic laminations. Moreover, only countably many geodesics of infinite type can give rise to the same subsurface sequence. It follows that there can only be countably many geodesics which are stabilized by a nontrivial element of B_n.

Point (iii) is proved in [**131**]. The idea of the proof is quite similar to the finite-type case. No proof will be given here.

As for (iv), the fact that there are uncountably many orderings induced by geodesics of infinite type follows immediately from statements (ii) and (iii). Indeed, there are even uncountably many conjugacy classes of such orderings, because we have the countable group B_n acting by conjugation on our uncountable set of orderings. The set of orbits is still uncountable. □

7.3. A proof of Property S for all Thurston-type orderings

The approach taken in this chapter yields a very natural proof of the left-orderability of B_n. However, it is not well-adapted to proving Properties **A** and **C** in isolation, which refer specifically to σ_1-positive and σ_1-negative braid words; in this chapter, we shall not persue these two properties further. It is, however, interesting and satisfying to see that all orderings of the Thurston type do satisfy Property **S**:

***Proposition 7.3.1* (Property S).** — *If $<_x$ is the ordering arising from our action of B_n on a point x of $\partial\widetilde{D_n}$ with* $\mathrm{Stab}(x) = \{1\}$*, then we have $\sigma_1\beta >_x \beta$.*

This implies, we recall, that inserting any conjugate of σ_1 anywhere in a braid word yields a $<_x$-larger braid word. This proposition is an immediate consequence of the following stronger result:

Lemma 7.3.2. — *If x is any point in $\partial\widetilde{D_n} \setminus \{*\}$ (which is homeomorphic to $\mathbb{R}$), then we have $\sigma_1(x) \geqslant x$ for the ordering induced by the ordering of $\mathbb{R}$.*

Roughly speaking, this means the following: let us suppose that we sit at the point $*$ of $\partial\widetilde{D_n}$ and look toward $\partial\widetilde{D_n} \setminus \{*\}$. During a σ_1-action, we will see a homeomorphism of $\partial\widetilde{D_n} \setminus \{*\}$ which moves some points further to the left, and leaves others fixed, but no points will jump to the right.

Proof of the lemma. — (Sketch) In order to prove that an orientation-preserving homeomorphism of the real line (for instance, the one induced by σ_1) sends every point x of $\mathbb{R}$ to an x' with $x' \geqslant x$, it suffices to prove that its square (for instance, the one induced by σ_1^2) has this property. The homeomorphism σ_1^2 of D_n is a positive Dehn

twist along a simple closed curve c in D_n. The preimage $\widetilde{c}$ of c in $\widetilde{D_n}$ consists of an infinite number of geodesics, each connecting two points of $\widetilde{D_n} \cap S^1_\infty$.

Now we have, as before, a geodesic $\widetilde{\gamma}_x$ in $\widetilde{D_n}$ connecting $*$ to x, but let us also consider all the other possible liftings of its projection γ_x: they are all geodesic rays in $\widetilde{D_n}$, and they are disjoint from each other. Now a curve in $\widetilde{D_n}$ whose endpoint is our desired $\sigma_1(x)$ can be described as follows: it starts at $*$, follows the path $\widetilde{\gamma}_x$ up to its first intersection with $\widetilde{c}$. There it turns left, following $\widetilde{c}$ up to the next intersection with another lifting of γ_x, where it has to turn back right. It follows this pattern, turning left at each intersection with $\widetilde{c}$, and right at each subsequent meeting with $p^{-1}(\gamma_x)$. We leave it to the reader to prove that the endpoint $\sigma_1(x)$ of this path lies to the left of x. □

7.4. A combinatorial approach and an extension to B_∞

At first glance, the geometrical approach developed in this chapter has two disadvantages: firstly, it uses rather deep results from analysis like the uniformization theorem, and, secondly, it seems to work only for B_n with finite n, not for B_∞. However, the definition of the orders coming from hyperbolic geometry and some of the proofs can be given in purely combinatorial manner similar to that of Chapter 5, and, then, the above mentioned disadvantages disappear. This interpretation was first suggested in [**65**], and the authors thank Jonathon Funk for his help in the preparation of this section.

7.4.1. From a geodesic to a word. — It was mentioned above that the choice of the hyperbolic metrics we use actually does not matter. Also, neither does the behaviour of the chosen geodesic γ at infinity or near $\partial\widetilde{D_n}$: for instance, the geodesics γ and γ' in Figure 7.7 lead to the same partial ordering of B_3.

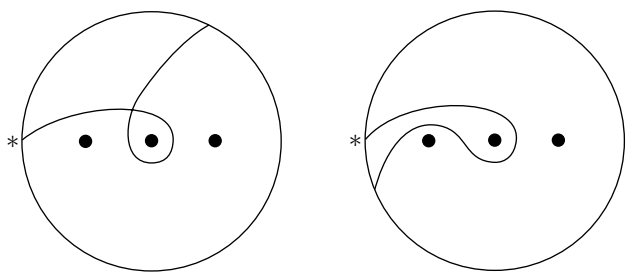


FIGURE 7.7. Two geodesics inducing the same partial order

Following the notation of Chapter 6, we introduce segments $e_1, e_2, \ldots$ between punctures and assume that a hyperbolic metric is chosen so that the interior of each segment e_i is an infinite geodesic. To each finite or infinite geodesic γ originating

at the distinguished point $*$ we attribute a sequence of letters $y_i^{\pm 1}$ in the following way. We go along γ, starting from the base point $*$, and each time we cross e_i we write y_i if we are going in the upward direction, and y_i^{-1} is we are going in the downward direction. The sequence obtained can be either finite, in which case we call it simply a word, or infinite, in which case we call it an *infinite word*. For instance, the geodesics shown in Figure 7.2(b, c, d) are attributed the words $y_1^{-1}y_2^{-1}y_1y_3^{-1}y_2y_4^{-1}y_3$, $y_2^{-1}y_3^{-1}y_1$, and $y_2^{-1}y_3y_4^{-1}y_1^{-1}y_4y_3^{-1}y_2y_3^{-1}$, respectively.

The hyperbolic geometry setting guarantees that all words corresponding to geodesics will be freely reduced words, *i.e.*, contain no subword of the form $y_iy_i^{-1}$ or $y_i^{-1}y_i$.

Once this encoding of geodesics by words has be defined, we can switch to a purely algebraic language. We shall leave it to the reader to convince themselves that the subsequent statements are a reformulation of the above mentioned geometrical results.

7.4.2. Algebraic description of Thurston-type orderings. — Let us denote by $\widehat{F_\infty}$ the set of all finite and infinite reduced nonempty words in the letters $y_i^{\pm 1}$, with $i = 1, 2, \dots$; we introduce an ordering $\lhd$ on $\widehat{F_\infty}$ as follows.

Consider the following *circular* list L involving the letters $y_i^{\pm 1}$ plus one additional symbol y_∞:

$$L: \ y_\infty < \cdots < y_2 < y_1 < y_1^{-1} < y_2^{-1} < \cdots < y_\infty.$$

Here $<$ is meant to indicate that our circular list L has a distinguished direction.

Definition 7.4.1. — For three pairwise distinct elements a, b, c of L, we say that a, b, c *go in the right order* if b is met before c when one goes from a in the direction of the list L.

If we denote y_i^{-1} as y_{-i}, then this definition is equivalent to saying that y_i, y_j, y_k are in the right order when we have

$$\Big(\frac{1}{i} - \frac{1}{j}\Big)\Big(\frac{1}{j} - \frac{1}{k}\Big)\Big(\frac{1}{i} - \frac{1}{k}\Big) > 0.$$

Definition 7.4.2. — Let w_1, w_2 be distinct elements of $\widehat{F_\infty}$. Let w denote the longest common prefix (*i.e.*, left subword) of w_1 and w_2. Let y_i^e be the last (rightmost) letter of w, let z_0 be the inverse letter y_i^{-e}, and let z_1, z_2 be the letters right-adjacent to w in w_1 and w_2, respectively. If w is empty, we set $z_0 = y_\infty$; if w equals w_i, we set $z_i = y_\infty$. Then we declare that $w_1 \lhd w_2$ is true if z_0, z_1, z_2 are in the right order.

This definition, which always makes sense as, by construction, the letters z_0, z_1, and z_2 are pairwise distinct, formalizes the notion of going to the left for geodesics.

The proof of the following statement is straightforward.

Lemma 7.4.3. — *The relation $\lhd$ is a linear ordering on $\widehat{F_\infty}$.*

Note that we do not claim that the ordering $\lhd$ is invariant under left or right multiplication.

Now we construct an action of B_∞ on $\widehat{F_\infty}$. First, we set

$$\sigma_i \cdot y_k = \begin{cases} y_1^{-1}y_2 & \text{for } i = k = 1, \\ y_{i-1}y_i^{-1}y_{i+1} & \text{for } i = k > 1, \\ y_k & \text{for } i \neq k, \end{cases} \quad \sigma_i \cdot y_k^{-1} = \begin{cases} y_2^{-1}y_1 & \text{for } i = k = 1, \\ y_{i+1}^{-1}y_iy_{i-1}^{-1} & \text{for } i = k > 1, \\ y_k^{-1} & \text{for } i \neq k, \end{cases}$$

and define the action of σ_i on a (finite or infinite) word w to be the result of freely reducing the concatenation of the images of the successive letters of w. In this way, the action of σ_i induces a bijection of $\widehat{F_\infty}$, so we can define the action of σ_i^{-1} to be the inverse bijection, and, finally, define the action of an arbitrary braid word on $\widehat{F_\infty}$.

For finite words, the action coincides with the action of B_∞ on $F_\infty \setminus \{1\}$ considered in Section 5.2, so it is not hard to check that the above formulas provide an action of B_∞ on $\widehat{F_\infty}$.

Proposition 7.4.4. — *The action of B_∞ on $\widehat{F_\infty}$ equipped with $\lhd$ is order-preserving.*

Since the linear ordering $\lhd$ of $\widehat{F_\infty}$ is described explicitly, Proposition 7.4.4 can be established by a direct verification of cases similar to those of Section 5.1. This has been done in [**65**] in a slightly different setting.

Proposition 7.4.4 implies that each element x of $\widehat{F_\infty}$ defines a partial ordering of B_∞ by setting $\beta_1 <_x \beta_1$ for $\beta_1 \cdot x \lhd \beta_2 \cdot x$. This ordering $<_x$ is linear if and only if the stabilizer of x in B_∞ is trivial.

Let

$$z = y_1^{-1}y_2^{-1}y_1y_3^{-1}y_2y_4^{-1}y_3 \dots,$$

and let z_n denote the truncated version: $z_n = y_1^{-1}y_2^{-1}y_1 \dots y_n^{-1}y_{n-1}$. Then the geodesic encoded by z_n is exactly the geodesic γ_b considered in Section 7.1.3 and shown in Figures 7.2(b) and 7.5(a). Then the relation $<_{z_n}$ is a linear ordering on B_n for each n, and the relation $<_z$ is a linear ordering on B_∞. The counterpart of Proposition 7.1.6 is then

Proposition 7.4.5. — *The linear ordering $<_{z_n}$ on B_n coincides with the σ-ordering for every n, and, therefore, so does the linear ordering $<_z$ on B_∞.*

In this way, we obtain the eighth equivalent definition of the σ-ordering mentioned in the Introduction:

Corollary 7.4.6. — *For any two braids β_1, β_2 in B_n, the inequality $\beta_1 < \beta_2$ is true if and only if we have $\beta_1 \cdot z_n \lhd \beta_2 \cdot z_n$ in $\widehat{F_\infty}$ (actually, in $F_\infty \setminus \{1\}$, since z_n is a finite word).*

We refer the reader to [**65**] for a further connection of the previous interpretation with the theory of toposes.

CHAPTER 8

TRIANGULATIONS

In this chapter we exhibit a technique which is frequently used for studying mapping class groups and various geometric structures (hyperbolic metrics, foliations, etc.) on surfaces. A triangulation of a surface plays a similar role to that of a basis in a vector space. There is a natural way to associate with every triangulation a "coordinate system" on the set of topological objects of certain type. For instance, the isotopy class of a simple closed curve disjoint from vertices of a triangulation can be uniquely determined from the knowledge of intersection numbers with all edges of the triangulation, provided that the curve is tight with respect to the triangulation (see below for an explanation of the term).

There is a naturally defined elementary operation on triangulations, called flip, which is an analogue of an elementary transformation of a matrix in linear algebra. Any two triangulations having the same set of vertices can be obtained from each other by finitely many flips. This allows one to express geometrical ideas from the previous two chapters in purely combinatorial terms and to construct algorithms for detecting the order in braid groups and, more generally, in mapping class groups.

We describe here two approaches, both using triangulations. The first one is based on the Mosher normal form of a braid and provides automatic ordering of the braid group. The second, lamination, approach was originally developed by one of us (I.D.) [**54, 53**] to detect braid triviality efficiently. It was Stepan Orevkov who suggested that the action of braids on laminations should be more informative and give a simple method for comparing braids with respect to the σ-ordering. These ideas result in a very efficient comparison algorithm, and provide a (very short) proof of Property **A** that can be given without any reference to the geometric origination of the approach.

The algorithms coming from both approaches are quadratic in the length of braids to be compared. The algorithm based on the Mosher normal form approach is implemented on a finite state automaton whose number of states is exponential in the

number of strands. Running time for the algorithm using laminations does not depend on the number of strands provided that the algorithm is implemented on a RAM machine.

8.1. Singular triangulations

The triangulation technique considered in this chapter can be equally well developed for an arbitrary compact surface with or without boundary and not necessarily oriented. But, since in this book we are interested in braid groups, it will suffice for our goals if we restrict ourself to considering only a two-dimensional sphere S^2 as a principal surface. This will allow us to skip many technical details relevant to the general case and concentrate better on the ideas. We fix a set $\mathcal{P}$ of pairwise distinct points of S^2, say $\mathcal{P} = \{P_0, \dots, P_{k-1}\}$ with $k \geqslant 3$, which will be referred to as *punctures*.

The two-sphere S^2 is supposed to carry a piecewise linear (PL) structure and an orientation. All self-homeomorphisms of S^2 considered in this chapter are assumed to be PL and to preserve the orientation. (However, in the figures, we will draw various curves as if they are smooth.)

8.1.1. The notion of a singular triangulation. —

Definition 8.1.1. — A *singular triangulation* of the sphere S^2 with vertices at $\mathcal{P}$ is a set T of simple proper arcs which will be referred to as *edges*, such that
(i) For any edge $e \in T$ we have $\partial e \subseteq \mathcal{P}$;
(ii) The edges in T do not intersect each other except at the ends;
(iii) The edges in T cut the sphere into triangles.
The latter means that each connected component of $S^2 \setminus \bigcup_{e \in T} e$ can be represented as the homeomorphic image of an open two-dimensional simplex Δ^2 under a mapping that can be continuously extended to the boundary $\partial\Delta^2$ and after that sends each side of the simplex onto an edge in T. We allow different edges of $Delta^2$ to map to the same edge in T. These triangles will be referred to as *faces* of the singular triangulation.

In this definition, the number of edges of a singular triangulation equals

$$|T| = 3k - 6$$

by Euler characteristic reason.

By this definition, the set of edges of any triangulation with the set of vertices $\mathcal{P}$ forms a singular triangulation, but not all singular triangulations are of this type. For instance, the patterns displayed in Figure 8.1 may appear in a singular triangulation, while they cannot appear in a triangulation. Thus, singular triangulations generalize the notion of triangulation.

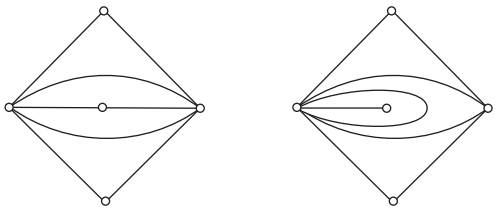


FIGURE 8.1. Singular triangulations

Remark 8.1.2. — In the literature, the most frequently used name for singular triangulations is *ideal* triangulations, the term coming from hyperbolic geometry where edges of a triangulation are geodesic paths.

Let φ be a homeomorphism of the sphere S^2 onto itself preserving the set of punctures: $\varphi(\mathcal{P}) = \mathcal{P}$. For any singular triangulation T of S^2 with vertices at $\mathcal{P}$, say $T = \{e_1, \ldots, e_{3k-6}\}$, we define $\varphi \cdot T$ to be the set of images of the edges of T: $\varphi \cdot T = \{\varphi(e_1), \ldots, \varphi(e_{3k-6})\}$.

Definition 8.1.3. — The orbit of a singular triangulation T under the action of the group of orientation preserving homeomorphisms of S^2 permuting the set $\mathcal{P}$ will be called the *combinatorial type* of T and is denoted by $[T]$. We shall also use the term 'combinatorial type' and similar notation for more complicated objects like a couple of triangulations etc.

In order to specify the combinatorial type of a singular triangulation, it suffices to do the following: cut the sphere along all the edges and enumerate the obtained triangles, for each triangle enumerate its sides in the clockwise order (we assume that an orientation has been chosen on the sphere), then list all pairs of sides of the triangles which are glued together in the sphere. This information can be encoded by a word of bounded length, which immediately implies the following

Lemma 8.1.4. — *There are only finitely many pairwise different combinatorial types of singular triangulations with vertices at the given finite set of punctures $\mathcal{P}$.*

8.1.2. Pulling triangulations tight. — The pulling tight procedure defined in Chapter 6 for curve diagrams extends naturally to the case of singular triangulations.

Definition 8.1.5. — Let T and T' be two singular triangulations of S^2 with vertices at $\mathcal{P}$.We say that T and T' are *transverse* to each other if, for all edges e in T and e' in T', the edges e and e' either coincide or intersect transversely (possibly in several points).

Definition 8.1.6. — By a *D-disk* of a transverse pair of singular triangulations T, T' we shall mean a 2-disk whose interior is disjoint from T and T' and whose boundary

consists of two arcs $\alpha \subseteq e$, $\alpha' \subseteq e'$, where e and e' are edges of T and T', respectively. If there is no D-disk of transverse singular triangulations T, T', then they are said to be *tight*.

Lemma 8.1.7. — *For any singular triangulations T, T' of S^2 with vertices at $\mathcal{P}$ there exists a homeomorphism φ of S^2 identical at punctures and isotopic to the identity relative to $\mathcal{P}$ such that the singular triangulations T and $\varphi \cdot T'$ are tight.*

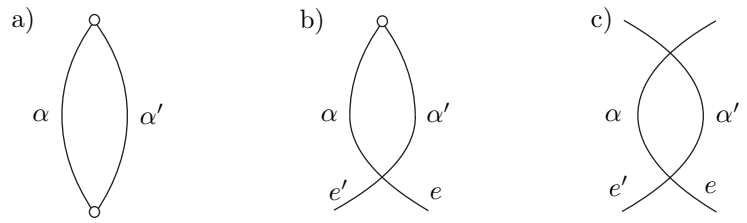


FIGURE 8.2. Three types of D-disks

Proof. — Without loss of generality, we may assume that T and T' are transverse. Suppose they have a D-disk bounded by arcs α and α' described in Definition 8.1.6. There are three possible cases, indicated in Figure 8.2: a) the arcs α and α' can be whole edges of T and T'; b) only one common end of α and α' is a puncture; c) no one common end of α and α' is a puncture. In all these cases there exists a homeomorphism φ sending α to α' and preserving the rest of the singular triangulation T'. In cases b) and c), by a small perturbation of φ we make the edge $\varphi(e')$ be transverse to T, see Figure 8.3. In this way, we obtain a singular triangulation $\varphi \cdot T'$ which is transverse to T and has a smaller number of transverse intersection points with T. After finitely many applications of this procedure we obtain the desired homeomorphism. This process is called *pulling tight*. □

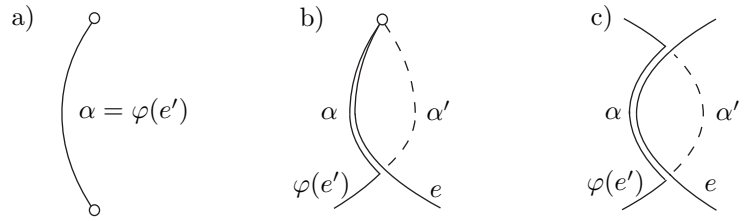


FIGURE 8.3. Pulling triangulations tight

Lemma 8.1.8. — *If two singular triangulations T, T' are isotopic and tight, then they coincide.*

Proof. — For two transverse singular triangulations T_1, T_2, let $\rho_1(T_1, T_2)$ be the number of transverse intersections of edges of T_1 with those of T_2, $\rho_2(T_1, T_2)$ the number of edges of T_1 that are not edges of T_2, and $\rho(T_1, T_2)$ the sum $\rho_1(T_1, T_2) + \rho_2(T_1, T_2)$. The idea of the pulling tight process is to decrease $\rho(T, T')$ as much as possible by deformation of T'.

The assumption of the lemma and the general position argument imply that there exists a sequence of singular triangulations $T' = T_0, T_1, \ldots, T_N = T$ transverse to T such that each passage $T_i \mapsto T_{i+1}$ is either eliminating a D-disk of T, T_i (in which case we have $\rho(T, T_{i+1}) < \rho(T, T_i)$) that we described above or the inverse operation. Among all such sequences let us choose the one for which $\sum_i \rho(T, T_i)$ takes the minimum value. Let $\rho(T, T_j) = \max_i \rho(T, T_i)$.

Suppose that $j \neq 0$. Then both passages $T_j \mapsto T_{j-1}$ and $T_j \mapsto T_{j+1}$ consist in eliminating a D-disk. Let D_1 be the D-disk for the first transform and D_2 for the second one. Then the pair of operations $T_{j-1} \mapsto T_j \mapsto T_{j+1}$ can be replaced by another pair $T_{j-1} \mapsto T'_j \mapsto T_{j+1}$, where the first passage $T_{j-1} \mapsto T'_j$ eliminates D_2 and the second one $T'_j \mapsto T_{j+1}$ creates D_1. We will have $\rho(T, T'_j) < \rho(T, T_j)$, which contradicts to the minimality of the sequence.

Thus, we have $j = 0$. The first passage $T_0 \mapsto T_1$ cannot be an elimination of a D-disk because T_0 and T are tight, and cannot be a creation of a D-disk because $\rho(T, T_0) = \max_i \rho(T, T_i)$ by construction. Hence, $N = 0$ and $T' = T$. □

Now suppose that, in addition to the hypothesis of Lemma 8.1.7, we have one more singular triangulation T'' which is tight with both T and T'. Then, at each step of the pulling tight process described in the proof of Lemma 8.1.7, the singular triangulation T'' remains tight with $\varphi \cdot T'$. Indeed, the intersection of any edge of T'' with a D-disk of T and T' must be an arc connecting a point at α with a point at α' (see Fig. 8.4). Therefore, the homeomorphism φ in the proof of Lemma 8.1.7 can be chosen so as to preserve the singular triangulation T''.

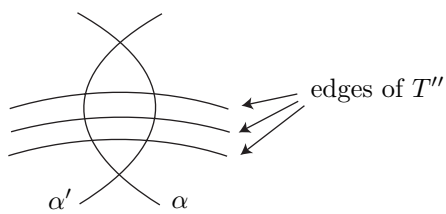


FIGURE 8.4. Triple tightening

In the particular case when T and T' are isotopic, we get the following result, which is an analogue of Proposition 6.2.1 for the case of triangulations. It says that the result of the pulling tight process does not depend on the order in which we reduce D-disks.

Proposition 8.1.9. — *For three singular triangulations T, T' and T'' of S^2 with vertices at $\mathcal{P}$ such that T, T'' are tight, T', T'', are also tight, and T and T' are isotopic, there exists a homeomorphism φ isotopic to the identity relative to $\mathcal{P}$ that preserves T'' and sends T to T'.*

It follows from this proposition that singular triangulations T and T' are tight if and only if they are transverse and, with the notation of the proof of Lemma 8.1.8, $\rho(T, T'')$ takes the minimal value at $T'' = T'$ among all singular triangulations T'' isotopic to T'.

Proposition 8.1.10. — *Arbitrarily many singular triangulations $T_1, \dots, T_q$ of S^2 with vertices at $\mathcal{P}$ can be pulled tight pairwise. More precisely, there exist homeomorphisms $\varphi_1, \dots, \varphi_q$ of S^2 isotopic to identity relative to $\mathcal{P}$ such that the singular triangulations $\varphi_1 \cdot T_1, \dots, \varphi_q \cdot T_q$ are pairwise tight.*

Proof. — We apply the pulling tight proces successively to pairs of singular triangulations: (T_1, T_2), (T_1, T_3), ..., (T_1, T_q), (T_2, T_3), (T_2, T_4), ..., (T_2, T_q),..., (T_{q-1}, T_q). As we have seen above, at each step of this process, singular triangulations that are already tight remain tight. □

In what follows, we shall not distinguish a singular triangulation T with vertices at $\mathcal{P}$ from any isotopic singular triangulation, *i.e.*, a singular triangulation of the form $\varphi \cdot T$, where φ is a homeomorphism isotopic to identity relative to $\mathcal{P}$. All simultaneously considered singular triangulations will be assumed to be pairwise tight.

8.1.3. Flips. — Let T be a singular triangulation and e be one of its edges that separates two different faces F_1 and F_2 of T. Then, the union $F_1 \cup F_2$ can be cut into two triangles in a different way as shown in Figure 8.5. By replacing the edge e with the edge e' we obtain another singular triangulation T'.

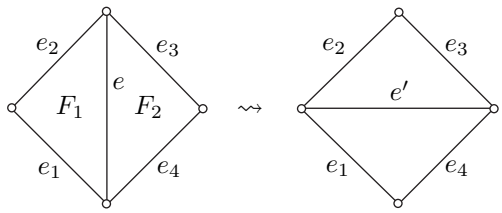


FIGURE 8.5. Flip

Definition 8.1.11. — In the situation above, we shall say that the singular triangulation T' is obtained from T by *flipping* the edge e.

The four vertices involved in a flip, *i.e.*, vertices of the faces F_1 and F_2 in the definition above may or may not be distinct. For instance, the two singular triangulations shown in Figure 8.1 can obtained from each other by a flip.

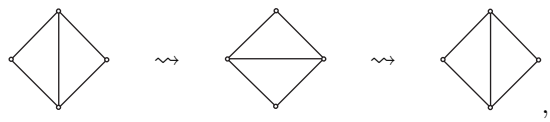
The importance of flips is due to their two well known properties given in Propositions 8.1.12 and 8.1.13.

Proposition 8.1.12. — *Any two singular triangulations T and T' having the same set of vertices $\mathcal{P}$ can be obtained from each other by finitely many flips.*

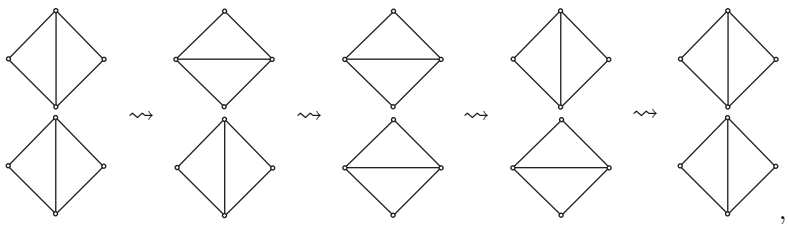
We shall prove this claim in the next section by providing an algorithm for constructing a sequence of flips transforming one triangulation to another. This sequence will be called the combing sequence. The finiteness of the length of the combing sequence will be established in Proposition 8.2.6.

Let us mention without proof another property of flips (which will not be used in the sequel). Consider the following sequences of flips that lead to the initial singular triangulation:

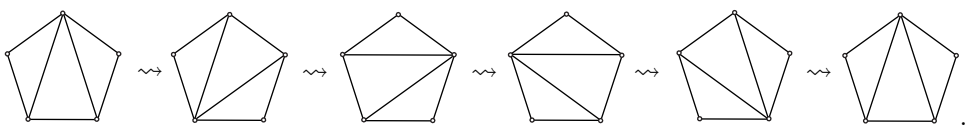
(i) a flip and its inverse:



(ii) four flips involving non-intersecting pairs of triangles and making a commutator:



(iii) five flips involving three triangles as follows:



Then we have

Proposition 8.1.13. — *Any sequence of flips that does not change a singular triangulation can be obtained from the empty one by applying finitely many operations of inserting or removing a subsequence of one of the three types listed above.*

8.2. The Mosher normal form

8.2.1. Finite state automata. — In this section we remind standard notions from the theory of automatic groups. We introduce here only a few basic notions of the theory and refer to [**57**] for further development.

Let A be a finite set, which will be referred to as an *alphabet* and its elements as *letters*. The set of all words on A will be denoted by A^*.

Definition 8.2.1. — A *finite state automaton* over the alphabet A is the collection of the following data:
(i) a finite oriented graph, whose vertices are called *states* and whose edges are called *arrows*;
(ii) for each state s, a bijection between A and the set of arrows coming from s (which are said to be *marked* with the corresponding letter);
(iii) a subset of states, whose elements are called *accept states* and all the other states are called *failure states*;
(iv) a distinguished state s_*, which is called the *start state*.

Clearly, for any word w on the alphabet A, say $w = a_1 \dots a_k$, and any state s of a finite state automaton M, there is a unique sequence $e_1, \dots, e_k$ of arrows of M such that e_1 starts at s, e_{i+1} starts at the end of e_i for $1 \leqslant i \leqslant k-1$, the arrow e_i is marked with a_i for $1 \leqslant i \leqslant k$. If e_k points to s', we say that the word w *reads* from the state s to s'.

Definition 8.2.2. — Let M be an automaton over A. A word w in A^* is said to be *accepted* by M if it reads from the start state s_* to an accept state. A subset L of A^* is said to be a *regular language* if there exists a finite state automaton over A that accepts a word w of A^* if and only if w lies in L.

It makes sense to say that a subset of $A^* \times A^*$ is a regular language. This means that elements of $A^* \times A^*$ are thought of as words in the alphabet $(A \cup \{\$\}) \times (A \cup \{\$\})$, where \$ is an artificially added symbol. If we are given a couple of words (w, w') in the alphabet A, we turn it into a word in $(A \cup \{\$\}) \times (A \cup \{\$\})$ as follows.

Let $w = a_1 \dots a_k$, $w' = a'_1 \dots a'_{k'}$, and assume that w is shorter than w'. Then the corresponding word in $(A \cup \{\$\}) \times (A \cup \{\$\})$ is this:

$$(a_1, a'_1) \dots (a_k, a'_k)(\$, a'_{k+1}) \dots (\$, a'_{k'}).$$

The case when the second word is shorter or they are of equal length is similar.

8.2.2. Combing sequence. — In this section we adopt the general construction of [**110**] to our specific case of the braid group B_n.

Definition 8.2.3. — By an *ordered oriented* singular triangulation we shall mean a singular triangulation whose edges are ordered and they are given an orientation.

Let T be a singular triangulation, T' an ordered oriented singular triangulation with edges $e_0, e_1, e_2, \ldots$. If triangulations T and T' do not coincide (as unordered nonoriented ones), let r be the least i satisfying $e_i \notin T$, and let α be the part of e_r between the starting point of e_r and the first intersection point with an edge of T.

Definition 8.2.4. — In the situation described above, we call the collection of arcs $\{e_i\,;\ 0 \leqslant i < r\} \cup \{\alpha\}$, which is supposed to keep ordering and orientation from T', the *leading part* of T' relative to T, see Figure 8.6. The edge f of T that cuts α will be called the *next-to-be-flipped* edge of T with target triangulation T'. The total number of transverse intersection points of edges of $T \setminus \{f\}$ with T' will be referred to as the *distance* from T to T' and denoted by $d(T, T')$ (notice that d is not symmetric).

Let $\lambda = \{e_i\,;\ 0 \leqslant i < r\} \cup \{\alpha\}$ be the leading part of T' relative to T; then we denote by $\overline{\lambda}$ the union $\left(\bigcup_{i=0}^{r-1} e_i\right) \cup \alpha$ of the arcs from λ.

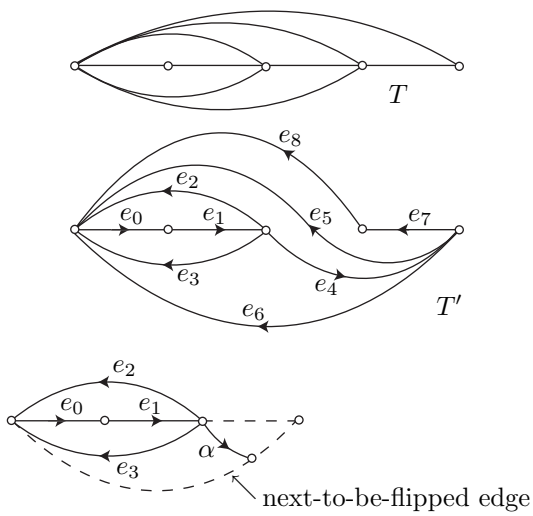


FIGURE 8.6. Leading part of a triangulation

Definition 8.2.5. — Let T be a singular triangulation, T' be an ordered oriented triangulation. The *combing sequence* of T relative to T' is the sequence of singular triangulations $T_0, T_1, T_2, \ldots$ that is uniquely defined by the rules: $T_0 = T$, and T_{i+1} is obtained from T_i by flipping the next-to-be-flipped edge of T_i with target triangulation T'.

Proposition 8.2.6. — *The combing sequence $T_0, T_1, \ldots$ terminates, i.e., for some N, we have $T_N = T'$.*

Proof. — Let f_i be the next-to-be-flipped edge of the singular triangulation T_i and g_{i+1} be the edge of T_{i+1} that replaces f_i after the flip. It is easy to see that, for any $i = 1, 2, \dots$, the edge g_i has a smaller number of transverse intersection points with the edges of T' than f_i does. Indeed, if the flips $T_{i-1} \to T_i$ and $T_i \to T_{i+1}$ are caused by the same edge of T', then f_i has one more intersection point with the edges of T' if compared with g_i, see Figure 8.7. If the edges causing those flips are different, then, by construction, g_i does not intersect T', but f_i does. In both cases, we have $d(T_{i-1}, T') > d(T_i, T')$. Indeed, we find $T_i \setminus \{f_i\} = (T_{i-1} \setminus \{f_{i-1}\}) \cup \{g_i\} \setminus \{f_i\}$. Since the distance from the starting triangulation T to the target triangulation T' is finite, for some N, we get $d(T_{N-1}, T') = 0$. For the next singular triangulation T_N, we will obviously have $T_N = T'$. □

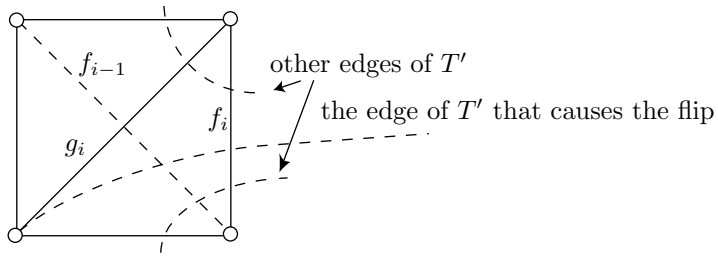


FIGURE 8.7. Combing terminates

8.2.3. Mosher normal form of a braid. — Let us fix some punctures $P_0, \dots, P_{n+1}$ on the sphere S^2 and an arc e_* connecting P_0 with P_{n+1} and disjoint from the other punctures.

In this section we shall consider only singular triangulations with vertices at $\mathcal{P}$ and having e_* as an edge, which will be referred to as the *distinguished edge*. All homeomorphisms of S^2 considered in this section are supposed to be fixed on e_*. Recall that all homeomorphisms are also assumed to preserve orientation. The group $\mathcal{MCG}(S^2, e_*; \mathcal{P})$ of isotopy classes of such homeomorphisms permuting the set $\mathcal{P}$ is clearly isomorphic to the mapping class group of an n-punctured disk, which, in turn, is isomorphic to B_n (see 1.1.2). Indeed, we may cut the sphere S^2 along e_* and think of a self-homeomorphism fixed at e_* as a self-homeomorphism of the 2-disk thus obtained (see Figure 8.8).

In order to simplify notation we use the same letter for a braid and any representative of the corresponding isotopy class from $\mathcal{MCG}(S^2, e_*; \mathcal{P})$. For instance, we shall write $\beta \cdot T$ for the image of a singular triangulation T under any homeomorphism presenting the braid β. This makes sense, since we have agreed above not to distinguish between isotopic singular triangulations. A singular triangulation of the cut sphere is

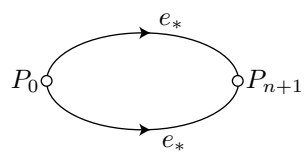


FIGURE 8.8. Cut sphere

a particular case of a diagram in the sense of 1.1.2. It is clear that a homeomorphism which fixes a singular triangulation is isotopic to the identity. In conjunction with Proposition 1.1.3, this implies the following

Lemma 8.2.7. — *In the settings above, the action of the braid group B_n on the set of singular triangulations (which are viewed up to isotopy) with $n+2$ vertices is free: only the trivial braid acts as the identity.*

In order to define the Mosher normal form we must choose a *base ordered oriented singular triangulation*. We denote the base singular triangulation by T_* and its edges by $e_0, e_1, \dots, e_{3n-1}$. For the sake of definiteness, we assume that the distinguished edge is the last one, $e_* = e_{3n-1}$, and that the latter is oriented from P_0 toward P_{n+1}.

Let β be a braid in B_n. Recall that we regard β as a self-homeomorphism of S^2.

Definition 8.2.8. — Let $T_0, \dots, T_N$ be the combing sequence of $T_0 = T_*$ relative to $\beta^{-1} \cdot T_*$, where the singular triangulation $\beta^{-1} \cdot T_*$ is supposed to inherit an ordering and orientation from T_*. The sequence

$$[T_N \to T_{N-1}], [T_{N-1} \to T_{N-2}], \dots, [T_1 \to T_0], \tag{8.2.1}$$

which consists of combinatorial types of flips, is called the *Mosher normal form* of the braid β.

Recall that 'combinatorial type' means up to a homeomorphism. In our case, this simply means that, for any braid β, we consider $[T \to T']$ and $[\beta \cdot T \to \beta \cdot T']$ to be equal. The singular triangulations appearing in the Mosher normal form of a braid are regarded as unordered and nonoriented. Since there are finitely many combinatorial types of singular triangulations and any singular triangulation can be flipped in finitely many different ways, there are only finitely many combinatorial types of flips. In other words, the Mosher normal form is a word in a *finite* alphabet.

Proposition 8.2.9. — *The set of Mosher normal forms of all braids from B_n is a regular language.*

Proof. — We shall construct an automaton M that is a word acceptor for Mosher normal forms. This automaton will play a key role in constructing an algorithm detecting the order in the braid group.

First, we establish conditions a sequence of flips must satisfy in order to be the Mosher normal form of a braid. Let

$$[T'_N \to T''_{N-1}], [T'_{N-1} \to T''_{N-2}], \dots, [T'_1 \to T''_0] \tag{8.2.2}$$

be a sequence of combinatorial types of flips. If it is the Mosher form of a braid, then the following condition is necessarily satisfied.

Condition 1. All subsequent flips in (8.2.2) are composable, the starting and the ending triangulation are combinatorially equivalent to T_*. In other words, we have

$$[T'_N] = [T''_0] = [T_*], \quad [T'_i] = [T''_i] \quad \text{for } i = 1, \dots, N-1. \tag{8.2.3}$$

The latter means that there exist braids $\beta_0, \dots, \beta_N$ satisfying

$$\beta_N \cdot T_*, = T'_N \quad \beta_0 \cdot T''_0 = T_*, \quad \beta_i \cdot T''_i = T'_i. \tag{8.2.4}$$

If Condition 1 holds, then the braids $\beta_0, \dots, \beta_N$ satisfying (8.2.4) are unique, since the action of braids on singular triangulations is free. Let $T_i = (\beta_0 \dots \beta_{i-1}) \cdot T'_i = (\beta_0 \dots \beta_i) \cdot T''_i$ and $\beta = (\beta_0 \dots \beta_N)^{-1}$. Then the sequence (8.2.2) can be rewritten as

$$[T_N \to T_{N-1}], [T_{N-1} \to T_{N-2}], \dots, [T_1 \to T_0], \tag{8.2.5}$$

and we have

$$T_0 = T_*, \quad T_N = \beta^{-1} \cdot T_*.$$

So, the braid β is the only candidate to be a braid of which the sequence (8.2.2) is the Mosher normal form. Let λ_i be the leading part of $\beta^{-1} \cdot T_*$ relative to T_i. The sequence (8.2.5), and hence the sequence (8.2.2), is the Mosher normal form of β if and only if the following holds:

Condition 2. We have $\lambda_{i-1} \neq \lambda_i$, and $\overline{\lambda_{i-1}} \subseteq \overline{\lambda_i}$ for $i = 1, \dots, N$.

To verify Condition 2 for a given i it is enough to know the combinatorial type of the pair (T_i, λ_i), which we denote by $[T_i, \lambda_i]$, and that of the flip $T_i \to T_{i-1}$. If it is satisfied, then the combinatorial type of the pair (T_{i-1}, λ_{i-1}) can be found from the knowledge of $[T_i \to T_{i-1}]$ and $[T_i, \lambda_i]$. It is now left to observe that there are only finitely many different combinatorial types of pairs (T, λ), where T is a singular triangulation and λ is an ordered collection of oriented arcs that *can* be a leading part of another singular triangulation T' relative to T.

We are now ready to construct an automaton M satisfying Conditions 1 and 2, which will complete the proof. □

Definition 8.2.10. — A *marked singular triangulation* is a pair (T, λ), where T is a singular triangulation and λ is either the same singular triangulation T provided with ordering and orientation or an ordered collection of pairwise disjoint oriented arcs $\lambda^0, \dots, \lambda^r$ such that, for $0 \leqslant i < r$, the arc λ^i is an edge of T and λ^r is contained entirely in a face of T and joins a vertex of this face with a point at the opposite side. This side is called the *next-to-be-flipped* edge of (T, λ).

Clearly, if λ is the leading part of an ordered oriented singular triangulation T' relative to T, then the next-to-be-flipped edge of a marked triangulation (T, λ) coincides with the next-to-be-flipped edge of T with target triangulation T'.

States of M. All the states of M, except one dead end state, are combinatorial types of marked triangulations. The pair (T_*, T_*) (where the first entry is considered as an unordered nonoriented singular triangulation) is the start state. The states of the form (T_*, λ) are accept states, all the other are failure states.

Arrows of M. We set up arrows of M so that, for any two marked singular triangulations (T, λ), (T', λ') such that T' is obtained from T by a flip, the automaton M has an arrow from $[T, \lambda]$ to $[T', \lambda']$ marked with the flip $[T \to T']$ if and only if the inverse flip $T' \to T$ is performed on the next-to-be-flipped edge of (T', λ'), $\overline{\lambda'} \subseteq \overline{\lambda}$ holds, and the enumeration of arcs in λ' is inherited from λ. All the other arrows point to the dead end failure state.

In other words, let (T, λ) be a marked singular triangulation, $T' \to T''$ be a flip. The arrow originating from the state $[T, \lambda]$ and marked with the letter $[T' \to T'']$ points to the dead end failure state unless the following conditions are satisfied:
(i) The combinatorial types of T and T' coincide, $[T] = [T']$, *i.e.*, there exists a braid β satisfying $\beta \cdot T = T'$, and (ii) The flip $T' \to T''$ cuts off a non-trivial part λ' of $\beta \cdot \lambda$. If the conditions hold, the arrow will point to the state $[T'', \lambda']$.

It is shown in [**110**] that the Mosher normal form satisfies certain conditions which are expressed by saying that the mapping class groupoid has an automatic structure (see definitions in [**110**] or [**57**]). This implies the following result.

Proposition 8.2.11. — *The Mosher normal form of a braid can be computed in quadratic time in the length of the given braid word, provided that the number of strands is fixed.*

This means, in particular, that the Mosher form itself has length at most quadratic in the length of the corresponding braid.

8.2.4. Braid ordering via Mosher normal form. — The following result is established in [**129**]:

Proposition 8.2.12. — *Under an appropriate choice of the base singular triangulation T_*, there exists an algorithm that, given the Mosher normal forms of two braids β_1, β_2 in B_n, detects their relative order in time linear in the length of the input, provided that n is fixed.*

Remark 8.2.13. — In [**129**] a more general statement is proved, where the braid group is replaced with a mapping class group of an arbitrary surface of finite type with nonempty boundary. The technique used in the general case is quite similar,

the only difference is in some technical details like the choice of the base singular triangulation.

Remark 8.2.14. — In the settings of [**129**], a right invariant order on the braid group was detected. Since we consider here a left invariant braid ordering, the Mosher normal forms of braids $\beta_1^{-1}, \beta_2^{-1}$ should be used instead.

First, we give a simpler proof of Proposition 8.2.12 than that in [**129**]. Then we outline the original method, which, in fact, proves a stronger result. Proposition 8.2.12 (with β_1, β_2 replaced with their inverses) will be a corollary to the following claim.

Proposition 8.2.15. — *Let M be the automaton constructed in Subsection 8.2.3 for some base singular triangulation T_* and let s_* be its start state. Under an appropriate choice of T_*, there exists a partial ordering $<$ on the set of states of M such that, for any two braids β, β' in B_n, the following holds:*

Let $a_N \dots a_1$, $a'_{N'} \dots a'_1$ be the Mosher normal forms of β^{-1} and ${\beta'}^{-1}$, respectively, let j be the least integer satisfying $a'_j \neq a_j$, and assume that the words $a_N \dots a_j$ and $a'_{N'} \dots a'_j$ read to states s and s', respectively. Then we have $\beta < \beta'$ if and only if s and s' are comparable and we have $s < s'$.

Proof. — As a base singular triangulation we can take any singular triangulation such that, for $i = 0, \dots, n$, the edge e_i starts at P_i and points to P_{i+1}, and for $i = 1, \dots, n-1$ the generator σ_i of the braid group B_n can be presented by a half-twist in a small neighbourhood of e_i, as shown in Figure 8.9. Then the union of $\bigcup\limits_{0 \leqslant i \leqslant n} \beta(e_i)$ is nothing else but the curve diagram of the braid β.

If the Mosher normal forms of two braids are different, then the braids themselves are different, and hence, their curve diagrams are different. The latter means that the first discrepancy in the combing sequences will be caused by the difference in $\beta(e_i)$ and $\beta'(e_i)$ with $i \leqslant n-1$. If this happens at the jth step of the combing process and we have $a_j^{-1} = [T_j \to T_{j+1}]$, ${a'_j}^{-1} = [T_j \to T'_{j+1}]$, then the leading parts λ_j and λ'_j of the target triangulations $\beta \cdot T_*$ and $\beta' \cdot T_*$ relative to T_j indicate the first divergence of the curve diagrams of β and β'.

Thus, the following partial ordering on the states of M will satisfy conditions of the claim. Comparable states are of the form $[T, \lambda]$, $[T, \lambda']$, where both λ and λ' consists of no more than n arcs, $\overline{\lambda}$ and $\overline{\lambda'}$ are simple curves coming from P_0, and no one of $\overline{\lambda}$, $\overline{\lambda'}$ is a part of the other. We set $[T, \lambda] < [T, \lambda']$ if and only if the first divergence of $\overline{\lambda'}$ from $\overline{\lambda}$ is to the left. □

So, we can detect the relative order of two braids β, β' given the Mosher normal forms of their inverses as follows. First, we read the given Mosher normal forms from the end, find the first difference, and cut off the coinciding parts. Second, we input the truncated words to two automata identical to M and read the final states. Third, we compare the final states.

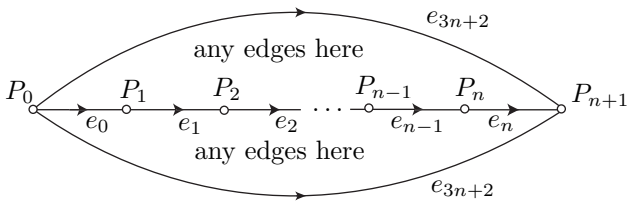


FIGURE 8.9. A base singular triangulation (only the important edges are displayed; the edge e_{3n+2} appears twice because this is a triangulation of the sphere S^2.)

As shown in [**129**], if the base singular triangulation T_* is chosen in a more inteligent way, we shall not need to read the inverted Mosher normal forms completely. It will suffice to read them a little further after the first discrepancy. Namely, the following result has been proved.

Lemma 8.2.16. — *Assume that the base singular triangulation T_* is chosen as shown in Figure 8.10. Let T_0, $T_1, \dots, T_N$ be the combing sequence of T_* with target triangulation $\beta \cdot T_*$, where β is a braid, and λ_i be the leading part of $\beta \cdot T_*$ relative to T_i for $i = 0, 1, \dots, N$. Then the combinatorial type of the marked singular triangulation (T_i, λ_i) can be found from the knowledge of the combinatorial types of the four succesive flips*

$$[T_i \to T_{i+1}], [T_{i+1} \to T_{i+2}], [T_{i+2} \to T_{i+3}], [T_{i+3} \to T_{i+4}]$$

in the general case $i \leqslant N-4$ (resp. of the $N-i$ flips $[T_i \to T_{i+1}], \dots, [T_{N-1} \to T_N]$ for $i = N-3, \dots, N-1$).

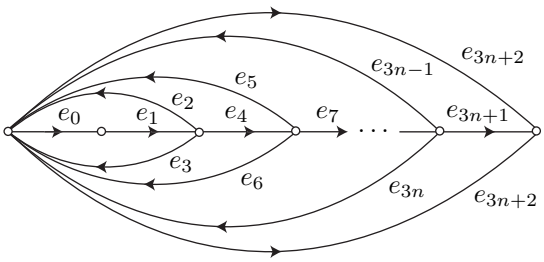


FIGURE 8.10. A better base singular triangulation

We skip the proof, which consists in searching finitely many possibilities. Notice, that the choice of the base triangulation in Lemma 8.2.16 is also well adopted for detecting the order as in the proof of Proposition 8.2.15. As a consequence of Lemma 8.2.16 we get the following two statements:

Corollary 8.2.17. — *Under the choice of T_* as in Lemma 8.2.16, the set of inverted Mosher normal forms of braids from B_n is a regular language.*

Corollary 8.2.18. — *Under the choice of T_* as in Lemma 8.2.16, there exists a finite state automaton that detects the relative order of two braids by reading the inverted Mosher normal forms of their inverses no more than 3 steps further after the first discrepancy.*

We conclude the section by observing that under the choice of T_* as in Lemma 8.2.16, the relative order of β and 1 can be detected from the last element of the Mosher normal form of β^{-1}, *i.e.*, the first flip in the combing sequence. The proof of the next assertion is then an easy exercise.

Proposition 8.2.19. — (i) *A braid β admits a σ-positive (respectively, σ-negative) representative braid word if and only if the first flip in the combing sequence of T_* relative to $\beta{\cdot}T_*$ is performed on the edge e_{3i-1} (respectively, e_{3i}) for some i, $1 \leqslant i \leqslant n$.*

(ii) *For two braids β_1 and β_2, we have $\beta_1 < \beta_2$ if and only if, under the choice of the base singular triangulation as in Figure 8.10, the last flip of the Mosher normal form of $\beta_2\beta_1^{-1}$ occurs in the upper half-sphere.*

We thus have obtained the ninth equivalent definition of the braid ordering mentioned in Introduction (note that, in our settings, $\beta_2\beta_1^{-1}$ should be used instead of $\beta_1^{-1}\beta_2$ because we use the Mosher normal form of the inverses of the braids considered).

8.3. Laminations

Algorithms based on finite state automata as described in the previous section are not so easy to use in practice. The reason is the size of the automata, which is comparable with the number of singular triangulations. The latter grows exponentially with the number of punctures.

Below we use essentially the same geometrical idea as in Chapter 6 for detecting the order of braids by computing their action on laminations. This will allow us to give a fairly simple algorithm for detecting the order, which can be turned into an actual computer program with little effort. A similar algorithm was constructed also in [**103**], but we describe it here in a more explicit way, obtaining as a by-product a set-theoretical solution to the Yang–Baxter equation.

Again, as described in 1.1.2, we visualize the action of a braid by showing the image of a certain diagram under the corresponding homeomorphism of a disk (which we extend to be a homeomorphism of the sphere). Our diagram here will be a certain collection of closed curves, which will be referred to as a lamination.

The notion of a lamination was introduced by William Thurston in [**139**], though he did not use the term "lamination". Laminations naturally appear as infinity points

of the Teichmüller space of a surface. The nice coordinate system on the so-called decorated Teichmüller spaces introduced by R.C. Penner in [**118**] has a natural counterpart for laminations. The action of the mapping class groups on laminations written in terms of this coordinate system becomes piecewise linear and easily computable.

The general definition of a lamination is quite complicated, but for our specific purposes it will not be needed. We shall give a definition of a geometrically simpler object, an integral lamination, which is a partial case of the general notion. We recommend to the reader the paper [**62**] as a nice reference on integral laminations and their connections with Teichmüller spaces.

8.3.1. Normal curves. —

Definition 8.3.1. — For $\mathcal{S}$ a two-dimensional surface, we call a *reduced curve* on $\mathcal{S}$ any compact one-dimensional submanifold of $\mathcal{S}$ without boundary no connected component of which is isotopic to zero, *i.e.*, bounds a disk. Notice that the reduced curves that we shall use are *not* assumed to be oriented.

Let T be a singular triangulation of the sphere S^2 with vertices at $\mathcal{P}$.

Definition 8.3.2. — We say that a reduced curve γ included in $S^2 \setminus \mathcal{P}$ is *normal* with respect to the singular triangulation T if:
(i) The curve γ intersects any edge of T transversely;
(ii) Among the connected components of $S^2 \setminus \gamma \cup \left(\bigcup_{e \in T} e\right)$, there is no disk whose boundary consists of two arcs one of which is a part of γ and the other is a part of an edge of T.

The pulling tight process for reduced curves is very similar to that of curve diagrams and triangulations. By using it, one can prove the following two results.

Proposition 8.3.3. — *For every reduced curve γ included in $S^2 \setminus \mathcal{P}$ there exists a normal curve γ' isotopic to γ.*

Proposition 8.3.4. — *Let γ_1 and γ_2 be two reduced curves on $S^2 \setminus \mathcal{P}$ normal with respect to a singular triangulation T. The normal curves γ_1 and γ_2 are isotopic in $S^2 \setminus \mathcal{P}$ if and only if, for any edge e in T, the number of intersection points of γ_1 with e is equal to that of γ_2.*

In other words, the isotopy class ψ of reduced curves on $S^2 \setminus \mathcal{P}$ can be uniquely determined from the knowledge of the number of intersections of a normal curve from ψ with every edge of a singular triangulation with vertices at $\mathcal{P}$. If these numbers are known for a singular triangulation T, then one can easily compute them for any other singular triangulation obtained from T by a flip. This is done as follows.

Let e be an edge of the singular triangulation T separating two faces F_1 and F_2, e_1, e_2, e_3, e_4 be the four sides of the quadrilateral formed by F_1 and F_2 as shown in Figure 8.5, and let γ be a curve normal with respect to T.

Lemma 8.3.5. — *In the situation described above, one can always flip the edge e so that the curve γ remains normal with respect to the new singular triangulation. Let e' be the flipped edge and $n_i = \#(\gamma \cap e_i)$, $n = \#(\gamma \cap e)$, $n' = \#(\gamma \cap e')$. Then we have*

$$n' = \max(n_1 + n_3, n_2 + n_4) - n. \tag{8.3.1}$$

The easy proof is left to the reader.

8.4. Integral laminations

Definition 8.4.1. — A connected component of a reduced curve on $S^2 \setminus \mathcal{P}$ is said to be *trivial* if it bounds a disk on S^2 with exactly one puncture inside.

Definition 8.4.2. — By an *integral lamination* on the punctured sphere $S^2 \setminus \mathcal{P}$ we shall mean an isotopy class of reduced curves on $\mathcal{S}$ having no trivial component. For L such a isotopy class, we call any reduced curve from L a *representative* of the lamination L.

The set of all integral laminations on the punctured sphere will be denoted by $\mathcal{L}(S^2; \mathcal{P})$. Any homeomorphism $\varphi : S^2 \to S^2$ preserving the set of punctures, *i.e.*, satisfying $\varphi(\mathcal{P}) = \mathcal{P}$, induces a bijection $\mathcal{L}(S^2; \mathcal{P}) \to \mathcal{L}(S^2; \mathcal{P})$ in the natural way: if γ is a representative of the lamination L, then $\varphi \cdot L$ is the isotopy class of $\varphi(\gamma)$. Clearly, the mapping $L \mapsto \varphi \cdot L$ depends on the isotopy class of φ only, thus, we get a left action of the mapping class group $\mathcal{MCG}(S^2; \mathcal{P})$ on the set $\mathcal{L}(S^2; \mathcal{P})$.

In principle, this action can already be computed in terms of the intersection numbers of a lamination with edges of a singular triangulation. Formally speaking, this action will be partial as an action in the corresponding coordinate space because the coordinates, *i.e.*, the intersection numbers, satisfy certain restrictions like triangle inequalities. In orde to make the construction more symmetric we introduce the following modification of the notion of lamination.

Definition 8.4.3. — An integral lamination L on the punctured sphere $S^2 \setminus \mathcal{P}$ together with a mapping $w : \mathcal{P} \to \mathbb{Z}$ will be called a *decorated integral lamination*. By a *representative* of a decorated integral lamination (L, w) we shall mean a reduced curve γ included in $S^2 \setminus \mathcal{P}$ whose connected components are attributed signs $+$ or $-$ so that:
(i) All non-trivial components of γ are attributed sign $+$ and their union is a representative of L;
(ii) For each P in $\mathcal{P}$, the algebraic number of trivial components of γ surrounding the puncture P equals $w(P)$.

The set of all decorated laminations on $S^2 \setminus \mathcal{P}$ will be denoted by $\widetilde{\mathcal{L}}(S^2;\mathcal{P})$. This set is nothing else but the Cartesian product $\mathcal{L}(S^2;\mathcal{P}) \times \mathbb{Z}^{\#(\mathcal{P})}$. Indeed, we have a natural projection

$$p : \widetilde{\mathcal{L}}(S^2;\mathcal{P}) \to \mathcal{L}(S^2;\mathcal{P}), \tag{8.4.1}$$

which forgets decoration and deletes trivial components, and an embedding

$$\iota : \mathcal{L}(S^2;\mathcal{P}) \to \widetilde{\mathcal{L}}(S^2;\mathcal{P}), \tag{8.4.2}$$

which sends a lamination to the same one with zeros assigned to each puncture. However, as we shall see, the set of decorated integral laminations is better adapted, by using singular triangulations, for introducing a global coordinate system on it than the set of non-decorated integral laminations.

As a partial case of Proposition 8.3.3 we have the following statement, whose proof is easy:

Lemma 8.4.4. — *For any singular triangulation T of S^2 with vertices at $\mathcal{P}$ and any (decorated) integral lamination L on $S^2 \setminus \mathcal{P}$, there exists a representative γ of L normal with respect to T.*

Let T be a singular triangulation with vertices at $\mathcal{P}$, let L be a *decorated* integral lamination on $S^2 \setminus \mathcal{P}$, and let e be an edge of T. According to Lemma 8.3.3 there exists a representative γ of L normal with respect to T. Let us denote by $\langle L, T\rangle(e)$ the *algebraic* number of intersections of γ with e, where the sign of an intersection point is defined by the sign attributed to the corresponding component of γ. Lemma 8.3.4 implies that the number $\langle L, T\rangle(e)$ does not depend on the choice of the representative γ.

Thus, for every singular triangulation T with vertices at $\mathcal{P}$ and every decorated lamination L on $S^2 \setminus \mathcal{P}$, we have constructed a mapping

$$\langle L, T\rangle : T \to \mathbb{Z}, \tag{8.4.3}$$

that is an element of $\mathbb{Z}^T$.

Definition 8.4.5. — We shall say that an element ψ of $\mathbb{Z}^T$ is *even* if, for the sides e_1, e_2, e_3 of any face of T, the sum

$$\psi(e_1) + \psi(e_2) + \psi(e_3)$$

is even. The set of all even elements of $\mathbb{Z}^T$ will be denoted by $(\mathbb{Z}^T)_{\text{even}}$.

Clearly, $(\mathbb{Z}^T)_{\text{even}}$ is a sublattice in $\mathbb{Z}^T$ of maximal rank.

Proposition 8.4.6. — *For every singular triangulation T with vertices at $\mathcal{P}$ the mapping ρ_T defined by*

$$L \mapsto \langle L, T\rangle \tag{8.4.4}$$

is a one-to-one map from $\widetilde{\mathcal{L}}(S^2;\mathcal{P})$ onto $(\mathbb{Z}^T)_{\text{even}}$.

Proof. — By a *configuration* at $P \in \mathcal{P}$ we shall mean a triple (e_1, e_2, e_3) of edges of T (not necessarily distinct) that are sides of a face of T listed counterclockwise such that e_1 and e_2 have P as an endpoint. We denote the set of all configurations at P by $\mathcal{C}(P)$.

Let (L, w) be a decorated integral lamination, let γ be its representative. Let ψ be the image of (L, w) under the mapping ρ_T.

Let (e_1, e_2, e_3) be a configuration at P, and let F be the corresponding face of T. Clearly, $(\psi(e_1) + \psi(e_2) - \psi(e_3))/2$ indicates the algebraic number of connected components of $\gamma \cap F$ that connect a point at e_1 with a point at e_2.

Since all the non-trivial components of γ have positive signs, for any puncture P, we have

$$\min_{(e_1,e_2,e_3)\in\mathcal{C}(P)} (\psi(e_1) + \psi(e_2) - \psi(e_3))/2 = w(P). \tag{8.4.5}$$

This means that from the knowledge of $\rho_T(L, w)$ we can determine the decoration w and hence, the number of intersection points of L with any edge of T. In view of (8.3.4) this means that a decorated lamination can be uniquely restored from its image under ρ_T, *i.e.*, that the mapping ρ_T is injective.

Let us show that ρ_T is surjective. Given $\psi \in (\mathbb{Z}^T)_{\text{even}}$, let us take an even integer N such that for any edge e in T we have $N > 3|\psi(e)|$. For each edge e in T, we mark $N + \psi(e)$ points at e. Then there will exist a unique up to isotopy normal curve γ intersecting edges of T precisely at the marked points. Indeed, topologically, there is only one way to continue γ into each face of T: if e_1, e_2, e_3 are sides of the same face, then we must connect $(\psi(e_1) + \psi(e_2) - \psi(e_3) + N)/2 > 0$ points at e_1 with the same number of points at e_2, etc. (see Figure 8.11) .

We attribute the connected components of γ the positive sign. Then we add $N/2$ trivial components attributed the negative sign around each puncture. It is easy to check that the obtained decorated lamination lies in $\rho_T^{-1}(\psi)$. □

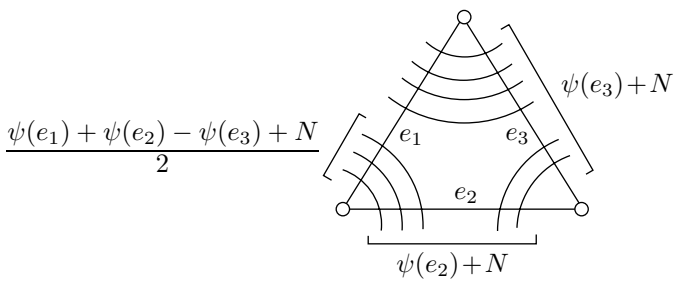


FIGURE 8.11. Recovering the lamination

8.5. Action of the braid group on laminations

We have already mentioned that there is a natural action of the mapping class group $\mathcal{MCG}(S^2;\mathcal{P})$ on the set $\mathcal{L}(S^2;\mathcal{P})$ of laminations as well as on the set $\widetilde{\mathcal{L}}(S^2;\mathcal{P})$ of decorated laminations. To construct the desired action of the group B_n let us first fix $k = n+3$ punctures on the sphere S^2, which we consider as the union $\mathbb{R}^2 \cup \{\infty\}$, at the points

$$P_i = (i, 0) \quad \text{for } i = 0, 1, \dots, n+1, \quad \text{and } P_{n+2} = \infty. \tag{8.5.1}$$

Let D be a disk in the plane $\mathbb{R}^2$ covering the points $P_1, \dots, P_n$, but not points P_0 and P_{n+1}. In Chapter 1, we saw that the braid group B_n can be identified with the mapping class group of the disk D with n punctures. Each homeomorphism of the disk D that is identical at the boundary can be extended to a homeomorphism of the the whole plane by making it identical on the complement to the disk D. In this way we obtain a homomorphism

$$\xi : B_n \to \mathcal{MCG}(S^2;\mathcal{P}), \tag{8.5.2}$$

which, in fact, is an injection as we shall see.

For any (decorated) integral lamination L on $S^2 \setminus \mathcal{P}$ and any braid $\beta \in B_n$ we define $\beta \cdot L$ to be

$$\varphi_\beta \cdot L, \tag{8.5.3}$$

where φ_β is a homeomorphism from the isotopy class $\xi(\beta)$. In the previous section we constructed a bijection $\rho_T : \widetilde{\mathcal{L}}(S^2;\mathcal{P}) \to (\mathbb{Z}^T)_{\text{even}}$ for any singular triangulation. By using it, we obtain a left action of the group B_n on $(\mathbb{Z}^T)_{\text{even}}$, which is easy to compute explicitly as follows.

Let e be an edge of the singular triangulation T. Clearly, we have

$$\rho_T(\beta \cdot L)(e) = \langle \varphi_\beta \cdot L, T\rangle(e) = \langle L, \varphi_\beta^{-1} \cdot T\rangle(\varphi_\beta^{-1}(e)). \tag{8.5.4}$$

Thus, in order to compute $\rho_T(\beta \cdot L)$ we need only to compute the intersection number of the same decorated lamination L with every edge of the image of the singular triangulation T under the homeomorphism φ_β^{-1}. This can be done by using Formula (8.3.1), since, according to Proposition 8.1.12, the singular triangulation $\varphi_\beta \cdot T$ can be obtained from T by finitely many flips.

From now on, we fix a singular triangulation T_0 of the sphere S^2, identified with $\mathbb{R}^2 \cup \{\infty\}$, that will be convenient for us. This singular triangulation consists of

the following edges e_i, $i = 0, 1, \dots, 3n+2$:

$$
\begin{aligned}
e_0 &= \{(x,y)\,;\ x \leqslant 0,\ y = 0\}, \\
e_{3i+1} &= \{(x,y)\,;\ x = i + 1/2\}, \qquad && i = 0, 1, \dots, n, \\
e_{3i-1} &= \{(x,y)\,;\ x = i,\ y \geqslant 0\}, && i = 1, \dots, n, \\
e_{3i} &= \{(x,y)\,;\ x = i,\ y \leqslant 0\}, && i = 1, \dots, n, \\
e_{3n+2} &= \{(x,y)\,;\ x \geqslant n+1,\ y = 0\}
\end{aligned}
\tag{8.5.5}
$$

(see Figure 8.12).

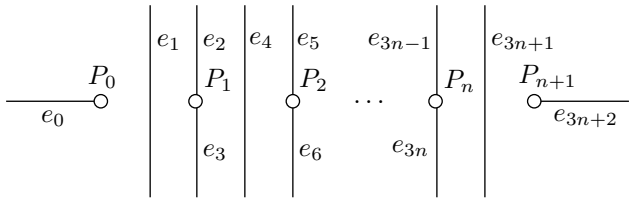


FIGURE 8.12. A distinguished triangulation

We shall now proceed with computing the action of the generator σ_i on the set of decorated laminations in the coordinate system $(z_0, \dots, z_{3n+2})$ given by the map ρ_{T_0}:

$$z_j(L) = \langle L, T_0 \rangle (e_j). \tag{8.5.6}$$

To this end, we present a sequence of four flips that transforms the singular triangulation T_0 to $\varphi \cdot T_0$, where φ is a representative of the isotopy class $\xi(\sigma_i^{-1})$. The sequence of obtained singular triangulations is shown in Figure 8.13. At each edge e of every singular triangulation T appearing in this sequence, the number $\langle L, T \rangle (e)$ of intersections of L with e is indicated. What we need are the numbers

$$z'_j = z_j(\sigma_i \cdot L) = \langle L, \varphi^{-1} \cdot T_0 \rangle (\varphi^{-1}(e_j)). \tag{8.5.7}$$

It can be seen from the figure that the coordinates of $\sigma_i \cdot L$ are

(8.5.8)

$$
\begin{aligned}
z'_{3i} &= z_{3i+3} \\
z'_{3i+2} &= z_{3i-1} \\
z'_{3i-1} &= \max(z_{3i-2} - z_{3i+1} + \max(z_{3i-1} + z_{3i+3}, z_{3i} + z_{3i+2}), z_{3i-1} + z_{3i+3}) - z_{3i}, \\
z'_{3i+3} &= \max(z_{3i+4} - z_{3i+1} + \max(z_{3i-1} + z_{3i+3}, z_{3i} + z_{3i+2}), z_{3i-1} + z_{3i+3}) - z_{3i+2}, \\
z'_{3i+1} &= \max(z_{3i-1} + z_{3i+3}, z'_{3i-1} + z'_{3i+3}) + z_{3i+1} - \max(z_{3i-1} + z_{3i+3}, z_{3i} + z_{3i+2}), \\
z'_j &= z_j \quad \text{for } j \leqslant 3i-2 \text{ and } j \geqslant 3i+4.
\end{aligned}
$$

Thus, we have obtained:

Proposition 8.5.1. — *Formulas (8.5.8) define a left action of the braid group B_n on $(\mathbb{Z}^{T_0})_{\text{even}}$.*

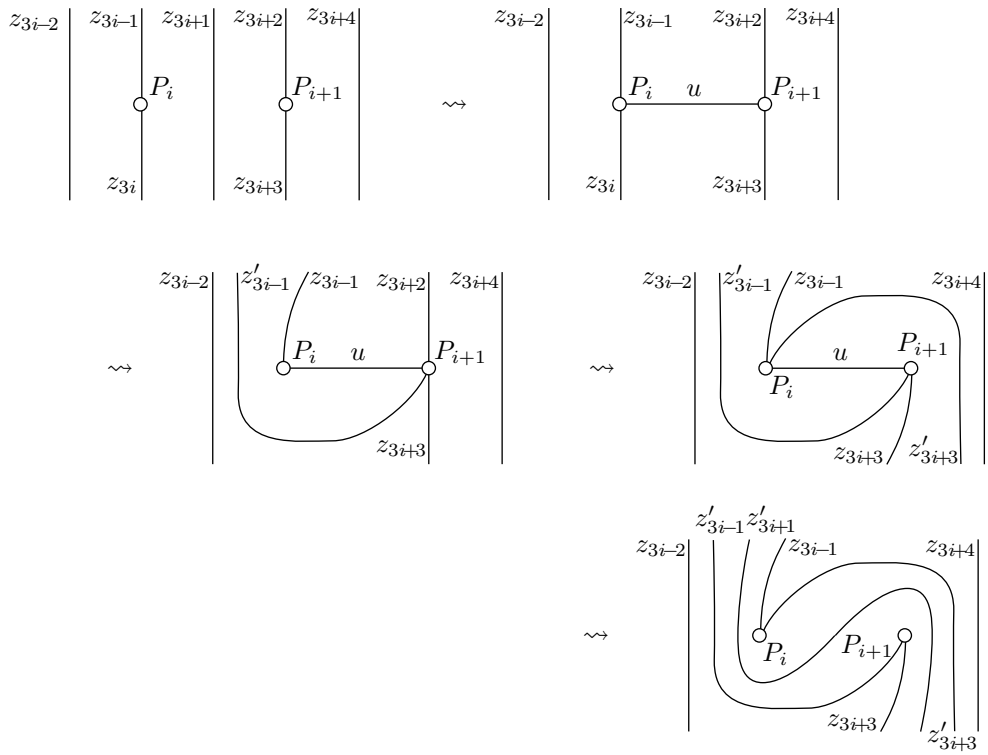


FIGURE 8.13. Action of σ_i on a decorated lamination; we have

$$\begin{aligned} u &= \max(z_{3i-1} + z_{3i+3}, z_{3i} + z_{3i+2}) - z_{3i+1}, \\ z'_{3i-1} &= \max(z_{3i-1} + z_{3i+3}, u + z_{3i-2}) - z_{3i}, \\ z'_{3i+3} &= \max(z_{3i-1} + z_{3i+3}, u + z_{3i+4}) - z_{3i+2}, \\ z'_{3i+1} &= \max(z_{3i-1} + z_{3i+3}, z'_{3i-1} + z'_{3i+3}) - u. \end{aligned}$$

Remark 8.5.2. — It is clear that, for every natural number λ, every η in $(\mathbb{Z}^{T_0})_{\text{even}}$, and every β in B_n, we have

$$\beta \cdot (\lambda\eta) = \lambda(\beta \cdot \eta). \tag{8.5.9}$$

Consequently, we can use the rule (8.5.9) to extend the action (8.5.8) to an action on $\mathbb{R}\mathbb{Z}^{T_0}$ and, since all functions in formulas (8.5.8) are continuous, to an action on $\mathbb{R}^{T_0}$. This space $\mathbb{R}^{T_0}$ can be interpreted as the space of all laminations on $S^2 \setminus \mathcal{P}$, and the action of the whole mapping class group of $S^2 \setminus \mathcal{P}$ can be extended to it.

Now, we shall introduce a different coordinate system on $\widetilde{\mathcal{L}}(S^2;\mathcal{P})$ by using the same singular triangulation T_0. We set

$$a_i = \frac{z_{3i-1} - z_{3i}}{2}, \quad b_i = \frac{z_{3i-2} - z_{3i+1}}{2}, \quad \text{for } i = 1, \dots, n, \tag{8.5.10}$$

and

$$w_j = w(P_j), \quad \text{for } j = 0, 1, \dots, n+2, \tag{8.5.11}$$

where the coordinates z_j are the same as above.

Proposition 8.5.3. — *The mapping*

$$L \mapsto (a_1(L), b_1(L), \dots, a_n(L), b_n(L), w_0(L), \dots, w_{n+2}(L)) \tag{8.5.12}$$

is a bijection from $\widetilde{\mathcal{L}}(S^2;\mathcal{P})$ to $\mathbb{Z}^{3n+3}$.

The first $2n$ coordinates $(a_1, b_2, \dots, a_n, b_n)$ do not depend on the decoration, i.e., they depend only on the underlying integral (non-decorated) lamination, and they define a bijection from $\mathcal{L}(S^2;\mathcal{P})$ to $\mathbb{Z}^{2n}$.

Proof. — It easy to check that the variables z_i are uniquely determined by relations (8.5.10) and relations (8.4.5), which look as

(8.5.13)

$$\begin{aligned} 2w_0 &= 2z_0 - z_1, \\ 2w_i &= z_{3i-1} + z_{3i} - \max(z_{3i+1}, z_{3i-2}), \quad \text{where } 1 \leqslant i \leqslant n, \\ 2w_{n+1} &= 2z_{3n+2} - z_{3n+1}, \\ 2w_{n+2} &= \min(z_1 - 2|a_1|, z_4 - 2|a_1|, z_4 - 2|a_2|, z_7 - 2|a_2|, \dots, z_{3n+1} - 2|a_n|) \end{aligned}$$

in our particular case. □

Now, we shall compute the action of the braid group B_n on the integral lattice $\mathbb{Z}^{2n}$ which is identifined with the set of integral laminations $\mathcal{L}(S^2;\mathcal{P})$ by using coordinates $(a_1, b_2, \dots, a_n, b_n)$.

Proposition 8.5.4. — *For η in $\mathbb{Z}^{2n}$, say $\eta = (a_1, b_2, \dots, a_n, b_n)$, we have $\sigma_i \cdot \eta = (a'_1, b'_2, \dots, a'_n, b'_n)$, with $a'_j = a_j$, $b'_j = b_j$ for $j \neq i$, $i+1$, and*

$$\begin{aligned} a'_i &= a_i + (\delta^+ + b_i)^+, \\ a'_{i+1} &= a_{i+1} - (\delta^+ - b_{i+1})^+, \\ b'_i &= b_i - (-\delta')^+ + \delta^+, \\ b'_{i+1} &= b_{i+1} + (-\delta')^+ - \delta^+, \end{aligned} \tag{8.5.14}$$

where we used the following notation: $\delta = a_{i+1} - a_i$, $\delta' = a'_{i+1} - a'_i$, $x^+ = \max(0, x)$. Further, the action of σ_i^{-1} is given by

$$\sigma_i^{-1} \cdot \eta = (\tau\sigma_i\tau) \cdot \eta, \tag{8.5.15}$$

with $\tau \cdot (a_1, b_1, \dots, a_n, b_n) = (-a_1, b_1, \dots, -a_n, b_n)$.

The proof is an easy computation. Notice that, once Formulae (8.5.14) are given, it is easy but tedious to check that they define an action of the braid group B_n on the set $\mathbb{Z}^{2n}$.

Remark 8.5.5. — The action of the braid group B_n on $\mathbb{Z}^{2n}$ defined above preserves the standard simplectic form $a_1 \wedge b_1 + \cdots + a_n \wedge b_n$. This can be checked easily by writing down the matrix of σ_i for each part of the space where the action of σ_i is linear.

8.5.1. A proof of Property A. — Now we shall apply the action of the braid group B_n defined by (8.5.14) in order to (re)-prove Property **A**, and to give a new characterization of braids bigger than 1 with respect to the σ-ordering.

Proposition 8.5.6. — *Let β be a braid in B_n, and $(a_1, b_1, \ldots, a_n, b_n)$ be the result of applying β to the sequence $(0, 1, \ldots, 0, 1)$ from $\mathbb{Z}^{2n}$ (see Figure 8.14 for a picture of the corresponding lamination). If β admits a representative braid word containing at least one letter σ_i (resp. σ_i^{-1}) and no letter σ_i^{-1} (resp. σ_i) or $\sigma_j^{\pm 1}$ with $j < i$, then we have $a_i > 0$ (resp. $a_i < 0$) and $a_j = 0$ for $j < i$.*

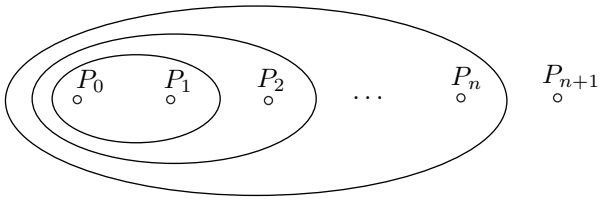


FIGURE 8.14. The lamination $(0, 1, 0, 1, \ldots, 0, 1)$

Proof. — It is clear from Formulae (8.5.14) that applying a σ_i-free braid leaves coordinates a_j, b_j with $j \leqslant i$ unchanged. Once σ_i is applied, the coordinate a_i becomes positive, since it is replaced with

$$a_i + (\delta^+ + b_i)^+ = 0 + (\delta^+ + 1)^+ \geqslant 1.$$

It is also clear from (8.5.14) that a_i cannot decrease if we apply a further σ_i-positive or σ_i-free braid word. □

Remark 8.5.7. — Notice that the elements y_i considered in Section 5.2 are presented by loops that (after a small deformation letting them to go "behind" 0) will look exactly as connected components of a $(0, 1, 0, 1, \ldots, 0, 1)$-lamination. So, it does not come at a surprise that this lamination is used to detect the order. Switching from words presenting elements of a free group to finite sequences of integers presenting an integral lamination is analogous to switching from a non-positional number

system to a positional one. This explains the algorithmic efficiency (see below) of the lamination approach.

***Corollary 8.5.8* (Property A).** — *A braid that admits at least one σ_1-positive braid word representative is not trivial.*

Proposition 8.5.6 and Formulae (8.5.14) provide also a simple algorithm for comparing braids: it suffices to compute $\beta\cdot(0,1,\dots,0,1)$ and apply Proposition 8.5.6. We thus obtain the tenth (and last) equivalent definition of the braid ordering mentioned in Introduction:

Proposition 8.5.9. — *For β_1, β_2 in B_n, the relation $\beta_1 < \beta_2$ is true if and only if the first nonzero coefficient of odd index in the sequence $\beta_1^{-1}\beta_2\cdot(0,1,\dots,0,1)$ is positive.*

We shall now give an estimation for the running time of the previous comparison algorithm. We shall use the notation

$$\Delta_{ij} = (\sigma_i\sigma_{i+1}\dots\sigma_{j-1})(\sigma_i\sigma_{i+1}\dots\sigma_{j-2})\dots\sigma_i \tag{8.5.16}$$

for Garside-like elements, with $i<j$. So, in particular, we have

$$\Delta_{i,i+1} = \sigma_i, \quad \Delta_{i,i+2} = \sigma_i\sigma_{i+1}\sigma_i.$$

Definition 8.5.10. — For a braid word w of the form

$$w = \Delta_{i_1j_1}^{p_1}\dots\Delta_{i_mj_m}^{p_m}, \tag{8.5.17}$$

we define the Δ-*length* of w to be

$$|w|_\Delta = \sum_{s=1}^{m}(1+\log(j_s-i_s)+\log|p_s|). \tag{8.5.18}$$

By the Δ-*length* of a braid β we mean the least Δ-length of a word presenting β and written in the form (8.5.17).

Clearly, for the ordinary length $|w|$ of a braid word we have

$$|w| \geqslant |w|_\Delta,$$

and some words become much shorter in the sense of Definition 8.5.10.

For a point η in $\mathbb{Z}^{2n}$, we denote $\max_j|\eta_j|$ by $\|\eta\|$.

Proposition 8.5.11. — *For every braid β in B_n, we have*

$$\log\|\beta\cdot(0,1,\dots,0,1)\| \leqslant 2|\beta|_\Delta. \tag{8.5.19}$$

Thus, if the braid β is given as a braid word w, then the element $\beta\cdot(0,1,\dots,0,1)$ can be computed algorithmically using $C\cdot|w|_\Delta\cdot|w|$ operations on a RAM machine, where C is a constant that does not depend on the number of strands n. In particular, the relative order of β and 1 can be detected in $O(|w|_\Delta\cdot|w|)$ time.

Proof. — It is not difficult to achieve Inequality (8.5.19) by considering the geometrical picture of applying a Garside-like element to a lamination. The other two assertions follow immediatly. □

The asymptotic estimation given in Proposition 8.5.11 for the running time of the algorithm computing $\beta \cdot (0, 1, \ldots, 0, 1)$ is sharp in the sense that, for certain class of braid words, the number of arithmetical operations needed for computing $w \cdot (0, 1, \ldots, 0, 1)$ by using formulae (8.5.14) will be bounded from below by $C' \cdot |w|^2$, where $C' > 0$ is some constant. This is due to the fact that the coordinates of $\beta^n \cdot (0, 1, \ldots, 0, 1)$ grow exponentially in n if β is a pseudo-Anosov braid. An example of such a braid could be $\sigma_1 \sigma_2^{-1}$.

CHAPTER 9

BI-ORDERING THE PURE BRAID GROUPS

We saw in Section 1.3 that the full braid group B_n is left-orderable, but not bi-orderable (for $n \geqslant 3$). In this chapter, we will see that the *pure* braid group P_n, a normal subgroup of B_n of index $n!$, can be given an ordering invariant under multiplication on both sides. The key is that free groups are bi-orderable, and P_n is a semidirect product of free groups, according to Artin's combing technique.

With appropriate choice of conventions, the ordering has the property that Garside positive pure braids (expressible in the generators σ_i using only positive exponents: $P_n^+ = P_n \cap B_n^+$) are all greater than the identity. We will also see that P_n^+ is *well ordered* under this ordering.

The ordering we will describe for P_n is radically different from those defined for B_n in earlier chapters. It is natural to ask if there is a possible uniform ordering: a left-ordering of B_n which restricts to a bi-ordering of P_n. Perhaps surprisingly, the answer is that this is impossible. That will be explained in the last section.

9.1. Descending central series

We should mention that the bi-orderability of P_n follows from the work of [**59**], which shows that the pure braid groups satisfy the hypothesis of the following proposition. Recall the definition of the descending central series associated with a group G,

$$G = G_0 \supseteq G_1 \supseteq G_2 \supseteq \cdots ,$$

is given inductively by $G_{n+1} = [G_n, G]$, the group generated by commutators $hgh^{-1}g^{-1}$, with $h \in G_n$ and $g \in G$. These are normal subgroups of G, and the quotient groups G_n/G_{n+1} are abelian.

Proposition 9.1.1. — *Suppose G is a group which is residually nilpotent, meaning $\bigcap G_i = \{1\}$, and such that each G_n/G_{n+1} is torsion-free. Then G is bi-orderable.*

Proof. — It is straightforward to bi-order countable torsion-free abelian groups, so take $<_n$ to be an arbitrary bi-ordering of G_n/G_{n+1}. For any distinct elements g, h in G, let $N = N(g,h) = N(h,g)$ be the greatest n such that $g^{-1}h$ belongs to G_n, so it represents a nontrivial class $[g^{-1}h]$ of G_N/G_{N+1}. Define $g < h$ if and only if $1 <_N [g^{-1}h]$. □

Corollary 9.1.2. — *For each n, the pure braid group P_n is bi-orderable.*

9.2. An effective ordering on P_n

We prefer a more constructive approach to bi-ordering P_n, which has the advantage of defining a well-ordering of P_n^+, as well as straightforward computability. The next few sections lay the groundwork for this construction, following [**83**].

9.2.1. Artin combing. — Recall that the pure braids are those which induce the trivial permutation under the canonical map $B_n \to \mathfrak{S}_n$, so there is an exact sequence

$$1 \to P_n \hookrightarrow B_n \to \mathfrak{S}_n \to 1.$$

The standard inclusion $B_n \subseteq B_{n+1}$ restricts to the pure braid groups: $P_n \subseteq P_{n+1}$. However, the pure braid groups have the advantage that the inclusion has a left inverse, the mapping

$$r : P_{n+1} \to P_n$$

defined by 'forgetting the last strand' is a homomorphism.

Moreover, the kernel of r is the set of pure $n+1$ strand braids representable so that the first n strands go straight across (the trivial braid). This is clearly the same thing as the fundamental group of a plane with n points removed, which is of course a free group F_n of rank n. Thus we have the (split) exact sequence

$$1 \longrightarrow F_n \xrightarrow{i} P_{n+1} \xrightarrow{r} P_n \longrightarrow 1.$$

This sequence, together with the fact that F_n can be bi-ordered by a particularly well-behaved ordering, provides an inductive step for ordering all the pure braid groups.

Since r is a retraction of groups, we see that P_{n+1} is a semidirect product of P_n with F_n, and the process may be iterated to present P_{n+1} as a semidirect product of free groups $F_1, \ldots, F_n$, often called the Artin combing of the braid. That is, each pure braid β in P_{n+1} has a unique expression $\beta = \beta_1\beta_2\cdots\beta_n$, where β_i is a braid with all strands straight, except the $(i+1)^{st}$, which can interact only with strands of lower index. We will call the vector $(\beta_1, \beta_2, \ldots, \beta_n)$ the *Artin coordinates* of β in P_{n+1}. Each β_i is a pure braid in the subgroup F_i of P_{n+1}.

We take generators $x_{i,j}$ for F_j defined by $x_{i,j} = \sigma_j\sigma_{j-1}\ldots\sigma_{i+1}\sigma_i^2\sigma_{i+1}^{-1}\ldots\sigma_{j-1}^{-1}\sigma_j^{-1}$ with $1 \leqslant i \leqslant j$, as depicted in Figure 9.1. When the context is understood, we may write $x_{i,j}$ as x_i. Figure 9.2 illustrates a typical combing.

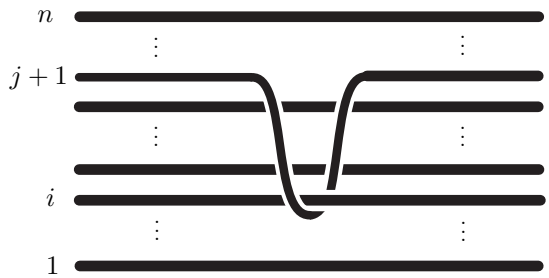


FIGURE 9.1. The generator $x_{i,j}$ of F_j in P_n, for $i \leqslant j$

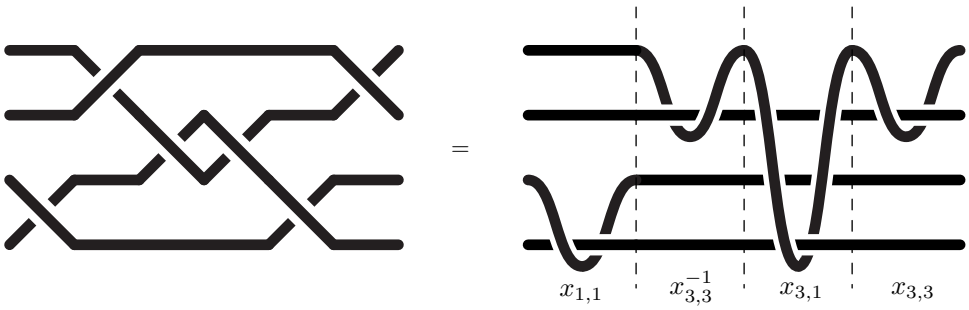


FIGURE 9.2. The pure 4-braid $\sigma_1\sigma_3^{-1}\sigma_2^2\sigma_1\sigma_3$ and its Artin combed form $x_{1,1}x_{3,3}^{-1}x_{1,3}x_{3,3}$

9.2.2. Bi-ordering extensions. — Left-orderability is inherited under extensions, but this is not necessarily true for bi-orderability. Suppose G is a group, with a normal subgroup F and denote the quotient $H := G/F$, with projection $p : G \to H$.

Lemma 9.2.1. — *Suppose F and H have left-invariant orderings. For g, $g' \in G$ define an ordering by declaring $g < g'$ if we have either $p(g) <_H p(g')$ or else $p(g) = p(g')$ and $1 <_F g^{-1}g'$. Then $<$ is a left-ordering of G. If F and H are bi-ordered, $<$ is a bi-ordering of G if and only if conjugation of F by G is order-preserving, that is, $f <_F f'$ implies $g^{-1}fg <_F g^{-1}f'g$ for all f, f' in F and g in G.*

The proof is straightforward and left to the reader.

Example 9.2.2. — The Klein bottle group $K = \langle x, y\,;\, x^{-1}yx = y^{-1}\rangle$ fits in an exact sequence $1 \to \mathbb{Z} \to K \to \mathbb{Z} \to 1$, where the infinite cyclic subgroup is generated by y. The group K is therefore left-orderable. However, K cannot be bi-ordered, for such an ordering must be invariant under conjugation and the defining relation would lead to the contradiction that $1 < y$ is equivalent to $1 < y^{-1}$. The problem here, of course, is that the map $y \to y^{-1}$ cannot possibly be order-preserving.

9.2.3. Magnus expansion. — If F_n is a free group based on $x_1, \cdots, x_n$, it is not obvious, at first glance, that F_n is bi-orderable. Of course it is known that F_n satisfies the hypotheses of Proposition 9.1.1. But a device of W. Magnus gives a uniform, and pretty, way of defining an ordering.

Consider the ring $\Lambda_n = \mathbb{Z}\langle\langle X_1, \cdots, X_n\rangle\rangle$ of formal power series in the non-commuting indeterminates X_i. The Magnus map $\mu : F_n \to \Lambda_n$ is defined by

$$\mu(x_i) = 1 + X_i, \quad \mu(x_i^{-1}) = 1 - X_i + X_i^2 - X_i^3 + \cdots .$$

It is an injective mapping of F_n into the group of units of Λ_n, and, more precisely, into the multiplicative subgroup $1 + O(1)$, where $O(k)$ denotes the subset of Λ_n made of the series involving only terms of total degree $\geqslant k$ [**100**].

Now we order all of Λ_n as follows. Write formal power series with terms in some fixed sequence, in which lower degree monomials come before higher degree. Within a degree one can be arbitrary, but for definiteness we list the monomials of a given degree lexicographically according to the subscripts. Then we compare two series by looking at the first term at which their coefficients differ, and order them according to that coefficient. This gives a linear ordering of Λ_n which is clearly invariant under addition, but not under multiplication (for instance by -1). However, one can check readily that, when restricted to the multiplicative subgroup $1+O(1)$, it is also invariant under multiplication on either side [**83**]. Finally, we order F_n via the embedding μ.

Example 9.2.3. — In this ordering of F_n, we find $1 < x_2 < x_1 < x_2^{-1}x_1x_2$, using the calculations

$$\mu(x_1) = 1 + X_1, \quad \mu(x_2) = 1 + X_2,$$
$$\mu(x_2^{-1}x_1x_2) = (1 - X_2 + X_2^2 - \cdots)(1 + X_1)(1 + X_2) \in 1 + X_1 + X_1X_2 - X_2X_1 + O(3).$$

The following can be established inductively. It is the key to Magnus' proof that free groups are residually nilpotent.

Lemma 9.2.4. — *The subgroup $\mu((F_n)_k)$ is included in $1 + O(k+1)$.*

Proposition 9.2.5. — *Suppose $\varphi : F_n \to F_n$ is an automorphism, and let $\varphi_{\mathrm{ab}} : F_n/[F_n, F_n] \to F_n/[F_n, F_n]$ be the induced automorphism on the abelianization of F_n. If φ_{ab} is the identity, then the ordering defined above is invariant under φ.*

Proof. — It suffices to show that $1 < w$ implies $1 < \varphi(w)$. The hypothesis means that $\varphi(x_i)x_i^{-1}$ is in the commutator subgroup $[F_n, F_n]$. By the lemma above, $\mu([F_n, F_n])$ lies in the subgroup $1 + O(2)$ of G. Write $\mu(\varphi(x_i)) = 1 + X_i'$. The Magnus expansion satisfies

$$\mu(\varphi(x_i)) \in (1 + O(2))(1 + X_i) = 1 + X_i + O(2),$$

and therefore $X_i' \in X_i + O(2)$. Now if w is a word in the free group F_n satisfying $w > 1$, its image under φ has Magnus expansion obtained from that of w by replacing

each occurence of X_i by some element of $X_i + O(2)$. It follows that the first non-zero non-constant terms of $\mu(w)$ and $\mu(\varphi(w))$ are identical, and $1 < w$ is equivalent to $1 < \varphi(w)$. □

We will touch briefly on an equivalent approach to this same ordering of F_n. Fox defined linear mappings $\partial/\partial x_i : \mathbb{Z}F_n \to \mathbb{Z}F_n, \quad i = 1, \ldots, n$, which are derivations (we will avoid the details here). There is, moreover, an augmentation map $\epsilon : \mathbb{Z}F_n \to \mathbb{Z}$. One of the utilities of these maps is that they give the coefficients of the Magnus expansion of a word $w \in F_n$. The coefficient of $X_{i_1} \cdots X_{i_r}$ in $\mu(w)$ is given by the appropriate rth partial derivative, followed by the augmentation, that is:

$$\epsilon\Big(\frac{\partial^r w}{\partial x_{i_1} \cdots \partial x_{i_r}}\Big).$$

Thus one could mechanize the ordering by comparing these calculations, beginning with first partials, then second partials, etc. until they differ for two input words, then compare them according to that coefficient.

9.2.4. Ordering P_n. — We now have the ingredients for bi-ordering P_n inductively. Certainly, P_1, which is $\{1\}$, and P_2, which is isomorphic to $\mathbb{Z}$, are bi-orderable.

Recalling the exact sequence

$$1 \longrightarrow F_n \xrightarrow{i} P_{n+1} \xrightarrow{r} P_n \longrightarrow 1,$$

we need to determine the conjugation action of P_{n+1} upon F_n.

Proposition 9.2.6. — *If $\varphi : F_n \to F_n$ is defined by $\varphi(x) = \beta x \beta^{-1}$, where x lies in F_n and β lies in P_{n+1}, then, for each $i = 1, \cdots, n$, there exists an element w_i in F_n (depending on β) satisfying*

$$\varphi(x_i) = w_i x_i w_i^{-1},$$

and, therefore, φ_{ab} is the identity.

Proof. — Note that, although F_n is normal in P_{n+1}, and P_{n+1} is normal in B_{n+1}, F_n is NOT normal in B_{n+1}. For example, F_n is not closed under conjugation by σ_n. However, conjugation by B_n does leave F_n invariant. For $i < n$, the generator σ_i acts as follows

$$\sigma_i(x_i) = x_i x_{i+1} x_i^{-1}, \quad \sigma_i(x_{i+1}) = x_i, \quad \sigma_i(x_j) = x_j \text{ for } \ j \neq i, i+1,$$

as is evident from the pictures of Figure 9.3. This is exactly the classical Artin representation B_n into $\mathrm{Aut}(F_n)$ as described in Chapter 5; each braid β in B_n sends x_i to some conjugate of $x_{\pi(i)}$, where π is the permutation corresponding to β. In particular, for $\beta \in P_n$, then $\beta(x_i)$ has the required form. Since P_{n+1} is generated by P_n and F_n, the proposition is proved. □

We are now ready to complete a new proof of the first part of Theorem I.2, namely of the bi-orderability of P_n.

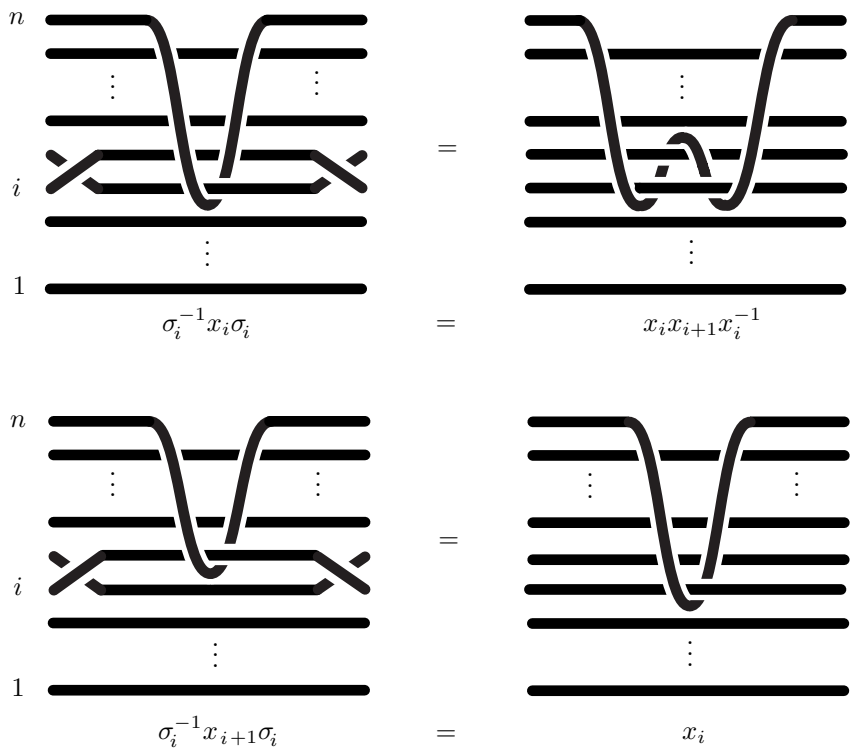


FIGURE 9.3. Conjugation action of σ_i on x_i $(= x_{i,n})$ and x_{i+1}, generators of F_n

Proof. — We now complete the inductive ordering of the pure braid groups. Assuming P_n has been bi-ordered, and using the bi-ordering of F_n, which by Proposition 9.2.5 is invariant under the action of P_{n+1}, we have, from Lemma 9.2.1, a bi-invariant ordering of P_{n+1}. □

The following completes Theorem I.2, and gives some additional information.

Proposition 9.2.7. — *The ordering described has the following properties:*
(i) *The ordering of P_{n+1} extends that of P_n, thus giving a bi-invariant ordering of P_∞;*
(ii) *For $n \geqslant 3$ it is a dense ordering;*
(iii) *If β lies in $P_n \cap B_n^+$, then we have $1 < \beta$ or $1 = \beta$;*
(iv) *The set $P_n \cap B_n^+$ is well-ordered, for each finite n.*

Proof. — Point (i) follows from the inductive definition. In fact, the ordering of P_{n+1} is just lexicographic ordering of the Artin components $(\beta_1, \cdots, \beta_n)$ of a pure braid β, and the inclusion $P_n \hookrightarrow P_{n+1}$, in Artin coordinates, is

$$(\beta_1, \cdots, \beta_n) \mapsto (\beta_1, \cdots, \beta_n, 1).$$

For (ii) it is enough to find a sequence of pure braids which converge to the identity. This can be accomplished by taking terms deeper in the lower central series of the free group F_{n-1}.

Part (iii) will be proved by induction on n. The statement is obvious for $n = 1$ and also clear for $n = 2$: elements of $P_2 \cap B_2^+$, are powers of σ_1^2, and $1 < \sigma_1^2 = x_{1,1}$. Inductively, suppose that (iii) is true for $P_n \cap B_n^+$ and let β be a braid in $P_{n+1} \cap B_{n+1}^+$. Now β has an expression in the σ_i with only positive exponents. Consider the braid $r(\beta)$ in P_n. Note that $r(\beta)$ is obtained from β by erasing the last strand, and therefore it, too, is expressed in the σ_i with only positive exponents, but with i strictly less than n. Therefore, by induction, we have $1 < r(\beta)$ (case (1)) or else $r(\beta) = 1$ (case (2)). In case (1), $1 < \beta$ by definition of the ordering. In case (2), β is an element of F_n; in fact the first n strands do not cross each other. We can read the expression for β_n (which equals β) in terms of the free generators of F_n from the places where the last strand passes under the first n strands. Each time it passes under the ith strand going upward *resp.* downward), we count $x_{n,i}$ with exponent $+1$ (*resp.* -1), by our convention. We see from this that β_n is a word in $x_{n,1}, \ldots, x_{n,n}$ without negative exponents and at least one positive exponent for $\beta \neq 1$. It follows that its Magnus expansion has positive leading coefficient (occurring at a linear term). Therefore we have $1 < \beta$ and the proof of (iii) is complete.

We now turn to the proof of (iv), again an induction. The statement is trivial for $n = 1$ and clear for $n = 2$. Consider a non-empty subset S of $P_{n+1} \cap B_{n+1}^+$, and assume (iv) is true for n. We need to show S has a least element. The set $r(S) = \{r(\beta) \in P_n\,;\, \beta \in S\}$, being a subset of $P_n \cap B_n^+$, has a least element (by induction); call it α_0. Let

$$S_0 = \{\gamma \in S\,;\, r(\gamma) = \alpha_0\};$$

clearly if there is a least element of S_0 it is also the least element of S. Note that the coordinates of γ in S_0 in the semidirect product of P_n and F_n are $\gamma = (\alpha_0, {\alpha_0}^{-1}\gamma)$. We now appeal to Lemma 9.2.8 to find a $\gamma_0 \in S_0$ whose last coordinate ${\alpha_0}^{-1}\gamma_0$ is minimal. Then γ_0 is the least element of S, establishing (iv). □

In the following, we refer to the ordering we have defined on the free group F_n, via the Magnus expansion μ, as the Magnus ordering. The notation F_n simultaneously refers to the subset of P_{n+1} which is the kernel of the retraction map $r : P_{n+1} \to P_n$.

Lemma 9.2.8. — *Consider a subset T of F_n of the form $T = \{{\alpha_0}^{-1}\gamma\,;\, \gamma \in S_0\}$, where S_0 is some nonempty subset of $P_{n+1} \cap B_{n+1}^+$ and α_0 is a fixed pure braid in P_n. Then T has a least element in the Magnus ordering of F_n.*

Proof. — The condition ${\alpha_0}^{-1}\gamma \in F_n$ is equivalent to the equation $r(\gamma) = \alpha_0$. As discussed above, the exponents of $x_{i,n}$ of a braid in F_n can be read from the linking numbers of the last strand of the braid with the ith strand, and those exponents are

non-negative if the braid is also in B^+_{n+1}. For $\gamma \in P_{n+1}$ and $\alpha_0 \in P_n$, those linking numbers for ${\alpha_0}^{-1}\gamma$ coincide with the analogous linking numbers of the strands of γ. This implies that the coefficients $(q_1, \dots, q_n)$ of the linear terms in the Magnus expansion of all elements of T are all non-negative. It follows that there is a lexicographically minimal one, say $(\widehat{q}_1, \dots, \widehat{q}_n)$. Let

$$T' = \{\tau \in T\,;\, \mu(\tau) \in 1 + \widehat{q}_1 X_1 + \cdots + \widehat{q}_n X_n + O(2)\}.$$

Then T' is nonempty, and its elements are strictly smaller than the other elements of T. We now claim that T' is finite. It follows that T' has a least element, which is therefore the least element of T. To verify the claim, observe that the exponent sum (as a word in $x_1, \dots, x_n$, where we have dropped the second subscript) of every element of T' is $\widehat{q}_1 + \cdots + \widehat{q}_n$. Each x_i has exponent sum $+2$ when expanded in the braid generators σ_j. The σ-exponent sum is an invariant of braids, and we conclude that for every ${\alpha_0}^{-1}\gamma$ in T', the σ-exponent sum of γ is exactly $2\sum_i \widehat{q}_i$ plus the σ-exponent sum of α_0. There are only finitely many distinct γ in B^+_{n+1} satisfying this, and the lemma follows. □

Note that our ordering is *not* a well-ordering of $P_\infty \cap B^+_\infty$, as we have the infinite descending sequence $\sigma_1^2 > \sigma_2^2 > \sigma_3^2 > \cdots$. Presumably this could be remedied, as with the canonical ordering of B_n, by a reversing trick, but we will not pursue this here. We conclude this section with a comparison of our ordering of P_n and certain partial braid orderings in the literature. Garside [**66**] defined a partial order $<_G$ on the semigroup B^+_n by defining $\alpha <_G \beta$ to be true if $\alpha\gamma = \beta$ holds for some γ in B^+_n. It is noted in [**57**] that this defines a lattice order on the set $\{\alpha \in B^+_n\,;\, \alpha <_G \Delta_n\}$ (connected with the Bruhat ordering on $\mathfrak{S}_n$). From Proposition 9.2.7(iii) and invariance under multiplication we immediately conclude the following.

Proposition 9.2.9. — *Our linear order on P_n extends the Garside partial order on $P_n \cap B^+_n$, i.e., $\alpha <_G \beta$ implies $\alpha < \beta$ for α, β in $P_n \cap B^+_n$.*

Elrifai and Morton [**55**] defined a partial order on B_n which is related to the Garside partial ordering. However, our ordering of P_n does *not* extend the Elrifai–Morton ordering, when restricted to P_n. We refer the interested reader to [**83**] for details.

9.3. Incompatibility of the orderings: local indicability

The purpose of this section is to argue that bi-orderings of P_n cannot be constructed so that they extend to left-invariant orderings of B_n.

Proposition 9.3.1. — *For $n \geqslant 5$, there is no left-invariant ordering of B_n which restricts to a bi-ordering on P_n.*

This argument, which appears in [**126**], makes use of a definition of Higman [**74**], who was motivated in part by the zero-divisor conjecture.

Definition 9.3.2. — A group G is *locally indicable* if every nontrivial finitely-generated subgroup H has an infinite cyclic quotient, or equivalently, there is a nontrivial map $H \to \mathbb{Z}$.

Proposition 9.3.3. — *Bi-orderable groups are locally indicable. Locally indicable groups are left-orderable. Neither of these implications can be reversed.*

Proof. — Let $H = \langle h_1, \cdots, h_r \rangle$ be a finitely generated subgroup of the bi-ordered group G. We may assume the set of generators is minimal and that $1 < h_1 < \cdots < h_r$ holds. One argues that the subgroup K of H, which is the union of convex subgroups which do not contain h_r, is normal and convex in H, and that the quotient H/K embeds in the positive additive reals (a generalization of Hölder's theorem, see [**26**]). Then it is easy to construct a nontrivial map from H to $\mathbb{Z}$. The proof that locally indicable groups are left-orderable is a result of Burns and Hale; we refer the reader to [**19**]. To see that the first implication is not reversible, one can cite the Klein bottle group (noting that local indicability is preserved under extensions). The reader might check that the braid groups B_3 and B_4 are also locally indicable but not bi-orderable. The first examples of left-orderable groups which are not locally indicable were published by Bergman [**5**]. The universal cover $\widetilde{SL(2,\mathbb{R})}$ of $SL(2,\mathbb{R})$ is one of his examples. □

According to [**126**], one can strengthen the first part of the above proposition, which we state without proof.

Proposition 9.3.4. — *If $(G, <)$ is a left-ordered group and H is a finite-index subgroup of G such that the ordering is bi-invariant when restricted to H, then G is locally indicable.*

The following together with Proposition 9.3.4 complete a proof of Proposition 9.3.1.

Proposition 9.3.5. — *For $n \geqslant 5$, the braid group B_n is not locally indicable.*

Proof. — This follows from a calculation of Goren and Lin [**71**], that the commutator subgroup $[B_n, B_n]$ is finitely-generated and perfect (equal to its own commutator subgroup) for $n \geqslant 5$. Any map of $[B_n, B_n]$ to an abelian group must be trivial, so B_n is not locally indicable. □

The pure braid group can be regarded as the fundamental group of the complement of the family of hyperplanes $z_i = z_j$ in the space $\mathbb{C}^n$ with coordinates $z_1, \ldots, z_n$. Our analysis of orderability applies to many other (but not all) complex hyperplane

arrangements. In particular, one can argue that the complement of a hyperplane arrangement of "fibre type" has biorderable fundamental group. For further details and a recent discussion of the fundamental groups of hyperplane arrangements, see [**115**].

CHAPTER 10

OPEN QUESTIONS

In this chapter we gather some open questions connected with the various aspects of braid orderings considered in this book. We refer to [**40**] and [**39**] for many others.

We should start, however, with a very general remark. There are many approaches to braid groups that have not been considered in this book. In fact, braid groups play a rôle in many areas of mathematics that have not even been mentioned (e.g., algebraic geometry or mathematical physics). We can therefore still hope that new, illuminating, perspectives on braid orderings will emerge in the future.

10.1. Well-ordering on positive braids

The result that the restriction of the linear ordering $<$ to B_n^+ is a well-ordering may be considered the deepest result known so far in the domain, but it remains mysterious in many ways.

10.1.1. Self-distributive algebra. — Lemma 2.1.18 tells us that, for every braid β in B_n and every left cancellative LD-system S, there exists at least one sequence $\vec{x}$ in S^n such that the action of β on $\vec{x}$ is defined, *i.e.*, $\vec{x} \cdot \beta$ exists. Let us reverse the point of view, and, starting with a sequence $\vec{x}$, consider the set $D(S, \vec{x})$ of all braids β such that $\vec{x} \cdot \beta$ exists. Thus β belongs to $D(S, \vec{x})$ if and only if $\vec{x} \cdot w$ exists for at least one braid word w representing β.

If S is a rack, then, for every sequence $\vec{x}$ in S^n, $D(S, \vec{x})$ is all of B_n. On the other hand, if S is the LD-system $(B_\infty, *)$, we can for instance consider the set $D(B_\infty, (1, \dots, 1))$ consisting of those braids β for which $(1, \dots, 1) \cdot \beta$ is defined. It is easy to see that $D(B_\infty, (1, \dots, 1))$ is a proper subset of B_n: for instance, σ_i^{-1} belongs to $D(B_\infty, (1, \dots, 1))$ for no i. We have seen in Section 2.2 that every braid in $D(B_\infty, (1, \dots, 1))$ necessarily admits a decomposition as a shifted product of special braids, and it can be shown that this condition is also sufficient. In the general case of an arbitrary left cancellative LD-system S, the action of positive braids on S^n

is always defined, so every set $D(S, \vec{x})$ with $\vec{x} \in S^n$ includes B_n^+. In some cases studied in [**89**], $D(S, \vec{x})$ coincides with B_n^+, and, then, the restriction of the braid order $<$ to $D(S, \vec{x})$ is a well-ordering.

Conjecture 10.1.1. — *(Laver's conjecture, braid form) For every sequence $\vec{x}$ in B_∞^n, the subset $D(B_\infty, \vec{x})$ of B_n is well ordered by $<$.*

10.1.2. Finite trees. — Using the coding of positive braid words by uniform trees allows us to associate with every positive braid a well defined ordinal that precisely characterizes its rank in the well ordering $(B_n^+, <)$.

Question 10.1.2. — How can one efficiently compute this ordinal in practice?

A general solution is proposed in [**16**], but it is algorithmically efficient only in the case of 3 strands, and further investigations are needed in the general case. Determining the rank of a braid word among all braid words is easy, but, in order to determine the rank of an (irreducible) braid word among irreducible braid words, one has to count how many reducible words are to be removed, a quite complicated question in the general case.

10.1.3. Hyperbolic geometry. — We have seen in Chapter 7 how to define an infinite family of distinct left-invariant orderings on B_n whose restriction to B_n^+ is a well-ordering, but we did not address the determination of the length of that well-ordering. The following question may well be quite easy to answer:

Question 10.1.3. — What is the order type of $(B_n^+, <)$, where $<$ is an ordering of Nielsen-Thurston type as alluded above?

10.1.4. Further applications. — Understanding the well-ordering better could pave the way to new applications, in particular along the lines already sketched in Subsection 1.3.3: after fixing some canonical way of decomposing every braid into a fraction consisting of two positive braids, we can define a well-ordering on arbitrary braids. Then each set of braids, for instance each conjugacy class in B_n, possesses a unique least element.

Question 10.1.4. — Can one compute this least element in a conjugacy class, or, equivalently, the associated pair of ordinals?

A similar question can be raised with "Markov equivalence class" replacing "conjugacy class"; a solution would associate a pair of ordinals with each knot.

Another natural question involving knots is:

Question 10.1.5. — For a given knot K, is there a $<$-minimal braid β of minimal braid index such that K is the closure of β?

10.2. Finding σ-positive representatives

Property **C** asserts that every nontrivial braid admits at least one representative braid word that is σ-positive, σ-negative, or the empty word.

Conjecture 10.2.1. — *For any $n > 3$ there exist numbers $c(n), c'(n)$ such that any n-strand braid with a representative word of length ℓ has a σ-positive or σ-negative representative of length at most $c(n) \cdot \ell$. Moreover, such a representative word can be found by an algorithm whose running time is bounded by $c'(n) \cdot \ell^2$.*

It is even conceivable that this statement holds with $c = \frac{7}{5}$, independent of n. However, the bound of $\frac{7}{5}$ is the best one can hope for, since the braid $\sigma_1\sigma_2\sigma_3^{-1}\sigma_2\sigma_1^{-1}$ has no σ-positive representative of length less than 7 (see [**60**], Theorem 5.1).

We have described several algorithmic methods for finding σ-positive representatives in practice: each of the proofs of Property **C** sketched in this text actually leads to an effective solution in theory. Some solutions turn out to be rather inefficient. For instance, the solution associated with braid colourings and self-distributive algebra in Chapter 2 has a huge complexity: the only proved upper bound is a tower of exponentials of exponential height! Similarly, the proof of Property **C** explained in Chapter 4 leads in theory to an algorithmic method for finding σ-positive or σ-negative representatives, but very little is known about the complexity of the method, for, as mentioned above, little is known about practically computing the rank of a braid.

Currently, the only known upper bounds for the length of a σ-positive or σ-negative braid word equivalent to a given word w are those given by Proposition 3.2.2 (handle reduction), and by Proposition 6.2.6 (the topological approach). Both of them are exponential with respect to the length of w,

As we shall see, the above conjecture is closely related to some quite far-reaching conjectures about Artin groups and mapping class groups.

Remark 10.2.2. — If, instead of finding an explicit σ-positive or σ-negative representative braid word for a given braid β, we consider the simpler question of deciding whether $\beta > 1$ is true, then—at least from a theoretical point of view—the best methods known so far are the algorithms described in Chapter 8, namely the one based on Lee Mosher's proof that mapping class groups are automatic [**129**] and the one using laminations. Both have a quadratic complexity.

10.2.1. Handle reduction. — We have seen that handle reduction, as described in Chapter 3, is a very efficient solution in practice. However, there remains a large gap between the complexity bound established in Proposition 3.2.2 and the experimental values of Tables 3.1 and 3.2. This suggests that the argument of Section 3.2 is far from optimal, but, despite serious efforts, nobody has yet been able to prove a polynomial bound on the complexity of this algorithm.

Conjecture 10.2.3. — *The computation time required for the handle reduction algorithm depends quadratically on the length of the input braid word w (for each fixed width). Moreover, the length of the output braid word w' is linearly bounded by the length of w.*

Clearly, Conjecture 10.2.3 implies Conjecture10.2.1. The second statement in Conjecture 10.2.3 would be a consequence of a positive solution to the following more general conjecture about word reversing—which extends without change to many group presentations, see [**42**]:

Conjecture 10.2.4. — *If w is a braid word of width n and length ℓ, and w' is a freely reduced braid word obtained from w by reversing and one-sided equivalence, then the length of w' is at most $n^2\ell$.*

10.2.2. Relaxation. — The method of useful arcs explained in Chapter 6 (as well as Larue's method of Chapter 5 which is essentially equivalent) is not efficient in practice, but we have described in Section 6.3 an improved method called relaxation algorithm. However, once again, the exact complexity of the method is still unknown. The following conjecture is supported by extensive computer experiments with some hundred million random braid words.

Conjecture 10.2.5. — *The computation time required for the relaxation algorithm depends quadratically on the length of the input braid word w. Moreover, the length of the output braid word w' is linearly bounded by the length of w.*

Remark 10.2.6. — Analogous relaxation algorithms can be constructed by using laminations instead of curve diagrams. The natural measure of complexity of a lamination is its *length*, defined by the formula $b_1 + 2b_2 + \cdots + nb_n$ in the parametrization of (8.5.10), provided that the lamination lies in the orbit of $(0, 1, \ldots, 0, 1)$ under action of B_n. The same questions as for curve diagrams apply in this case.

It seems likely that the output braid words of the relaxation algorithms represent $(\lambda, 1)$-quasigeodesics in the Cayley graph of B_n, where the factor λ is a positive integer that depends only on the number of strings n. (And indeed the same should be true for the handle-reduction algorithm, when applied to a quasigeodesic input braid word.) Unfortunately, we are currently far from having a complete understanding of quasi-geodesics in mapping class groups. A very large class of quasi-geodesics was constructed by Masur and Minsky [**105**].

10.2.3. Complexity issues. — For several of the algorithms presented in this book, the best upper bounds on the computational complexity that have been proved so far are far worse than the conjectured optimal bounds. One may hope that this is the manifestation of some deep and yet unknown aspect of the geometry of braids.

It is easy to compare the combinatorial algorithms one to the other, as was done in Table 3.1 for handle reduction *vs.* greedy normal form. As already noted several times, handle reduction proves to be very efficient in practice.

Making similar comparisons for other algorithms, notably the laminations algorithm in chapter 8 and the relaxation algorithm in chapter 6, is not so easy. The reason is that these algorithms are based on a completely different principle: they do not work by a step–by–step modification of a given braid word. Moreover, in contrast to the handle reduction algorithm or the classical Garside algorithms, they require long integer arithmetic. Thus the efficiency of any actual implementation would depend heavily on the quality of the long integer tools used. This makes any comparison with the other algorithms questionable.

In the case of the lamination algorithm, the upper bound of Proposition 8.5.11 is likely to be sharp, and the method should be efficient in practice. However, no experiment has been made so far to estimate the parameters.

As for the relaxation algorithm, experiments have only been made for relatively short braids (of length at most 50), for which no long integer arithmetic was required. These experiments suggest (see [**143**]) that the length of the output braid word may be bounded linearly by the length of the input braid word. If true, this would imply a quadratic bound on the computational complexity. To give an idea of the performance, a C implementation of the standard relaxation algorithm, a random braid word of length 50 in B_6 took on average 0.15 millisec. on a Pentium 4 Processor at 1.5 GHz. The output braid was on average of length 45, and in the worst case (in a sample of 40.000.000) of length 178. We remark that the relaxation algorithm uses a version of the laminations algorithm as its first step.

Let us finally observe that the practical efficiency of the above algorithms could make them convenient for cryptographical applications of braids, such as those considered in [**84**]. In particular, the encoding of B_n into $\mathbb{Z}^{2n}$ given by Formulas (8.5.14) could be used to define a perfect collision-free hash-function on B_n.

10.3. Topology of the space of orderings

Recently, Adam S. Sikora [**135**] proposed to study the topology of the space of left-invariant linear orderings on a (semi)group. Let G be a semigroup. Let us denote by $\mathrm{LO}(G)$ the set of all left-invariant orderings on G equipped with the smallest topology for which all sets of the form $U_{a,b} = \{x \in \mathrm{LO}(G)\,;\, a <_x b\}$, with $a, b \in G$, are open. We also denote by $\mathrm{LWO}(G)$ the topological subspace of $\mathrm{LO}(G)$ consisting of well-orderings. It can be proved that, for any countable semigroup G the topological space $\mathrm{LO}(G)$ is compact and totally discontinuous, a result that gives serious restrictions on topological spaces that can appear as $\mathrm{LO}(G)$ for some semigroup G. If one shows that $\mathrm{LO}(G)$ is non-empty and perfect, then this implies that the space $\mathrm{LO}(G)$ is homeomorphic to the Cantor set. In [**135**] this was shown to be the case

for $G = \mathbb{Z}^n$ or $\mathbb{N}^n$: each of the spaces $\mathrm{LO}(\mathbb{Z}^n)$, $\mathrm{LO}(\mathbb{N}^n)$, $\mathrm{LWO}(\mathbb{N}^n)$ is homeomorphic to the Cantor set. Let us mention that Sikora used the compactness of $\mathrm{LWO}(\mathbb{N}^n)$ for a very short and nice interpretation of the existence of the universal Gröbner basis in the polynomial ring $k[x_1, \ldots, x_n]$.

The perfectness of the space $\mathrm{LO}(G)$ can be characterized in the following way. $\mathrm{LO}(G)$ is perfect if and only if for any $a_1, b_1, \ldots, a_n, b_n \in G$ there exist either no or infinitely many orderings x such that $a_i <_x b_i$ for all $i = 1, \ldots, n$. In other words, $\mathrm{LO}(G)$ is *not* perfect if and only if it is either empty or there is an ordering in $\mathrm{LO}(G)$ that can be defined by fixing the relative order of finitely many elements.

It is also noticed in [**135**] that $\mathrm{LWO}(G)$ is not always a closed subset of $\mathrm{LO}(G)$, which was true in the case $G = \mathbb{N}^n$.

So, for any group G generated by finitely many elements $g_1, \ldots g_n$ and admitting infinitely many left-invariant linear orderings it is natural to ask the following questions: (i) What is the topology of $\mathrm{LO}(G)$? (ii) What is the topology of $\mathrm{LO}(G^+)$, where G^+ is the submonoid of G generated by $g_1, \ldots, g_n$? (iii) Is $\mathrm{LWO}(G^+)$ a closed subset in $\mathrm{LO}(G^+)$ and what is the topology of $\mathrm{LWO}(G^+)$?

Even for the free group with more than one generator all three questions are open so far. As we saw in Chapter 7, the braid group B_n admits infinitely many left-invariant linear orderings, some of which become well-orderings when restricted to B_n^+. So, we would like to set the questions above as an open problem for the group B_n.

Conjecture 10.3.1. — *For $n \geqslant 3$, the spaces $\mathrm{LO}(B_n)$, $\mathrm{LO}(B_n^+)$, and $\mathrm{LWO}(B_n^+)$ are homeomorphic to the Cantor set.*

10.4. Generalizations and extensions

The discovery of a left-invariant order on Artin braid groups naturally leads to the much more general question of whether the numerous groups appearing in low-dimensional topology are orderable.

10.4.1. Generalized braid groups. — As was mentioned in Section 6.3, the notion of a braid group can be extended to any surface.

Question 10.4.1. — If $\mathcal{S}$ is a compact orientable surface without boundary, is the surface braid group $B_n(\mathcal{S})$ left-orderable?

This is a natural question, since the corresponding *pure* surface braid groups *are* bi-orderable—this was shown by Juan Gonzalez-Meneses [**70**], using an extension of the techniques presented in Chapter 9.

10.4.2. Torelli groups. — The *Torelli group* of a surface $\mathcal{S}$ is defined to be the subgroup of $\mathcal{MCG}(\mathcal{S})$ of those elements which act trivially on the homology $H_1(\mathcal{S})$. In the special case of a genus-two surface, the Torelli group is an infinitely generated free group, and hence bi-orderable.

Question 10.4.2. — Is the Torelli group of a closed surface $\mathcal{S}$ without boundary left-orderable, or even bi-orderable?

Even stronger, one might ask about the orderability properties of the kernel of the action on $H_1(\mathcal{S}, \mathbb{Z}/p)$, where p is a prime. (For a proof that this group is torsion-free see [**78**], Chapter 1.) Currently no finite-index subgroup of $\mathcal{MCG}(\mathcal{S})$ is known to be left-orderable.

10.4.3. Surface and 3-manifold groups. — It is shown in [**127**] that the fundamental group (or, equivalently, the one-string braid group) of any compact surface, except for the projective plane $\mathbb{RP}^2$, is left-orderable. Moreover, with the further exception of the Klein bottle, all surface fundamental groups are actually bi-orderable.

The situation is more subtle when considering the case of fundamental groups of compact 3-manifolds, which we will refer to simply as 3-manifold groups. A study of these groups is initiated in the paper [**10**], where necessary and sufficient conditions are derived for the left-orderability and bi-orderability of fundamental groups of the important class of Seifert-fibred 3-manifolds (manifolds which are foliated by topological circles). It is also shown there that for each of the eight 3-dimensional geometries, there exist manifolds modelled on that geometry which have left-orderable group and also there exist examples whose groups are not left-orderable.

Recall that a 3-manifold is called irreducible if every smooth 2-sphere bounds a 3-ball in the manifold. An important general result of [**10**] is that all compact irreducible orientable 3-manifolds with positive first Betti number have left-orderable groups. In particular, all knot and link groups are left-orderable.

Question 10.4.3. — Given an automorphism $\varphi : G \to G$ of a surface group (or more generally of any bi-orderable group), under what conditions does there exist a bi-ordering of G which is φ-invariant, meaning $x < y$ implies $\varphi(x) < \varphi(y)$?

This is relevant to the study of 3-manifolds which are bundles over S^1, with surface fibres. If φ is the monodromy associated with such a fibration, then a φ-invariant bi-ordering of the fibre's group naturally leads to a bi-ordering of the fundamental group of the total space, and vice versa. In [**120**], this observation, as well as the techniques described in chapter 9, are used to prove that certain fibred knots with pseudo-Anosov monodromy have bi-orderable groups. By contrast, the group of any torus knot cannot be bi-ordered, because they contain elements which do not commute, while a power of one of those elements commutes with the other, which cannot occur in a bi-orderable group.

Conjecture 10.4.4. — *If G is the fundamental group of a closed orientable (irreducible) 3-manifold, then G is virtually bi-orderable, i.e., there exists a subgroup of finite index which is bi-orderable.*

It is shown in [**10**] that Conjecture 10.4.4 holds for Seifert-fibred 3-manifolds, and more generally for all manifolds with a geometric structure, except possibly hyperbolic manifolds. We do not even know if hyperbolic manifold groups are virtually left-orderable.

To put the difficulty of these questions into perspective, we point out that from general properties of orderable groups and covering space theory one can show that any 3-manifold satisfying Conjecture 10.4.4 also satisfies a certain well-known conjecture in 3-manifold theory; this conjecture states that any closed, orientable, irreducible 3-manifold $\mathcal{M}$ with infinite fundamental group has a finite-sheeted cover $\widetilde{\mathcal{M}}$ with positive first Betti number. In particular, such a manifold $\mathcal{M}$ is Haken, and hence satisfies Thurston's geometrization conjecture.

10.4.4. Artin–Tits groups. — Another natural generalisation of braid groups are, of course, Artin–Tits groups (also called Artin groups) and, more generally, Garside groups (see [**42**]).

Question 10.4.5. — Which Artin–Tits groups are left-orderable or bi-orderable?

Currently, the only Artin–Tits groups known to be left-orderable are those that embed in mapping class groups. Among the finite Coxeter type ones, these are all but those of type E_6, E_7, and E_8 [**141, 121**]. Among the infinite Coxeter type ones, there is one well-known family of groups that are biorderable [**52**], namely the right-angled Artin–Tits groups, which have only commutation relations—also called partially commutative groups. Indeed, these groups embed in pure surface braid groups [**27**].

BIBLIOGRAPHY

[1] S.I. Adyan – "Fragments of the word Delta in a braid group", *Mat. Zam. Acad. Sci. SSSR* **36** (1984), no. 1, p. 25–34, (Russian); English translation in *Math. Notes of the Acad. Sci. USSR* **36** (1984), no. 1, p. 505–510.

[2] E. Artin – "Theorie der Zöpfe", *Abh. Math. Sem. Univ. Hamburg* **4** (1925), p. 47–72.

[3] ———, "Theory of braids", *Ann. of Math.* **48** (1947), p. 101–126.

[4] V.G. Bardakov – "On the theory of braid groups", *Mat. Sb.* **183** (1992), no. 6, p. 3–42, (Russian. English summary); English translation in *Acad. Sci. Sb. Math.* **76** (1993), no. 1, p. 123–153.

[5] G. Bergman – "Right orderable groups which are not locally indicable", *Pacific J. Math.* **147** (1991), p. 243–248.

[6] S. Bigelow – "Braid groups are linear", *J. Amer. Math. Soc.* **14** (2001), no. 2, p. 471–486.

[7] J. Birman – "On braid groups", *Comm. Pure Appl. Math.* **22** (1969), p. 41–72.

[8] ———, *Braids, Links, and Mapping Class Groups*, Annals of Math. Studies, vol. 82, Princeton Univ. Press, 1974.

[9] N. Bourbaki – *Algèbre, chapitres I–III*, Hermann, Paris, 1970.

[10] S. Boyer, D. Rolfsen & B. Wiest – "Orderable 3-manifold groups", preprint, 2001.

[11] E. Brieskorn – "Automorphic sets and braids and singularities", in *Braids*, Contemporary Mathematics, vol. 78, American Mathematical Society, 1988, p. 45–117.

[12] E. Brieskorn & K. Saito – "Artin-Gruppen und Coxeter-Gruppen", *Invent. Math.* **17** (1972), p. 245–271.

[13] W. Burau – "Über Zopfgruppen and gleichsinnig verdrillte Verkettungen", *Abh. Math. Sem. Univ. Hamburg* **11** (1936), p. 179–186.

[14] S. Burckel – "L'ordre total sur les tresses positives", Ph.D. Thesis, Université de Caen, 1994.

[15] ———, "The wellordering on positive braids", *J. Pure Appl. Algebra* **120** (1997), no. 1, p. 1–17.

[16] ———, "Computation of the ordinal of braids", *Order* **16** (1999), p. 291–304.

[17] ———, "Syntactical methods for braids of three strands", *J. Symb. Comput.* **31** (2001), p. 557–564.

[18] G. Burde & H. Zieschang – *Knots*, Studies in Mathematics, vol. 5, de Gruyter, 1985.

[19] R. Burns & V. Hale – "A note on group rings of certain torsion-free groups", *Canad. Math. Bull.* **15** (1972), p. 441–445.

[20] J.W. Cannon, W.J. Floyd & W.R. Parry – "Introductory notes on Richard Thompson's groups", *Enseign. Math.* **42** (1996), p. 215–257.

[21] P. Cartier – "Développements récents sur les groupes de tresses, applications à la topologie et à l'algèbre", in *Séminaire Bourbaki*, Astérisque, vol. 189-190, Société Mathématique de France, 1990, exposé 716 (novembre 1989), p. 17–67.

[22] A. Casson & S. Bleiler – *Automorphisms of surfaces after Nielsen and Thurston*, LMS student texts, vol. 9, Cambridge University Press, 1988.

[23] R. Charney – "Artin groups of finite type are biautomatic", *Math. Ann.* **292** (1992), no. 4, p. 671–683.

[24] R. Charney, J. Meier & K. Whittlesey – "Bestvina's normal form complex and the homology of Garside groups", Preprint, 2001.

[25] W.L. Chow – "On the algebraic braid group", *Ann. of Math.* **49** (1948), p. 654–658.

[26] P.F. Conrad – "Right-ordered groups", *Michigan Math. J.* **6** (1959), p. 267–275.

[27] J. Crisp & B. Wiest – "Graph braid groups, surface groups, graph groups and surface braid groups", in preparation.

[28] M. Davis – "Non-positive curvature and reflection groups", p. 373–422, North Holland, 2002.

[29] P. Dehornoy – "Infinite products in monoids", *Semigroup Forum* **34** (1986), p. 21–68.

[30] ———, "Free distributive groupoids", *J. Pure Appl. Algebra* **61** (1989), p. 123–146.

[31] ———, "Sur la structure des gerbes libres", *C. R. Acad. Sci. Paris Sér. I Math.* **309** (1989), p. 143–148.

[32] ———, "Deux propriétés des groupes de tresses", *C. R. Acad. Sci. Paris Sér. I Math.* **315** (1992), p. 633–638.

[33] ———, "Structural monoids associated to equational varieties", *Proc. Amer. Math. Soc.* **117** (1993), no. 2, p. 293–304.

[34] ———, "Braid groups and left distributive operations", *Trans. Amer. Math. Soc.* **345** (1994), no. 1, p. 115–151.

[35] ———, "The structure group for the associativity identity", *J. Pure Appl. Algebra* **111** (1996), p. 59–82.

[36] ———, "Weak faithfulness properties for the Burau representation", *Topology Appl.* **69** (1996), p. 121–143.

[37] ———, "A fast method for comparing braids", *Adv. in Math.* **125** (1997), p. 200–235.

[38] ———, "Groups with a complemented presentation", *J. Pure Appl. Algebra* **116** (1997), p. 115–137.

[39] ———, "Strange questions about braids", *J. Knot Th. and its Ramifications* **8** (1999), no. 5, p. 589–620.

[40] ———, *Braids and Self-Distributivity*, Progress in Math., vol. 192, Birkhäuser, 2000.

[41] ———, "Construction of self-distributive operations and charged braids", *Internat. J. Algebra and Comput.* **10** (2000), no. 1, p. 173–190.

[42] ———, "Groupes de Garside", *Ann. scient. Éc. Norm. Sup.* 4e *série* **35** (2002), p. 267–306.

[43] ———, "Study of an identity", *Algebra Universalis* **48** (2002), p. 223–248.

[44] ———, "Thin groups of fractions", in *Combinatorial and Geometric Group Theory*, Contemporary Mathematics, vol. 296, American Mathematical Society, 2002, p. 95–128.

[45] P. Dehornoy & Y. Lafont – "Homology of Gaussian groups", *Ann. Inst. Fourier (Grenoble)*, to appear.

[46] P. Dehornoy & L. Paris – "Gaussian groups and Garside groups, two generalisations of Artin groups", *Proc. London Math. Soc. (3)* **79** (1999), no. 3, p. 569–604.

[47] P. Deligne – "Les immeubles des groupes de tresses généralisés", *Invent. Math.* **17** (1972), p. 273–302.

[48] F. Digne – "Artin groups of finite type are linear", Preprint, 2001.

[49] R. Dougherty – "Critical points in an algebra of elementary embeddings", *Ann. Pure Appl. Logic* **65** (1993), p. 211–241.

[50] R. Dougherty & T. Jech – "Finite left-distributive algebras and embedding algebras", *Adv. in Math.* **130** (1997), p. 201–241.

[51] A. Drápal – "Persistence of cyclic left-distributive algebras", *J. Pure Appl. Algebra* **105** (1995), p. 137–165.

[52] G. Duchamp & J.-Y. Thibon – "Simple orderings for free partially commutative groups", *Internat. J. Algebra Comput.* **2** (1992), no. 3, p. 351–355.

[53] I. Dynnikov – "Integral laminations and braid ordering", in preparation.

[54] ———, "On a Yang-Baxter mapping and the Dehornoy ordering", *Uspekhi Mat. Nauk* **57** (2002), no. 3, p. 151–152, (Russian); English translation in *Russian Math. Surveys* **57** (2002), no. 3.

[55] E.A. El-Rifai & H.R. Morton – "Algorithms for positive braids", *Quart. J. Math. Oxford Ser. (2)* **45** (1994), no. 2, p. 479–497.

[56] D. Epstein – "Curves on 2-manifolds and isotopies", *Acta Math.* **115** (1966), p. 83–107.

[57] D. Epstein, J.W. Cannon, D.F. Holt, S.V.F. Levy, M.S. Paterson & W.P. Thurston – *Word Processing in Groups*, Jones and Bartlett Publ., 1992.

[58] P. Etingof, T. Schedler & A. Soloviev – "Set-theoretical solutions to the quantum Yang-Baxter equation", *Duke Math. J.* **100** (1999), no. 2, p. 169–209.

[59] M. Falk & R. Randell – "The lower central series of a fiber-type arrangement", *Invent. Math.* **82** (1985), p. 77–88.

[60] R. Fenn, M.T. Greene, D. Rolfsen, C. Rourke & B. Wiest – "Ordering the braid groups", *Pacific J. Math.* **191** (1999), p. 49–74.

[61] R. Fenn & C.P. Rourke – "Racks and links in codimension 2", *J. Knot Th. and its Ramifications* **1** (1992), p. 343–406.

[62] V.V. Fock – "Dual Teichmüller spaces", http://front.math.ucdavis.edu/dg-ga/9702018.

[63] H. Friedman – "Higher set theory and mathematical practice", *Ann. Math. Logic* **2** (1971), p. 325–357.

[64] ———, "On the necessary use of abstract set theory", *Adv. in Math.* **41** (1981), p. 209–280.

[65] J. Funk – "The Hurwitz action and braid group orderings", *Theory and Applic. of Categories* **9** (2001), no. 7, p. 121–150.

[66] F.A. Garside – "The braid group and other groups", *Quart. J. Math. Oxford Ser. (2)* **20** (1969), p. 235–254.

[67] K.F. Gauss – *Handbuch 7*, Univ. Göttingen collection.

[68] E. Ghys – "Groups acting on the circle", *Enseign. Math.* **47** (2001), no. 2, p. 329–407.

[69] J. González-Meneses – "Presentations of surface braid groups", *J. Knot Th. and its Ramifications*, to appear.

[70] ———, "Ordering pure braid groups on compact, connected surfaces", *Pacific J. Math.* **203** (2002), p. 369–378.

[71] E.A. Gorin & V.Ya. Lin – "Algebraic equations with continuous coefficients, and certain questions of the algebraic theory of braids", *Math. USSR Sbornik* **78** (1969), no. 120, p. 579–610.

[72] A. Hatcher & W. Thurston – "A presentation for the mapping class group of a closed orientable surface", *Topology* **19** (1980), no. 3, p. 221–237.

[73] M. Hertweck – "A counterexample to the isomorphism problem for integral group rings", *Ann. of Math.* **154** (2001), p. 115–136.

[74] G. Higman – "The units of group rings", *Proc. London Math. Soc. (3)* **46** (1940), no. 2, p. 231–248.

[75] ———, "Ordering by divisibility in abstract algebras", *Proc. London Math. Soc. (3)* **2** (1952), p. 326–336.

[76] O. Hölder – "Die Axiome der Quantität und die Lehre vom Mass", *Math.-Phys. Kl* **53** (1901), p. 1–64.

[77] J.E. Humphreys – *Reflection groups and Coxeter groups*, Cambridge Studies in Advanced Mathematics, vol. 29, Cambridge University Press, Cambridge, 1990.

[78] S. Ivanov – *Subgroups of Teichmüller modular groups*, Translations of Mathematical Monographs, vol. 115, American Mathematical Society, Providence, RI, 1992.

[79] V. Jones – "Hecke algebra representations of braid groups and link polynomials", *Ann. of Math.* **126** (1987), p. 335–388.

[80] D. Joyce – "A classifying invariant of knots: the knot quandle", *J. Pure Appl. Algebra* **23** (1982), p. 37–65;.

[81] A. Kanamori – *The Higher Infinite*, Perspectives in Mathematical Logic, Springer Verlag, 1994.

[82] C. Kassel – "L'ordre de Dehornoy sur les tresses", in *Séminaire Bourbaki*, Astérisque, vol. 276, Société Mathématique de France, 2002, exposé 865 (novembre 1999), p. 7–28.

[83] D.M. Kim & D. Rolfsen – "An ordering of pure braids and hyperplane arrangements", *Canad. J. Math.*, to appear.

[84] K.H. Ko, S. Lee, J.H. Cheon, J.W. Han, J. Kang & C. Park – "New public-key cryptosystem using braid groups", in *Proc. Crypto 2000*, Lecture notes in Comput. Sci., vol. 1880, Springer Verlag, 2000, p. 166–184.

[85] D. Krammer – "The braid group B_4 is linear", *Invent. Math.* **142** (2000), p. 451–486.

[86] ———, "Braid groups are linear", *Ann. of Math.* **155** (2002), no. 1, p. 131–156.

[87] R.H. La Grange & A.H. Rhemtulla – "A remark on the group rings of order preserving permutation groups", *Canad. Math. Bull.* **11** (1968), p. 679–680.

[88] D.M. Larue – "Left-distributive and left-distributive idempotent algebras", PhD. Thesis, University of Colorado, Boulder, 1994.

[89] ———, "On braid words and irreflexivity", *Algebra Universalis* **31** (1994), p. 104–112.

[90] R. Laver – "Elementary embeddings of a rank into itself", *Abstracts Amer. Math. Soc.* **7** (1986), p. 6.

[91] ———, "The left distributive law and the freeness of an algebra of elementary embeddings", *Adv. in Math.* **91** (1992), no. 2, p. 209–231.

[92] ———, "A division algorithm for the free left distributive algebra", in *Logic Colloquium '90* (Oikkonen & *al*, eds.), Lect. notes in Logic, vol. 2, Springer Verlag, 1993, p. 155–162.

[93] ———, "On the algebra of elementary embeddings of a rank into itself", *Adv. in Math.* **110** (1995), p. 334–346.

[94] ———, "Braid group actions on left distributive structures and well-orderings in the braid group", *J. Pure Appl. Algebra* **108** (1996), no. 1, p. 81–98.

[95] A. Levy – *Basic Set Theory*, Springer Verlag, 1979.

[96] P.A. Linnell – "Zero divisors and $L^2(G)$", *C. R. Acad. Sci. Paris Sér. I Math.* **315** (1992), no. 1, p. 49–53.

[97] P.A. Linnell & T. Schick – "Finite group extensions and the Atiyah conjecture", Preprint.

[98] D. Long & M. Paton – "The Burau representation is not faithful for $n \geqslant 6$", *Topology* **32** (1993), no. 2, p. 439–447.

[99] J.H. Lu, M. Yan & Y.C. Zhu – "On the set-theoretical Yang-Baxter equation", *Duke Math. J.* **104** (2000), no. 1, p. 1–18.

[100] W. Magnus, A. Karrass & D. Solitar – *Combinatorial group theory*, J. Wiley and sons, New York, 1966.

[101] A.I. Malcev – "On the embedding of group algebras in division algebras", *Doklady Akad. Nauk SSSR (N.S.)* **60** (1948), p. 1499–1501.

[102] A.V. Malyutin – "Orderings on braid groups, operations on closed braids, and confirmation of Menasco's conjecture", *Topology and dynamics* **267** (2000), p. 163–169, (Russian); in memory of V. A. Rokhlin (St. Petersburg, 1999). Zap. Nauchn. Sem. POMI.

[103] ———, "Fast algorithms for the recognition and comparison of braids", *Zap. Nauchn. Sem. POMI* **279** (2001), p. 197–217, (Russian).

[104] ———, "Dehornoy ordering and Markov-Birman-Menasco operations", preprint, 2002.

[105] H. Masur & Y. Minsky – "Geometry of the complex of curves II: hierarchical structure", *GAFA, Geom. funct. anal.* **10** (2000), p. 902–974.

[106] S.V. Matveev – "Distributive groupoids in knot theory", *Math. Sbornik* **119** (1982), no. 1-2, p. 78–88.

[107] R. McKenzie & R.J. Thompson – "An elementary construction of unsolvable word problems in group theory", in *Word Problems* (Boone & *al*, eds.), Studies in Logic, vol. 71, North Holland, 1973, p. 457–478.

[108] J. Michel – "A note on words in braid monoids", *J. Algebra* **215** (1999), p. 366–377.

[109] J. Moody – "The Burau representation of the group B_n is unfaithful for large n", *Bull. Amer. Math. Soc. (N.S.)* **25** (1991), no. 2, p. 379–384.

[110] L. Mosher – "Mapping class groups are automatic", *Ann. of Math.* **142** (1995), p. 303–384.

[111] B.H. Neumann – "On ordered division rings", *Trans. Amer. Math. Soc.* **66** (1949), p. 202–252.

[112] J. Nielsen – "Untersuchungen zur Topologie des geschlossenen zweiseitigen Flächen", *Acta Math.* **50** (1927), p. 189–358.

[113] ———, *Collected Mathematical Papers, edited by V.L. Hansen*, Birkhäuser, Boston-Basel-Stuttgart, 1986.

[114] S.Yu. Orevkov – "Strong positivity in the right-invariant order on a braid group and quasipositivity", *Mat. Zametki* **68** (2000), no. 5, p. 692–698, (Russian); English translation in *Math. Notes* **68** (2000), no. 5-6, 588-593.

[115] L. Paris – "On the fundamental group of the complement of a complex hyperplane arrangement", in *Singularities and Arrangements, Sapporo and Tokyo, 1998*, Adv. Stud. Pure Math., vol. 27, Kinokuniya, 2000, p. 257–272.

[116] ———, "Artin monoids embed in their groups", *Comment. Math. Helv.* **77** (2002), no. 3, p. 609–637.

[117] D.S. Passman – *The Algebraic Structure of Group Rings*, Pure and Appl. Math, Wiley Interscience, 1977.

[118] R.C. Penner – "The decorated Teichmüller space of punctured surfaces", *Comm. Math. Phys.* **113** (1987), no. 2, p. 299–339.

[119] R.C. Penner & J.L. Harer – *Combinatorics of train tracks*, Annals of Math. Studies, vol. 125, Princeton Univerity Press, 1992.

[120] B. Perron & D. Rolfsen – "Ordering groups of fibred knots", *Math. Proc. Cambridge Philos. Soc.*, to appear.

[121] B. Perron & J. Vannier – "Groupe de monodromie géométrique des singularités simples", *Math. Ann.* **306** (1996), no. 2, p. 231–245.

[122] M. Picantin – "The center of thin Gaussian groups", *J. Algebra* **245** (2001), no. 1, p. 92–122.

[123] ———, "The conjugacy problem in small Gaussian groups", *Comm. Algebra* **29** (2001), no. 3, p. 1021–1038.

[124] J. Przytycki – "Classical roots of knot theory", *Chaos, Solitons and Fractals* **9** (1998), no. 4, 5, p. 531–545.

[125] K. Reidemeister – *Knotentheorie*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 1, Julius Springer, Berlin, 1932, English translation: Knot theory, BCS associates, Moscow, Idaho (1983).

[126] A. Rhemtulla & D. Rolfsen – "Local indicability in ordered groups: braids and elementary amenable groups", *Proc. Amer. Math. Soc.* **130** (2002), no. 9, p. 2569–2577.

[127] D. Rolfsen & B. Wiest – "Free group automorphisms, invariant orderings and applications", *Algebraic and Geometric Topology* **1** (2001), p. 311–320 (electronic).

[128] D. Rolfsen & J. Zhu – "Braids, orderings and zero divisors", *J. Knot Th. and its Ramifications* **7** (1998), p. 837–841.

[129] C. Rourke & B. Wiest – "Order automatic mapping class groups", *Pacific J. Math.* **194** (2000), no. 1, p. 209–227.

[130] G.P. Scott – "Braid groups and the group of homeomorphisms of a surface", *Math. Proc. Cambridge Philos. Soc.* **68** (1970), p. 605–617.

[131] H. Short & B. Wiest – "Orderings of mapping class groups after Thurston", *Enseign. Math.* **46** (2000), p. 279–312.

[132] W. Shpilrain – "Representing braids by automorphisms", *Internat. J. Algebra and Comput.* **11** (2001), no. 6, p. 773–777.

[133] H. Sibert – "Partial orderings in Artin groups", in preparation.

[134] ———, "Extraction of roots in Garside groups", *Comm. Algebra* (2002), no. 6, p. 2915–2927.

[135] A.S. Sikora – "Topology on the spaces of orderings of groups", Preprint, 2001.

[136] C. Squier – "The homological algebra of artin groups", *Math. Scand.* **75** (1994), p. 5–43.

[137] K. Tatsuoka – "An isoperimetric inequality for artin groups of finite type", *Trans. Amer. Math. Soc.* **339** (1993), no. 2, p. 537–551.

[138] W. Thurston – "Finite state algorithms for the braid group", Circulated notes, 1988.

[139] ———, "On the geometry and dynamics of diffeomorphisms of surfaces", *Bull. Amer. Math. Soc. (N.S.)* **19** (1988), no. 2, p. 417–431.

[140] M. Wada – "Group invariants of links", *Topology* **31** (1992), no. 2, p. 399–406.

[141] B. Wajnryb – "An elementary approach to the mapping class group of a surface", *Geometry and Topology* **3** (1999), p. 405–466.

[142] B. Wiest – "Dehornoy's ordering of the braid groups extends the subword ordering", *Pacific J. Math.* **191** (1999), p. 183–188.

[143] ———, "An algorithm for the word problem in braid groups", Preprint, 2002.

INDEX

Acyclicity Property, *see Property* **A**
Address (in a uniform tree), 82
Arrow (of an automaton), 142
Artin coordinates (of a braid), 164
Artin group, *see Artin–Tits group*
Artin–Tits group, 6
  spherical, 7
Automaton, 142

Blueprint, 44
Braid, 2
  fundamental, 6
  geometric, 1
  group, 1
    of a surface, 3
  isotopic, 2
  order, 8
  positive, 5
  pure, 2
  special, 34
Braid word
  absolute value, 59
  equivalent, 28
  σ_1-free, 8
  σ_1-negative, 8
  σ_1-positive, 8
  σ-negative, 8
  σ-positive, 8
  positively equivalent, 60
  reducible, 79, 84
  representative, 7
  reversible, 28
Breadth (of a braid word), 72

Cayley graph, 59
Code (of a braid word), 72, 74
Colouring, 31
Combinatorial type
  of a triangulation, 137
  of a flip, 145
Combing, 143
Companion, 81, 86
Comparison Property, *see Property* **C**
Conjugate orderings, 114
Convex subgroup, 115
Critical prefix, 63
Curve
  normal, 151
  reduced, 151
Curve diagram, 103

D-disk (of a pair of triangulations), 137
Decomposition (main), 83
Dehn half-twist, 110
Delta-length, 160
Dense ordering, 114
Diagram (on a surface), 5
Discrete ordering, 114
Division normal form, 50

Expression, 7

Flip (of an edge), 141

Garside group, 7
Geodesic
  filling, 123
  finite type, 123
  infinite type, 123
Group
  bi-orderable, vii
  left-orderable, vii
  left-well-orderable, vii
  locally indicable, 171
  orderable, see *bi-orderable*

Handle, 55
  permitted, 56
  reduction, 56
Height (of a braid word), 62

Iterated left divisor, 36

Labelled binary tree, 37
Lamination,
  geodesic, 129,
  integral, 152
    decorated, 152
LD-expansion, 37
LD-monoid, 49
LD-system, 23
  acyclic, 27
  free, 37
  ordered, 27
Leading part (of a triangulation), 143

Magnus map, 166
Mapping class group, 3
Mosher normal form, 145

Neighbour (in a tree), 83

Property **A**, 8
  proof of, 47, 95, 107, 160
Property $\mathbf{A}_i$, 10
Property $\mathbf{A}_{LD}$, 40
Property **C**, 9
  proof of, 40, 57, 79, 102, 107
Property $\mathbf{C}^+$, 9
Property $\mathbf{C}_n$, $\mathbf{C}_\infty$, 10
Property $\mathbf{C}_{LD}$, 36
Property **S**, 14
  proof of, 51, 89, 111, 130
Property $\mathbf{S}^+$, 14

Rack, 26
Regular language, 142
Relaxation, 112
Reversing, 28

Shift endomorphism, 8
State (of an automaton), 142
  accept, 142
  failure, 142
Subsurface sequence, 124
  conjugated, 125
  of infinite type, 129
Subword Property, *see Property* **S**

Term, 37
  LD-equivalent, 37
Torelli group, 179
Tree (n-uniform), 72
Triangulation, 136
  marked, 146
  ordered oriented, 142
  singular, 136
  tight, 138
  transverse, 137

Useful arc, 108

Width (of a braid word), 57
Word
  freely reduced, 94
  traced in a graph, 59

INDEX OF NOTATION

$\mathbb{N}$ (nonnegative integers)
$\mathbb{Z}$ (integers)
$\mathbb{R}$ (reals)

Introduction

B_n (braid group), vii
B_n^+ (braid monoid), vii
$\mathfrak{S}_n$ (symmetric group), vii
P_n (pure braid group), vii

Chapter 1

B_n, B_∞ (braid group), 1
D^2 (disk), 1
σ_i (braid), 2
P_n (pure braid group), 2
$B_n(\mathcal{S})$ (surface braid group), 3
D_n (punctured disk), 3
$\mathcal{MCG}(\mathcal{S}, \mathcal{P})$ (mapping class group), 3
B_n^+ (braid monoid), 5
Δ_n (fundamental braid), 6
$<$ (braid order), 8
sh (shift endomorphism), 8
$\mathbf{A}$ (Property), 8
$\mathbf{A}_i$ (Property), 10
$\mathbf{C}$ (Property), 9
$\mathbf{C}^+$ (Property), 9
$\mathbf{C}_n$, $\mathbf{C}_\infty$ (Property), 10
$\mathbf{S}$ (Property), 14
$\mathbf{S}^+$ (Property), 14
$<^\phi$ (well-order on B_∞^+), 16
ω (ordinal), 15

Chapter 2

(LD) (left self-distributivity law), 23
ε (empty word), 24
$\equiv$ (braid word equivalence), 28
$D(w)$ (denominator of a braid word), 29
$N(w)$ (numerator of a braid word), 29
$/$ (braid word complement), 29
$\widetilde{D}(w)$ (right denominator of a braid word), 30
$\widetilde{N}(w)$ (right numerator of a braid word), 30
$*$ (braid exponentiation), 34
B_{sp} (special braids), 34
$\prod^{sh}$ (shifted product), 35
$\sqsubset$ (iterated left divisor), 36
$\mathbf{A}_{LD}$ (Property), 40
$\mathcal{F}_n$ (free LD-system), 37
T_n (free magma), 37
$=_{LD}$ (LD-equivalence), 37
$\mathrm{left}^k(t)$ (iterated left subterm), 38
$\sqsubset_{LD}$ (LD-prefix), 38
$t^{[k]}$ (right power), 38
∂ (term derivation), 39
$\mathbf{C}_{LD}$ (Property), 36
LD_α (LD operator), 42
$\mathcal{G}_{LD}$ (geometry monoid), 42
G_{LD} (geometry group), 43
$\mathcal{G}_{LD}^+$, G_{LD}^+ (positive monoids), 43
χ_t (blueprint of a term), 44
$P_<$, $P_=$ (preordering on G_{LD}), 45

Chapter 3

$\Gamma(w)$ (Cayley graph), 59
$\overline{w}$ (class of w), 61
$\pi_p(w)$ (critical prefix), 63

Chapter 4

W_n (positive braid words), 71
$<^{\mathrm{ShortLex}}$ (ShortLex order), 75
$\prec$ (order), 75
$\mathrm{rk}(u)$ (rank of a word), 78
$\mathrm{red}(u)$ (reduction of a word), 79, 84

red$^*(u)$ (iterated reduction of a word), 80
$(\mathcal{S}_\rho)$ (induction clause), 80, 86
$W_{i,j}$, $W^\bullet_{i,j}$ (braid words), 83
$\sigma_{i,j}$ (braid word), 84
W_n^{irred} (irreducible words), 88

Chapter 5

F_n, F_∞ (free group), 93
$\widehat{\sigma}_i$, $\widehat{\beta}$ (automorphism of F_n), 93
$S(x)$ (words in free group), 94
sh (shifted automorphism), 94

Chapter 6

D_n (punctured disk), 103
$P_1, \dots, P_n$ (punctures), 103
$e_0, \dots, e_n$ (segments), 103
E (main diameter), 103
$<_{\mathrm{CD}}$ (braid ordering), 106

Chapter 7

$<_x$ (ordering), 114
$\widetilde{D_n}$ (universal cover), 115
$\mathbb{H}^2$ (hyperbolic plane), 115
S^1_∞ (circle at infinity), 115
γ_x (geodesic), 116
$\widehat{F_\infty}$ (completion), 132
$\lhd$ (ordering), 132

Chapter 8

S^2 (two sphere), 136
$\mathcal{P}$ (set of punctures), 136
A^* (set of all words), 142
$\mathcal{L}(S^2;\mathcal{P})$ (laminations), 152
$\widetilde{\mathcal{L}}(S^2;\mathcal{P})$ (decorated laminations), 153
$(\mathbb{Z}^T)_{\mathrm{even}}$ (even sequences), 153
$|w|_\Delta$ (Δ-length), 160
$\|\eta\|$ (norm of an integral point), 160

Chapter 9

P_n (pure braid group), 163
P_n^+ (positive pure braid monoid), 163
$x_{i,j}$ (generators of P_n), 164
μ (Magnus map), 166
$O(k)$ (terms of degree $\geqslant k$), 166

Chapter 10

$D(S,\vec{x})$ (action on LD-systems), 173
LO(G) (space of orders), 177
LWO(G) (space of well-orders), 177