UNIVERSITAT DE BARCELONA

Facultat de Matemàtiques
i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

The Burau Representation of the Braid Group

Autora: Raquel Revilla Bouso

Director: Dr. Ricardo García López
Realitzat a: Departament de Matemàtiques i Informàtica
Barcelona, 19 de gener de 2020

Abstract

This work deals with two relevant subjects of modern mathematics: representation theory and braid theory. It also includes a relation between them; the Burau representation, as well as a new concept; the Fox partial derivatives and its relation with the braid theory, more specifically with the Burau representation. The structure of the work is as follows. Firstly, the basic notions of representation and braid theory will be given to understand the following results. Therefore, the Burau representation will be studied giving its definition, the one of its reduced form and proving its faithfulness. A different interpretation of this representation, introduced by V. Jones will also be studied and it has to be outlined due to its relation with the probability. Finally, the Fox calculus and its partial derivatives will be introduced to show an alternative approach to the Burau representation.

Resum

Aquest treball presenta dos temes rellevants de la matemàtica moderna: la teoria de representacions i la teoria de trenes. També inclou una relació entre ambdós; la representació de Burau, així com un nou concepte; el de les derivades parcials de Fox i la seva relació amb la teoria de trenes, més exactament amb la representació de Burau. L'estructura del treball és la següent. Primerament, es donaran les nocions bàsiques tant de la teoria de representacions com de la de trenes per tal de poder comprendre i enunciar els resultats següents. A continuació, s'estudiarà la representació de Burau donant la seva definició, la de la seva forma reduïda i comprovant la seva fidelitat en alguns casos. Cal destacar que també s'observarà una interpretació diferent d'aquesta representació relacionada amb la probabilitat deguda a V. Jones. Finalment, s'acabarà amb la introducció del càlcul de Fox i de les seves derivades parcials per tal de donar una altre perspectiva per a definir la representació de Burau.

2010 Mathematics Subject Classification. 20F36, 20G05

Acknowledgements

I would like to express my deep gratitude to Dr. Ricardo García for his helpful advice and guidance and, also, for his dedication, support and patience during all these months. Without him this work would not have been possible.

I would also like to give special thanks to my friends and my family for supporting and encouraging me throughout my degree.

It has been a wonderful experience. Thank you.

Contents

Introduction **1**

1 Representation theory **3**

1.1 Basic language . . . 3

2 Braid theory **7**

2.1 Basic concepts . . . 7

2.2 Definition of the braid group . . . 12

3 The Burau representation **20**

3.1 Definition . . . 20

3.1.1 The reduced Burau representation . . . 21

3.2 A probabilistic interpretation . . . 24

3.3 Faithfulness of the Burau representation . . . 26

3.3.1 ψ_1 is faithful . . . 26

3.3.2 ψ_2 is faithful . . . 27

3.3.3 ψ_3 is faithful . . . 27

4 Free differential calculus **34**

4.1 The Fox partial derivatives . . . 34

4.1.1 The Chain Rule . . . 37

4.2 The Fox calculus and the Burau representation . . . 39

Conclusions **44**

Bibliography **45**

Introduction

A mathematician born in Germany called Emil Artin, in the early part of the 20th century, began to study what nowadays is known as braid theory. The first ideas were based on the physical and tractable nature of braids, but, over the course of the century, this theory gradually evolved and now it is recognized as a basic theory in mathematics, with applications in such branches as topology, category theory and algebraic geometry. The central object is the braid group with n strings.

On the other hand, representation theory is the part of mathematics that studies abstract algebraic structures by representing their elements as linear transformations of modules over a ring or of vector spaces. Thus, a representation makes an abstract object more concrete by describing its elements and operations in terms of matrices. This is most effective when the representation is faithful, what means that no information is lost when passing from the abstract object to the linear data.

So, one can ask which are the representations of the braid group, and, moreover, if there are faithful representations of it. In 1936, Werner Burau, also a German mathematician, introduced a remarkable representation of the braid group, nowadays known as the Burau representation.

In this work we introduce some basic notions of braid theory and representation theory, and we apply them to study some particular aspects of the Burau representation. We provide a detailed proof of the fact that the Burau representation is faithful if $n \leqslant 3$, and we roughly indicate how it is proved that for $n \geqslant 5$ it is not. The case $n = 4$ is still open at present. [2]

While the Burau representation can be defined very quickly just giving the matrices which correspond to each Artin generator, such a definition seems unnatural and its motivation is unclear. At the beginning of the '50s, Ralph Fox, an American

[2] It is know that there are faithful representations of all of the braid groups, but this is out of the scope of this work.

mathematician, developed in a series of five papers the so-called free differential calculus, also known as Fox calculus. He defined the Fox derivatives, which are an algebraic construction in the theory of free groups with some similarities to the conventional derivatives in calculus. The final goal of this work is to define the Burau representation in terms of the Fox calculus. This is not straightforward, but the definition obtained looks much more natural and gives another viewpoint on the representations of Burau.

This work is organized in three parts:

The first part covers chapters 1 and 2 and it is intended to give some basic concepts which will be a fundamental tool in the sequel. Chapter 1 introduces some definitions of representation theory and some examples. Special attention is payed to the definition of faithfulness and irreducibility due to the importance this notions will have in chapter 3. In chapter 2 we state three different definitions of the braid group. We begin with the geometrical definition of a braid with n strings, of braid equivalence and of the braid group and we study its operation and a few properties. We continue giving the definition of the braid group in algebraic terms, that is, in terms of a presentation with generators and relations, this is the Artin presentation of the braid group. Finally, we define braid automorphisms and we state an important theorem which relates braids with automorphisms of free groups.

The second part covers chapter 3 and it deals with the relation between the notions introduced in the two previous chapters. It starts defining the Burau representation of the braid group and proving that it is reducible. Then, we prove the faithfulness of the Burau representation of the braid group with three strings, this is the most difficult part. We also give the details of an observation of Vaughan Jones about a probabilistic (or as he says, mechanical) interpretation of the Burau representation for positive braids.

The last part of this work is chapter 4, where we introduce the Fox calculus and we explain in detail how the Burau representation can be defined in terms of it.

The fundamental bibliographical sources have been [3], [6], [7] and [8]. The remaining references have been used as secondary bibliography and the article mentioned is just referenced for completeness. The present work merely aspires to be an exposition and a synthesis of the different sources we have used, but, as far as possible, we have tried to give a slightly original perspective on the topics treated.

Chapter 1

Representation theory

This chapter presents some fundamental ideas and results of representation theory which will be important throughout the work.

1.1 Basic language

From now and on, R will always be a commutative, unitary ring, M a free R-module and we will denote by $GL(M)$ the group of its automorphisms. If a basis $e_1, \ldots, e_n$ of M as a free R-module is fixed, we can identify $GL(M)$ with $GL(n, R)$, the group of invertible $n \times n$ matrices with coefficients in R.

Definition 1.1. Let G be a group. A *linear representation of G defined over R* is a group homomorphism

$$\rho : G \longrightarrow GL(M),$$

The rank of M is called the rank of ρ or its dimension if R is a field.

Remark 1.2. If the ring R is clear from context, it is customary to just say that *ρ is a representation of G*. Similarly, when the homomorphism ρ is clear from context, one may say only that M is *a representation of G*.

Given a representation $\rho : G \longrightarrow GL(M)$ and an element $g \in G$, we usually write

$$\rho(g)\, v$$

for the image of $v \in M$ under the linear transformation $\rho(g)$. Such vectors are also sometimes called *G-translates of v* or, simply, *translates of v* when the context is clear. Similarly, when ρ is clearly understood, one may simply write

$$gv = \rho(g)\, v.$$

Representations exist in plenty, we will see some examples.

Example 1.3. The following examples are easily defined representations just to clarify the concept:

1. If we take $R = \mathbb{C}$, then we have the exponential $z \mapsto e^z$, which is a group homomorphism from $(\mathbb{C}, +)$ to $(\mathbb{C}^\times, \cdot)$, or in other words, to $GL(1, \mathbb{C}) = GL(\mathbb{C})$. This means the exponential is a one-dimensional representation of the additive group of the complex numbers.

2. The additive group of the real numbers admits a two-dimensional real representation:

$$\begin{aligned} (\mathbb{R}, +) &\longrightarrow GL(2, \mathbb{R}) \\ t &\longmapsto \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \end{aligned}$$

3. Any subgroup $GL(n, \mathbb{R})$ can be thought of as being given with a natural representation, for example:

   (a) $\mathfrak{S}_n \hookrightarrow GL(n, \mathbb{R})$, where $\mathfrak{S}_n$ denotes the symmetric group of n elements. If $e_1, \ldots, e_n$ is the canonical basis of $\mathbb{R}^n$, to each $\sigma \in \mathfrak{S}_n$, we associate the automorphism given by

$$\begin{aligned} \rho(\sigma) : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ e_i &\longmapsto e_{\sigma(i)}. \end{aligned}$$

   It is easy to see that ρ is a group monomorphism.

   (b) $O_n = \{A \in GL(n, \mathbb{R}) \mid A^T A = A A^T = Id\} \hookrightarrow GL(n, \mathbb{R})$

   (c) $SL_n = \{A \in GL(n, \mathbb{R}) \mid det(A) = 1\} \hookrightarrow GL(n, \mathbb{R})$

Definition 1.4. Let G be a group. A representation ρ of G defined over R is *faithfull* if ρ is injective, i.e., if $Ker(\rho) = \{1\}$.

Definition 1.5. Let G be a group. A R-representation ρ of G is *trivial* if $\rho(g) = 1_M$ is the identity map of M for all $g \in G$, i.e., if $Ker(\rho) = G$.

Remark 1.6. Sometimes only the representation of degree 1, with $M = R$, mapping g to $1 \in R^\times$ is called *"the" trivial representation*. We will denote by $\mathbf{1}$ this one-dimensional representation when G and R are clear from the context, or $\mathbf{1}_G$ if only R is.

Although there are "many" representations, many of them are actually equivalent. In other words, quite often, the representations of G over R can be classified in a useful way. To go into this, we must explain how to relate possibly different representations.

Definition 1.7. Let G be a group. A *morphism, or homomorphism, between representations* ρ_1 and ρ_2 of G, both defined over R and acting on free modules M_1 and M_2, respectively, is a R-linear map

$$\phi : M_1 \longrightarrow M_2$$

such that

$$\phi\left(\rho_1\left(g\right)v\right) = \rho_2\left(g\right)\left(\phi\left(v\right)\right) \in M_2,$$

for all $g \in G$ and $v \in M_1$.

Remark 1.8. One also says that ϕ *intertwines* ρ_1 and ρ_2 and one denotes this by $\phi : \rho_1 \to \rho_2$.

This definition is also better visualized as saying that, for all $g \in G$, the square diagram

$$\begin{tikzcd}
M_1 \arrow[r, "\phi"] \arrow[d, "\rho_1(g)"] & M_2 \arrow[d, "\rho_2(g)"] \\
M_1 \arrow[r, "\phi"] & M_2
\end{tikzcd}$$

of linear maps commutes or, even more briefly, by omitting the mention of the representations and writing

$$\phi\left(gv\right) = g\phi\left(v\right)$$

for $g \in G$, $v \in M_1$.

If a morphism of representation ϕ is bijective , its inverse ϕ^{-1} is also a morphism, between ρ_2 and ρ_1, and it is therefore justified to call ϕ an isomorphism between ρ_1 and ρ_2. Indeed, using the diagram above, we find that the relation

$$\rho_2\left(g\right) \circ \phi = \phi \circ \rho_1\left(g\right)$$

is equivalent, in that case, to

$$\phi^{-1} \circ \rho_2\left(g\right) = \rho_1\left(g\right) \circ \phi^{-1}.$$

Definition 1.9. Let G be a group and let $\rho : G \to GL(M)$ be a representation of G. If a free submodule $F \subset M$ is stable under all operators, i.e., $\rho(g)(F) \subset F$ for all $g \in G$, then the restriction of $\rho(g)$ to F defines a homomorphism

$$\tilde{\rho} : G \longrightarrow GL(F)$$

which is therefore a R-representation of G, and the inclusion map

$$i : F \hookrightarrow M$$

is a morphism of representations. One speaks, naturally, of a *subrepresentation* of ρ or, if the action is clear from the context, of M itself.

The following type of representation are the fundamental building blocks for representations in general.

Definition 1.10. Let G be a group. A representation ρ of G on a R-free module M is *irreducible* if and only if $M \neq \{0\}$ and there is no free submodule of M stable under ρ, except $\{0\}$ and M itself, that is to say that the only submodules $N \subset M$ such that for all $g \in G \ \ \rho(g)(N) \subset N$ are $N = \{0\}$ and $N = M$.

If a representation is not irreducible, it is called *reducible*.

Chapter 2

Braid theory

In this chapter we give some definitions of braid theory and we state some important results that will be useful for the following chapters. We have mainly consulted the references [6] and [8] to develop it.

2.1 Basic concepts

Definition 2.1. Let D be a unit cube, so $D = \{(x, y, z) \mid 0 \leq x, y, z \leq 1\}$. On the top face of the cube, place n points, $A_1, A_2, \ldots, A_n$ and, similarly, place n points on the bottom face, $B_1, B_2, \ldots, B_n$. Now, join the n points $A_1, A_2, \ldots, A_n$ with $B_1, B_2, \ldots, B_n$ by means of n segments or arcs $d_1, d_2, \ldots, d_n$. However, the arcs can only be attached in such a way that the following three conditions hold:

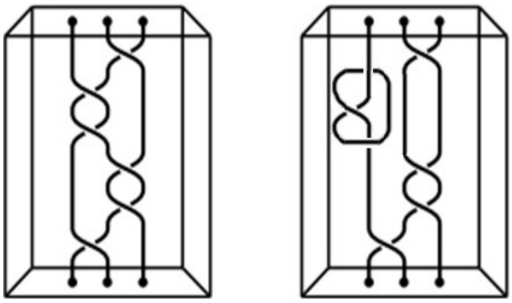
1. $d_1, d_2, \cdots, d_n$ are mutually disjoint.

2. Each d_i connects some A_j to some B_k where j and k may or may not be equal, but d_i is not permitted to connect A_j to A_k or B_j to B_k.

3. Each plane $E_s = \{z = s\}$ where $0 \leq s \leq 1$, in other words, each plane in D parallel to the xy-plane, intersects each arc d_i at one and only one point.

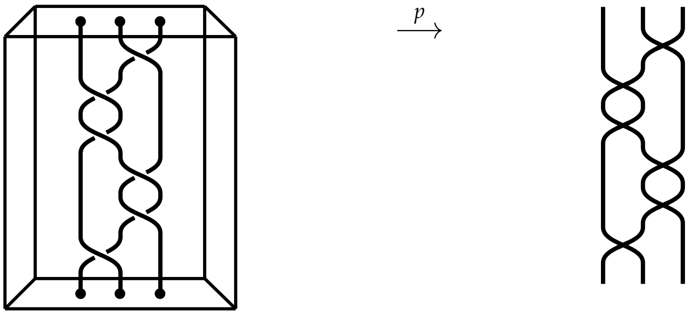
Such a configuration of n arcs $d_1, \ldots, d_n$ with end points $A_1, \ldots, A_n, B_1, \ldots, B_n$, is called *n-braid, or a braid with n strings*. As might be expected, d_i is called a *braid string* or equivalently the i^{th} *braid string*.

Example 2.2. The first figure is an example of a braid while the second one is an example of what can not be considered a braid.



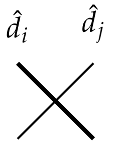
Let us describe an alternative way of representing a braid. Firstly, retract the cube D backwards onto the yz-plane by means of the projection, p, given by $p\,(x,y,z) = (0,y,z)$. Then we have a set of n simple curves, $d_1, d_2, \ldots, d_n$, on the yz-plane. We shall denote $p\,(\beta)$ by $\hat{\beta}$.

Example 2.3. In the following example it can be seen how a braid is projected:

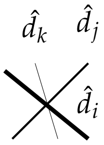


In addition, we may assume that the "new" curves $\hat{d}_i$ satisfy the following conditions:

1. $p\,(\beta)$ has at most a finite number of intersection points.
2. At most 2 distinct points of β are mapped onto the same point in $\hat{\beta}$. In such cases, Q is said to be a *double point* or *intersection point* of $\hat{\beta}$.



Double point

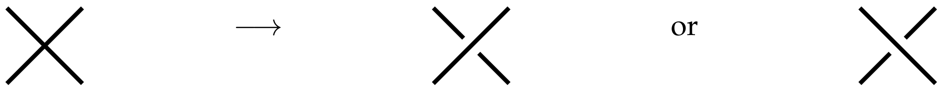


Ruled out by this condition

3. A vertex of β is never mapped onto a double point of $\hat{\beta}$.

A projection $\hat{\beta}$ that satisfies the above conditions is a *regular projection of* β.

Now, let β be a braid and $\hat{\beta}$ be a regular projection of β. In essence, $\hat{\beta}$ represents the braid except at the double points where it is not clear which string is in front of the other. To solve this problem, we cut away near a double point a small piece from either side of one of the strings as we can see in the following diagram:



So, a projection that has been altered in the above fashion is said to be a *regular diagram* or just simply a *diagram*.

Remark 2.4. It must be noted that these diagrams always exist. This is not a trivial result and the reader might consult the reference [6, Chapter 1, Subsection 1.2.2] for a proof.

Now, let $\hat{\beta}$ be a regular diagram of β.

A natural question to ask is, when are two braids regarded as the same or equivalent? The intuitive answer is when their differences can be removed in a reasonably straight-forward manner. So, what we need to find is either a move or a sequence of moves that will produce "essentially" the same braids.

Definition 2.5. Two n-braids, b and b', are *isotopic* if there is a continuous map $F : b \times I \to \mathbb{R}^2 \times I$ such that for each $s \in I$, the map $F_s : b \to \mathbb{R}^2 \times I$ sending $x \in b$ to $F(x, s)$ is an embedding whose image is a n-braid where $F_0 = id_b : b \to b$ and $F_1 = b'$. Each F_s automatically maps every endpoint of b itself.

In fact, isotopy is an equivalence relation. If two n-braids are equivalent, then we will treat them as if they are the same braid. We denote B_n the quotient of the set of n-braids by this equivalence relation, $\sim$.

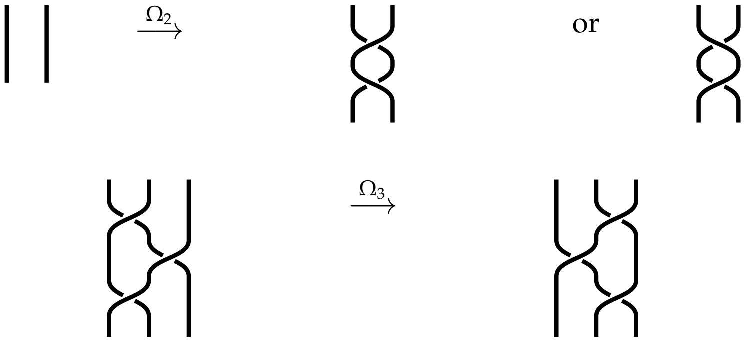
Now, we are going to provide two theorems to relate this concept with the idea of braid diagrams and with the polygonal image of the strings, respectively.

To state the first theorem, we will need the following previous concepts.

Definition 2.6. Two braid diagrams, $\hat{\beta}$ and $\hat{\beta}'$, with n strings are said to be *isotopic* if there is a continuous map $F : \hat{\beta} \times I \longrightarrow \mathbb{R} \times I$ such that for each $s \in I$ the set $\hat{\beta}_s = F\left(\hat{\beta} \times s\right) \subset \mathbb{R} \times I$ is a braid diagram with n strings, $\hat{\beta}_0 = \hat{\beta}$ and $\hat{\beta}_1 = \hat{\beta}'$. It is understood that F maps the crossings of $\hat{\beta}$ to the crossings of $\hat{\beta}_s$ for all $s \in I$

preserving the under or overgoing data. The family of braid diagrams $\{\hat{\beta}_s\} \in I$ is called an isotopy of $\hat{\beta}_0 = \hat{\beta}$ into $\hat{\beta}_1 = \hat{\beta}'$.

Definition 2.7. The following transformations of braid diagrams Ω_2 and Ω_3, as well as, the inverse transformations Ω_2^{-1} and Ω_3^{-1} are called *Reidemeister moves*.



The moves Ω_2 involves 2 strings and create two additional crossings (and there are 2 types of Ω_2-moves), while Ω_3 involves 3 strings and preserves the number of crossings.

Definition 2.8. Two braid diagrams, $\hat{\beta}$ and $\hat{\beta}'$, are *R-equivalent* if $\hat{\beta}$ can be transformed into $\hat{\beta}'$ by a finite sequence of isotopies and Reidemeister moves.

Now, we are ready to reformulate the notion of isotopic braids in topological terms.

Theorem 2.9. *Two braid diagrams present isotopic braids if and only if their diagrams are R-equivalent.*

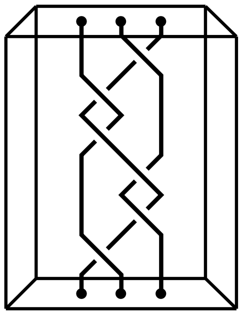
The proof of this theorem is far from the objective of this work. This is why we have not added it. [1]

To state the second theorem, let us denote the set of all n-braids by B_n, suppose D is a unit cube and within this cube there are n strings.

We will work with the polygonal image of a string, that is to say that any n-braid can be approximated by polygonal braids. [2] We redraw the braid in Example 2.2 with polygonal strings to clarify what we refer to.

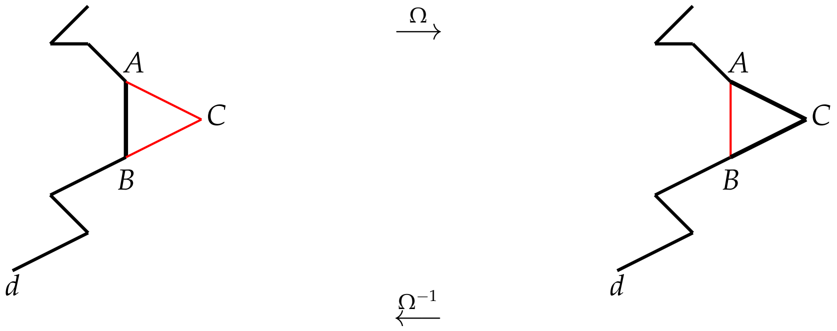
[1] It can be found in [6, Chapter 1, Theorem 1.6].

[2] We can find the proof in [6, Chapter 1, pg 10, step 2].



We now reformulate the notion of isotopy of braids in the polygonal setting. To this end, we introduce the following moves.

Definition 2.10. Let AB be an edge of a string d. Let C be a point in D such that the triangle ΔABC in D does not intersect any other strings and only meets d along AB. Further, suppose $AC \cup CB$ intersects every level plane E_s for $0 \leq s \leq 1$ at most at one point. If the above holds, we define an operation, which we shall denote by Ω, which replaces AB by the two edges $AC \cup CB$. We will also consider the inverse operation of Ω, namely, if $AC \cup CB$ is a part of a string and ΔABC does not intersect any other strings. Then, the operation, Ω^{-1}, replaces $AC \cup CB$ by the edge AB.



These are called *elementary moves* on a braid.

Theorem 2.11. *Two polygonal n-braids, b and b′, are isotopic if and only if b can be transformed into b′ by a finite number of elementary moves.*

The proof of this theorem is far from the objective of this work. This is why we have not added it. [3]

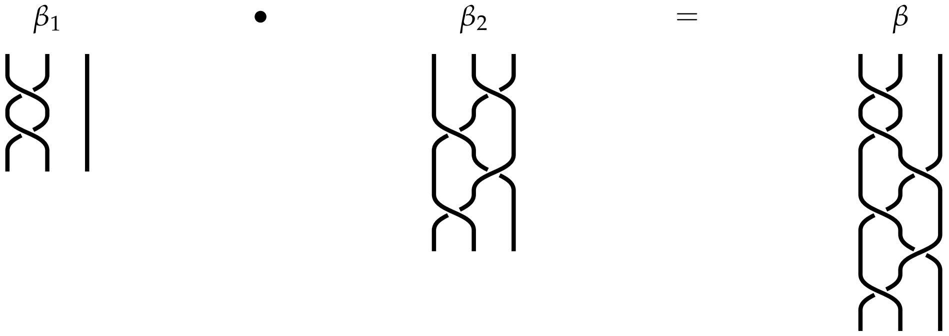
[3] It can be found in [6, Chapter 1, Claim 1.7].

2.2 Definition of the braid group

Now, we will denote a polygonal n-braid and, by a slight abuse of notation, also its diagram by β. In order to make the diagrams that we will draw easier to view, we shall draw these strings as smooth curves.

Definition 2.12. Let β_1 and β_2 be two n-braids in B_n. Identify the bottom edge of β_1 with the top edge of β_2 and remove only the identified edge. This provides us with a new n-braid which is defined to be the *product of the two n-braids* β_1 and β_2 and it is denoted by $\beta = \beta_1\beta_2$.

Example 2.13. Product of two braids:

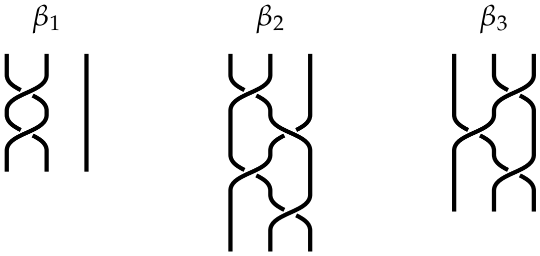


Proposition 2.14. *The product of braids is associative, i.e.,*

$$(\beta_1\beta_2)\,\beta_3 \sim \beta_1\,(\beta_2\beta_3).$$

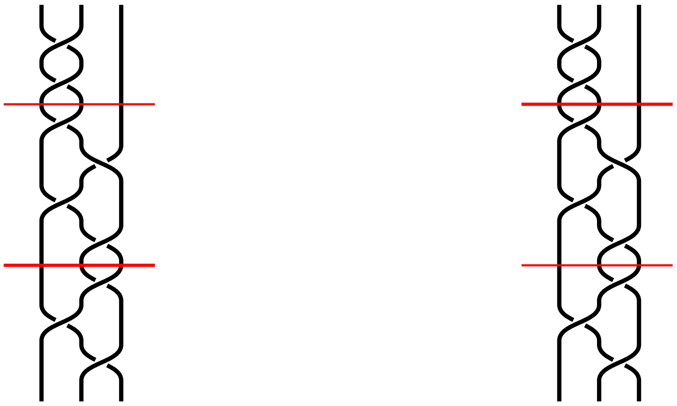
Althought the product of braids is associative, it need not to be commutative, i.e., $\beta_1\beta_2$ need not to be equivalent to $\beta_2\beta_1$.

Example 2.15. Consider the following braids,



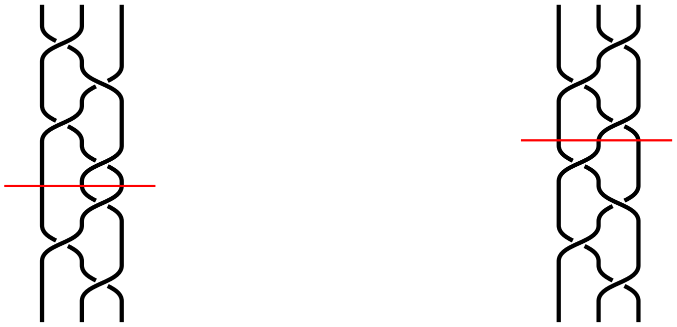
we can see that this braid product is associative:

$$(\beta_1\beta_2)\,\beta_3 \qquad \sim \qquad \beta_1\,(\beta_2\beta_3)$$



and, non commutative:

$$\beta_2\beta_3 \qquad \nsim \qquad \beta_3\beta_2$$



We can affirm that this two braids are not equivalent because we can check that, for example, in $\beta_2\beta_3$ the first string ends in the second position while in $\beta_3\beta_2$ it ends in the first position and the elementary moves do not touch where the braids start and finish. Then, $\beta_2\beta_3$ never can be transformed into $\beta_3\beta_2$ and they will never be equivalent.

Proposition 2.16. *Let e be the n-braid shown in the following diagram, where for $i = 1, 2, \ldots, n$ we join by a straight line segment the point A_i to the B_i. Then, for any n-braid β,*

$$\beta e \sim \beta \text{ and } \beta \sim \beta e$$

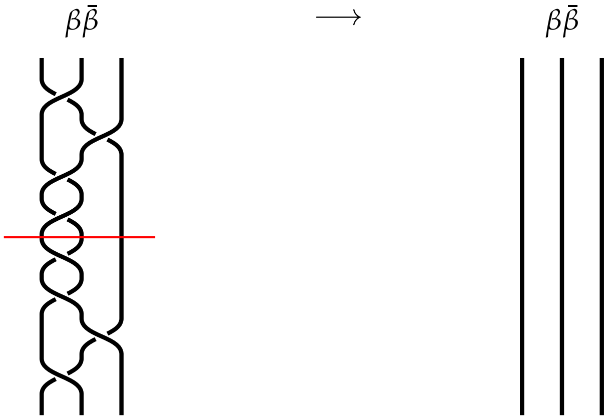
Such a braid e is called the *identity or trivial braid* and we shall denote it by 1_n.

A_1
A_2
...
A_{n-1}
A_n
B_1
B_2
...
B_{n-1}
B_n

Identity or trivial braid

Now, let β be a n-braid and let us construct a new n-braid $\bar{\beta}$ from β. Imagine that the bottom edge that contains β acts as a plane of a mirror. By taking the mirror image of β, we can construct a new n-braid, $\bar{\beta}$.

Example 2.17. In the following example it can be seen how a product of $\beta\bar{\beta}$ is equivalent to the identity braid by elementary moves.



Observe that the first string passes under the other ones and the second one passes under, over, over and under the third one what let us check that the second and third string can be separated. Then, we see how $\beta\bar{\beta}$ is transformed into the identity braid.

Proposition 2.18. *For each n-braid β, there exists a n-braid $\bar{\beta}$ such that*

$$\beta\bar{\beta} \sim 1_n \text{ and } \bar{\beta}\beta \sim 1_n.$$

For the proof we refer to [6, Chapter 1, Lemma 1.10].

Such a n-braid is called the *inverse of* β and denoted by β^{-1}.

Taking into account the definitions and propositions above and considering the equivalence classes of n-braids, we have all the necessary requirements for B_n to be a non-commutative group.

Theorem 2.19. *The set of equivalence classes of n-braid, B_n, with the product of two n-braids, forms a group. This group is usually called the n-braid group.*

Proof.

The binary operation is the product given by Definition 2.13; associativity is a consequence of Proposition 2.15; the identity element is 1_n given by Proposition 2.17 and the inverse element is $\bar{\beta}$ denoted by β^{-1} in Proposition 2.19. □

Next, we give a definition in purely algebraic terms of the n-braid group, B_n, that is in terms of generators and relations.

Definition 2.20. The *Artin braid group* $\widetilde{B_n}$ is the group generated by $n-1$ generators $\sigma_1,\ \sigma_2,\ \ldots,\ \sigma_{n-1}$ and the *braid relations*

$$\sigma_i\sigma_j = \sigma_j\sigma_i$$

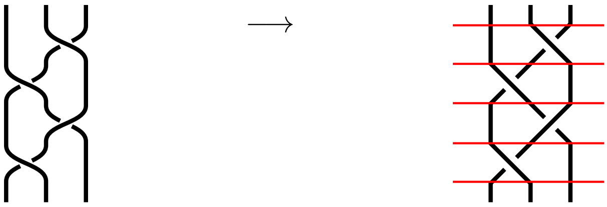
for all $i,\ j = 1,\ 2,\ \ldots,\ n-1$ with $|i-j| \geq 2$, and

$$\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}$$

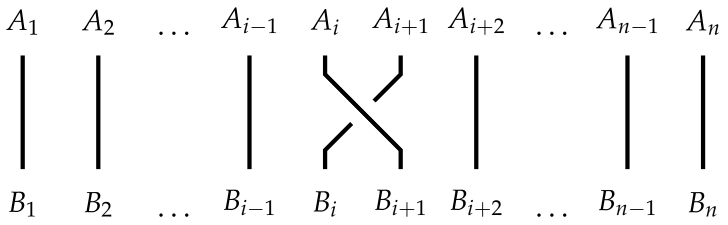
for $i = 1,\ 2,\ \ldots,\ n-2$.

The above generators $\sigma_1,\ \sigma_2,\ \ldots,\ \sigma_{n-1}$ arise from the following idea; taking a braid projection and diving it by means of level planes such that between two consecutive level planes only two strings are braided with a unique double point and the remaining strings hang vertically downwards.

Example 2.21. In the following example it can be seen how a braid is divided to obtain a partition of it.

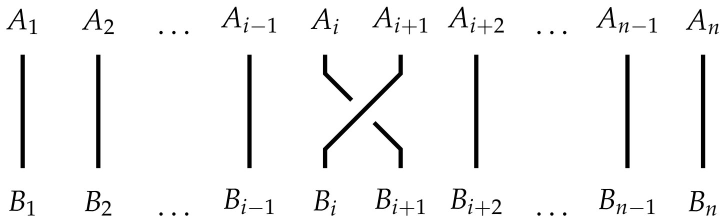


It is important to say that the configuration that lies between two consecutive planes is also, by our definition, a n-braid. For any n-braid there are $2n-2$ such possible generators and this set may be divided into two sets, each containing $n-1$ generators.


A_1 A_2 ... A_{i-1} A_i A_{i+1} A_{i+2} ... A_{n-1} A_n
B_1 B_2 ... B_{i-1} B_i B_{i+1} B_{i+2} ... B_{n-1} B_n

Generator σ_i

We shall denote these n-braids by $\sigma_1,\ \sigma_2,\ \ldots,\ \sigma_{n-1}$.


A_1 A_2 ... A_{i-1} A_i A_{i+1} A_{i+2} ... A_{n-1} A_n
B_1 B_2 ... B_{i-1} B_i B_{i+1} B_{i+2} ... B_{n-1} B_n

Generator σ_i^{-1}

And these by $\sigma_1^{-1},\ \sigma_2^{-1},\ \ldots,\ \sigma_{n-1}^{-1}$.

Due to Artin's seminal work on braids, this set of braids is sometimes referred to as *Artin braids* and the importance of these n-braids is that they generate the n-braid group, B_n.

Theorem 2.22. *For any $n \geq 1$ the n-braid group, B_n, is isomorphic to the Artin braid group, $\widetilde{B_n}$.*

Unfortunately, the proof of this theorem is beyond the scope of this work. This is why we have not added it. [4]

Remark 2.23. Notice that given a group homomorphism f from B_n to a group G, the elements $\{s_i = f(\sigma_i)\}_{i=1,\ \ldots,\ n-1}$ of G satisfy the braid relations of $\widetilde{B_n}$. And conversely, if $s_1,\ \ldots,\ s_{n-1}$ are elements of a group G satisfying the braid relations, then there is a unique group homomorphism $f : B_n \to G$ such that $s_i = f(\sigma_i)$ for all $i = 1,\ 2,\ \ldots,\ n-1$.

Finally, we give still one more way of seeing B_n. We realize the braid group as a group of automorphisms of the free group, F_n, on n generators $x_1,\ x_2,\ \ldots,\ x_n$.

[4] It can be found in [6, Chapter 1].

Definition 2.24. An automorphism φ of F_n is said to be a *braid automorphism* if it satisfies the following two conditions:

i) There is a permutation μ of the set $\{1,\ 2,\ \ldots,\ n\}$ such that $\varphi\left(x_k\right)$ is conjugate in F_n to $x_{\mu(k)}$ for all $k \in \{1,\ 2,\ \ldots,\ n\}$.

ii) $\varphi\left(x_1 x_2 \ldots x_n\right) = x_1 x_2 \ldots x_n$

It is easy to check that the following formulas define two mutually inverse braid automorphism $\check{\sigma}_i$ and $\check{\sigma}_i{}^{-1}$ of F_n for $i = 1,\ 2,\ \ldots,\ n-1$:

$$\check{\sigma}_i\left(x_k\right) = \begin{cases} x_k x_{k+1} x_k^{-1} & if \quad k = i \\ x_{k-1} & if \quad k = i+1 \\ x_k & \quad otherwise \end{cases}$$

$$\check{\sigma}_i{}^{-1}\left(x_k\right) = \begin{cases} x_{k+1} & if \quad k = i \\ x_k^{-1} x_{k-1} x_k & if \quad k = i+1 \\ x_k & \quad otherwise \end{cases}$$

Denote the set of braid automorphisms of F_n by $\check{B}_n$. It follows from the definitions that the inverse of a braid automorphism and the composition of two braid automorphisms are again braid automorphisms. Therefore, $\check{B}_n$ is a group with respect to the composition.

We now state the main theorem relating braids to braid automorphisms.

Theorem 2.25. *The formula $\sigma_i \mapsto \check{\sigma}_i$ with $i = 1,\ 2,\ \ldots,\ n-1$ defines a group isomorphism $B_n \to \check{B}_n$.*

Proof.

We will only prove that the formula $\sigma_i \mapsto \check{\sigma}_i$ defines a group homomorphism $B_n \to \check{B}_n$. The rest of the proof can be found in [6, Chapter 1, Theorem 1.31]. It must be taken into account that what we have here denoted by $\check{\sigma}_i$, in the just outlined reference is denoted by $\tilde{\sigma}_i{}^{-1}$.

Firstly, we prove $\check{\sigma}_i\check{\sigma}_j = \check{\sigma}_i\check{\sigma}_j$ for all $i, j = 1,\ 2,\ \ldots,\ n-1$ with $|i-j| \geq 2$.

- If $k = i$:

$\breve{\sigma}_j\left(\breve{\sigma}_i\left(x_i\right)\right) = \breve{\sigma}_j\left(x_i x_{i+1} x_i^{-1}\right) = x_i x_{i+1} x_i^{-1}$

$\breve{\sigma}_i\left(\breve{\sigma}_j\left(x_i\right)\right) = \breve{\sigma}_i\left(x_i\right) = x_i x_{i+1} x_i^{-1}$

- If $k = i+1$:

$\breve{\sigma}_j\left(\breve{\sigma}_i\left(x_{i+1}\right)\right) = \breve{\sigma}_j\left(x_i\right) = x_i$

$\breve{\sigma}_i\left(\breve{\sigma}_j\left(x_{i+1}\right)\right) = \breve{\sigma}_i\left(x_{i+1}\right) = x_i$

- If $k \neq i,\ i+1$:

$\breve{\sigma}_j\left(\breve{\sigma}_i\left(x_k\right)\right) = \breve{\sigma}_j\left(x_k\right) = x_k$

$\breve{\sigma}_i\left(\breve{\sigma}_j\left(x_k\right)\right) = \breve{\sigma}_i\left(x_k\right) = x_k$

The cases when $k = j$, $k = j+1$ and $k \neq j,\ j+1$ are similar to the previous.

Secondly, we prove $\breve{\sigma}_i \breve{\sigma}_{i+1} \breve{\sigma}_i = \breve{\sigma}_{i+1} \breve{\sigma}_i \breve{\sigma}_{i+1}$ for all $i = 1,\ 2,\ \ldots,\ n-2$.

- If $k = i$:

$\breve{\sigma}_i\left(\breve{\sigma}_{i+1}\left(\breve{\sigma}_i\left(x_i\right)\right)\right) = \breve{\sigma}_i\left(\breve{\sigma}_{i+1}\left(x_i x_{i+1} x_i^{-1}\right)\right) = \breve{\sigma}_i\left(x_i x_{i+1} x_{i+2} x_{i+1}^{-1} x_i^{-1}\right) = x_i x_{i+1} x_{i+2} x_{i+1}^{-1} x_i^{-1}$

$\breve{\sigma}_{i+1}\left(\breve{\sigma}_i\left(\breve{\sigma}_{i+1}\left(x_i\right)\right)\right) = \breve{\sigma}_{i+1}\left(\breve{\sigma}_i\left(x_i\right)\right) = \breve{\sigma}_{i+1}\left(x_i x_{i+1} x_i^{-1}\right) = x_i x_{i+1} x_{i+2} x_{i+1}^{-1} x_i^{-1}$

- If $k = i+1$:

$\breve{\sigma}_i\left(\breve{\sigma}_{i+1}\left(\breve{\sigma}_i\left(x_{i+1}\right)\right)\right) = \breve{\sigma}_i\left(\breve{\sigma}_{i+1}\left(x_i\right)\right) = \breve{\sigma}_i\left(x_i\right) = x_i x_{i+1} x_i^{-1}$

$\breve{\sigma}_{i+1}\left(\breve{\sigma}_i\left(\breve{\sigma}_{i+1}\left(x_{i+1}\right)\right)\right) = \breve{\sigma}_{i+1}\left(\breve{\sigma}_i\left(x_{i+1} x_{i+2} x_{i+1}^{-1}\right)\right) = \breve{\sigma}_{i+1}\left(x_i x_{i+2} x_i^{-1}\right) = x_i x_{i+1} x_i^{-1}$

- If $k \neq i,\ i+1$:

$\breve{\sigma}_i\left(\breve{\sigma}_{i+1}\left(\breve{\sigma}_i\left(x_k\right)\right)\right) = \breve{\sigma}_i\left(\breve{\sigma}_{i+1}\left(x_k\right)\right) = \breve{\sigma}_i\left(x_k\right) = x_k$

$\breve{\sigma}_{i+1}\left(\breve{\sigma}_i\left(\breve{\sigma}_{i+1}\left(x_k\right)\right)\right) = \breve{\sigma}_{i+1}\left(\breve{\sigma}_i\left(x_k\right)\right) = \breve{\sigma}_{i+1}\left(x_k\right) = x_k$ □

Remark 2.26. Since product of braids is usually written from left to right and composition of automorphisms from right to left, we have $\check{\eta\tau} = \check{\tau} \circ \check{\eta}$ if η and τ are braids.

Chapter 3

The Burau representation

This chapter relates the two previous ones by defining a representation of the braid group, B_n, called the Burau representation. We will also see an interpretation of this representation from [5] and we will discuss its main properties.

3.1 Definition

For all $n \geq 1$, Werner Burau introduced a linear representation of the braid group, B_n, by $n \times n$ matrices over the ring of Laurent polynomials over a field k, these are linear combinations of positive and negative powers of the variable with coefficients in k. We denoted it by

$$\Lambda = k\left[t, t^{-1}\right].$$

Fixed $n \geq 2$ and for $i = 1, \ldots, n-1$, we consider the following $n \times n$ matrix over the ring Λ :

$$U_i = \begin{pmatrix} I_{i-1} & 0 & 0 & 0 \\ 0 & 1-t & t & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & I_{n-i-1} \end{pmatrix}$$

where I_k denotes the unit $k \times k$ matrix. Note that when $i = 1$, there is no identity matrix in the upper left corner of U_i and when $i = n-1$, there is no unit matrix in the lower right corner of U_i.

Thus each matrix U_i has a diagonal block form with blocks being the unit matrices and the 2×2 matrix

$$U = \begin{pmatrix} 1-t & t \\ 1 & 0 \end{pmatrix}.$$

By the Cayley-Hamilton theorem, any 2×2 matrix M over the ring Λ satisfies $M^2 - tr\,(M)\,M + det\,(M)\,I_2 = 0$. So, for $M = U$, this gives $U^2 - (1-t)\,U - tI_2 = 0$. Since identity matrices also satisfy this equation, we have

$$U_i^2 - (1-t)\,U_i - tI_n = 0$$

for all i. This can be rewritten as $U_i\,(U_i - (\,1-t)\,I_n) = tI_n$. Hence, U_i is invertible over Λ and its inverse is:

$$U_i^{-1} = t^{-1}\,(U_i - (\,1-t)\,I_n) = \begin{pmatrix} I_{i-1} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & t^{-1} & 1-t^{-1} & 0 \\ 0 & 0 & 0 & I_{n-i-1} \end{pmatrix}.$$

The block form of the matrices $U_1, \ldots, U_{n-1}$ implies that

$$U_iU_j = U_jU_i$$

for all i, j with $|i-j| \geq 2$. And it is easy to check that

$$U_iU_{i+1}U_i = U_{i+1}U_iU_{i+1}$$

for all $i = 1, \ldots, n-2$.

By Remark 2.23, $\psi\,(\sigma_i) = U_i$ with $i = 1, \ldots, n-1$ defines a group homomorphism ψ_n from the braid group B_n with $n \geq 2$ to the group $GL\,(n, \Lambda)$ of invertible $n \times n$ matrices over Λ. That is to say, we have the following representation:

$$\begin{aligned} \psi_n : B_n &\longrightarrow GL\,(n,\Lambda) \\ \sigma_i &\longmapsto U_i \end{aligned}$$

This representation is called the *Burau representation of B_n*.

3.1.1 The reduced Burau representation

We will show that the Burau representation is reducible.

Proposition 3.1. *Let $n \geq 3$ and $V_1,\ V_2,\ \ldots,\ V_{n-1}$ be the $(n-1) \times (n-1)$ matrices over Λ given by*

$$V_1 = \begin{pmatrix} -t & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & I_{n-3} \end{pmatrix},\ V_{n-1} = \begin{pmatrix} I_{n-3} & 0 & 0 \\ 0 & 1 & t \\ 0 & 0 & -t \end{pmatrix},$$

and for $1 < i < n-1$

$$V_i = \begin{pmatrix} I_{i-2} & 0 & 0 & 0 & 0 \\ 0 & 1 & t & 0 & 0 \\ 0 & 0 & -t & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & I_{n-i-2} \end{pmatrix}.$$

Then, for all $i = 1, \ldots, n-1$,

$$C^{-1}U_iC = \begin{pmatrix} V_i & 0 \\ *_i & 1 \end{pmatrix},$$

where C is the $n \times n$ matrix

$$C = C_n = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

and $*_i$ is the row of lenght $n-1$ equal to 0 if $i < n-1$ and to $(0,\ \ldots,\ 0,\ 1)$ if $i = n-1$.

Proof.

For $i = 1,\ \ldots,\ n-1$ set

$$V_i' = \begin{pmatrix} V_i & 0 \\ *_i & 1 \end{pmatrix}.$$

It is sufficient to prove that

$$U_iC = CV_i' \text{ for all i.}$$

Given i, it is easy to check that U_iC is obtained from C by:

a) Replacing the entry $(i,\ i)$ by $1-t$.

b) Replacing the entry $(i+1,\ i)$ by 1.

Similarly, it is also easy to check that $CV_i^{'}$is obtained from C by the same modifications as above. □

Corollary 3.2. *The Burau representation is reducible for all $n \geq 2$.*

Proof.

Firstly, consider $n \geq 3$ and let $e_1, \ldots, e_n$ be the canonical base of Λ^n. Taking the Λ-module $\langle \; Ce_n \; \rangle$, we are going to see that it is invariant by the Burau representation: $\psi_n(\sigma_i) = U_i$. Using the previous theorem, we obtain that:

$$U_i C e_n = C \begin{pmatrix} V_i & 0 \\ *_i & 1 \end{pmatrix} e_n = Ce_n.$$

Secondly, consider $n = 2$ and observe that:

$$C^{-1}U_1C = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1-t & t \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -t & 0 \\ 1 & 1 \end{pmatrix}.$$

Now, let e_1, e_2 be the canonical base of Λ^2 and take the Λ-module $\langle \; Ce_2 \; \rangle$. We are going to see that it is invariant by the Burau representation. Using the previous equality, we obtain that:

$$U_1Ce_2 = C \begin{pmatrix} -t & 0 \\ 1 & 1 \end{pmatrix} e_2 = Ce_2.$$

In these two cases, the definition of irreducible representation has been contradicted because $\langle \; Ce_n \; \rangle$ is ψ_n-stable. So, for any $n \geq 2$ the Burau representation is reducible. □

Definition 3.3. Since the matrices $U_1, \ldots, U_{n-1}$ satisfy the braid relations, the conjugate matrices $C^{-1}U_1C, \ldots, C^{-1}U_nC$ also do. The relation

$$C^{-1}U_iC = \begin{pmatrix} V_i & 0 \\ *_i & 1 \end{pmatrix}$$

implies that matrices V_i also satisfy them. It is easy to see that these matrices are invertible over Λ and therefore belong to $GL_{n-1}(\Lambda)$, so we have a group homomorphism

$$\begin{array}{rcl} \psi_n^r : B_n & \longrightarrow & GL_{n-1}(\Lambda) \\ \sigma_i & \longmapsto & V_i \end{array}$$

$$\begin{array}{rcl} \psi_2^r : B_2 & \longrightarrow & GL_1(\Lambda) \\ \sigma_1 & \longmapsto & -t \end{array}$$

which is called the *reduced Burau representation*.

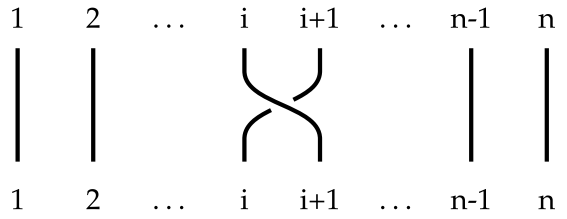
3.2 A probabilistic interpretation

It was V. Jones who, in his paper [5], notes that there is a "probabilistic" interpretation of the Burau representation of positive braids, which are n-braids given by words on $\sigma_1, \sigma_2, \ldots, \sigma_{n-1}$ (not on $\sigma_1^{-1}, \sigma_2^{-1}, \ldots, \sigma_{n-1}^{-1}$). He affirmed that:

"For positive braids there is also a mechanical interpretation of the Burau matrix: Lay the braid out flat and make it into a bowling alley with n lanes, the lanes going over each other according to the braid. If a ball traveling along a lane has probability $1-t$ of falling off the top lane, and continuing in the lane below, at every crossing then the (i,j) entry of the non-reduced Burau matrix is the probability that a ball bowled in the i^{th} lane will end up in the j^{th}."

In this section we are going to give a proof of Jones's interpretation and a specific example.

First of all, it is easy to see that the previous quote is true for the generators $\sigma_1, \sigma_2, \ldots, \sigma_{n-1}$ which generate the positive n-braid group, B_n^+. If we take any σ_i with the following diagram:



and apply the Burau representation, we obtain:

$$\psi_n(\sigma_i) = U_i = \begin{pmatrix} I_{i-1} & 0 & 0 & 0 \\ 0 & 1-t & t & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & I_{n-i-1} \end{pmatrix}.$$

Considering that the bowling alley goes from the top to the bottom of the braid diagram and that $1-t$ is the probability of a ball falling off the top lane and continuing in the lane below, the (i,j) entries of U_i are equal to the probabilities mentioned before. For instance, for the generator σ_i, the probability that the ball bowled in the i^{th} finish in the same one is that it falls off, that is to say $1-t$, which coincides with (i,i) entry.

We know that any braid of B_n^+ can be written as a product of braid generators:

$$\xi = \sigma_{i_1}\sigma_{i_2}\dots\sigma_{i_k} \in B_n^+ \text{ where } k \in \mathbb{N}$$

We will prove Jones's affirmation by induction on k. The case when $k = 1$ has just been seen, we check next the case $k = 2$; $\xi = \sigma_{i_1}\sigma_{i_2} \in B_n^+$. This case is not indispensable, but we have added it to have a more concrete idea of what we will do in the general case. Applying the Burau representation:

$$\psi_n\left(\sigma_{i_1}\sigma_{i_2}\right) = \psi_n\left(\sigma_{i_1}\right)\psi_n\left(\sigma_{i_2}\right) = U_{i_1}U_{i_2} = U_2'.$$

Denote $U_{i_1} = \left(a_{ij}\right)_{i,j}$, $U_{i_2} = \left(b_{kl}\right)_{k,l}$ and $U_2' = \left(c_{rs}\right)_{r,s}$.

By definition of the matrix product

$$c_{rs} = \textstyle\sum_{j=1}^{n} a_{rj}b_{js}$$

where a_{rj} is the probability that a ball bowled in r finish in j and b_{js} is the probability that a ball bowled in j finish in s, so c_{rs} is the probability that a ball bowled in r finish in s, not minding where it ends between the first braid generator, σ_{i_1}, and the second one, σ_{i_2}. So, this case has also been verified.

Now, we suppose that Jones' affirmation is true for any braid in B_n^+ with k generators, that is to say that the inductive hypothesis is:

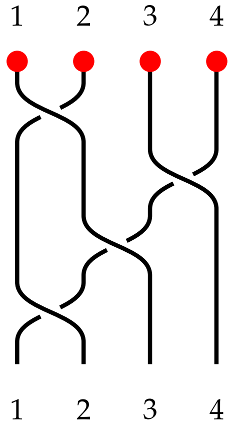
$$\psi_n\left(\sigma_{i_1}\sigma_{i_2}\dots\sigma_{i_k}\right) = U_k'$$

where U_k' is the Burau matrix whose entries coincide with the probabilities mentioned before. And, we want to see that it is true for any braid with $k+1$ generators. So, we take $\xi = \sigma_{i_1}\sigma_{i_2}\dots\sigma_{i_m}\sigma_{i_{k+1}} \in B_n^+$ and then,

$$\psi_n\left(\sigma_{i_1}\sigma_{i_2}\dots\sigma_{i_k}\sigma_{i_{k+1}}\right) = \psi_n\left(\sigma_{i_1}\sigma_{i_2}\dots\sigma_{i_k}\right)\psi_n\left(\sigma_{i_{k+1}}\right) = U_k'U_{i_{k+1}} = U_{k+1}'$$

where the first equality holds because ψ_n is a morphism, the second one is because of the inductive hypothesis and the case when $k = 1$ and the last one is obtained as in the case $k = 2$ above.

Example 3.4. We will consider the following positive braid ξ in B_4^+:



Considering that the bowling alley goes from the top to the bottom of the braid diagram, we note that

$$\xi = \sigma_1\sigma_3\sigma_2\sigma_1$$

and applying the Burau representation

$$\psi_4\left(\sigma_1\sigma_3\sigma_2\sigma_1\right) = \psi_4\left(\sigma_1\right)\psi_4\left(\sigma_3\right)\psi_4\left(\sigma_2\right)\psi_4\left(\sigma_1\right) = U_1U_3U_2U_1.$$

Finally,

$$U_1U_3U_2U_1 = \begin{pmatrix} 1-t & (1-t)\,t & t^2 & 0 \\ 1-t & t & 0 & 0 \\ 1-t & 0 & 0 & t \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

and it is easy to check that these entries are the probabilities mentioned by Jones.

3.3 Faithfulness of the Burau representation

We are going to see, depending on n, whether the Burau representation is faithful or not. We will start with the easy cases $n = 1$ and $n = 2$, then we analyze in detail when $n = 3$ and we state the case where $n \geq 5$. Case $n = 4$ is not solved yet; it is unknown whether ψ_4 is faithful or not.

3.3.1 ψ_1 is faithful

This is trivial because $B_1 = \{1\}$.

3.3.2 ψ_2 is faithful

It is easy to see that $B_2 \cong \mathbb{Z}$. If we consider a 2-braid, it will be or the identity braid or it will formed by a product of σ_1 or σ_1^{-1}, that is, it will have a determined number of positive or negative crossings (if we have a product of σ_1 and σ_1^{-1}, they will be canceled until we have a braid as the ones mentioned before). Then, the isomorphism consists in sending the braid to the number of crossings with its symbol (positive or negative). It is immediate to check that the product of braids corresponds to the sum in $\mathbb{Z}$.

Then, the generator $\sigma_1 \in B_2$ has image $U_1 \in GL(\Lambda)$ and observe that:

$$(1,-1)\,U_1 = (1,-1)\begin{pmatrix} 1-t & t \\ 1 & 0 \end{pmatrix} = (t,-t) = -t\,(1,-1)$$

which implies, $(1,-1)\,U_1^k = (-t)^k\,(1,-1)$ for all $k \in \mathbb{Z}$ and $U_1^k \neq I_2$ for all $k \in \mathbb{Z} - \{0\}$.

3.3.3 ψ_3 is faithful

To analyze ψ_3 we will use the reduced Burau representation, but there are some previous results that will be needed before.

Definition 3.5. Let

$$SL_2(\mathbb{Z}) = \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mid a,b,c,d \in \mathbb{Z} \text{ and } ad - bc = 1\}.$$

Then $SL_2(\mathbb{Z})$ is a group, its center is $\{+I_2,\ -I_2\}$ and its quotient group $SL_2(\mathbb{Z})/\{\pm I_2\}$ is called the *modular group*. It is denoted by $PSL_2(\mathbb{Z})$.

We have

$$PSL_2(\mathbb{Z}) = \{\hat{\mathbb{C}} \longrightarrow \hat{\mathbb{C}} \text{ where } \hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\},\, z \longmapsto \tfrac{az+b}{cz+d} \text{ and } ad - bc = 1\}.$$

We consider the following group presentations:

1. $\left\langle a,\ b \mid aba = bab,\ (aba)^4 = 1 \right\rangle$
2. $\langle s,\ t \mid s^3 = t^2,\ t^4 = 1 \rangle$
3. $\langle s,\ t \mid s^3 = t^2 = 1 \rangle$

then, we have the next result.

Lemma 3.6. **a)** *Presentations* 1 *and* 2 *define the same group* G.

b) *The group* H *defined by presentation* 3 *is isomorphic to* $G/\left\langle t^2\right\rangle$.

Proof.

a) Considering the substitutions $s = ab$, $t = aba$ and $a = s^{-1}t$, $b = t^{-1}s^2$, then the result is easy to verify.

b) It is easy to verify if we take the presentation 2 of G and compare with 3.

□

Consider the matrices $A,\ B \in SL_2\left(\mathbb{Z}\right)$ given by

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},\ B = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$$

which verify that

$$ABA = BAB \text{ and } (ABA)^4 = 1.$$

Therefore, we have a group morphism:

$$\begin{aligned} f : G &\longrightarrow SL_2\left(\mathbb{Z}\right) \\ a &\longmapsto A \\ b &\longmapsto B \end{aligned}$$

which in terms of s and t is

$$f\left(s\right) = AB = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},\ f\left(t\right) = ABA = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

and, $f\left(t^2\right) = -I_2$. Then, we have an induced morphism:

$$\bar{f} : H = G/\left\langle t^2\right\rangle \longrightarrow PSL_2\left(\mathbb{Z}\right).$$

Theorem 3.7. *The morphisms* $f : G \longrightarrow SL_2\left(\mathbb{Z}\right)$ *and* $\bar{f} : H \longrightarrow PSL_2\left(\mathbb{Z}\right)$ *are isomorphisms.*

Proof.

It follows from the previous definitions, that f induces an isomorphism:

$$\left\langle t^2\right\rangle \overset{\sim}{\to} \left\{\pm I_2\right\}$$

which lets us affirm:

a) f is injective $\Leftrightarrow$ $\bar{f}$ is injective.

b) f is exhaustive $\Leftrightarrow$ $\bar{f}$ is exhaustive.

Now, it will be sufficient to verify that f is exhaustive and that $\bar{f}$ is injective to prove the theorem.

a) To see that f is exhaustive we prove that the matrices A and B generate $SL_2(\mathbb{Z})$.

Let $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$, which means that $ad - bc = 1$ and denote $b = b(M)$ and $d = d(M)$.

We define $T = f(t) = f(aba) = f(a) f(b) f(a) = ABA = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$.

It has to be proved that M can be written as a word on matrices $A^{\pm 1}$ and $B^{\pm 1}$. We distinguish the following cases:

- If $b = 0 \Rightarrow a = d = \pm 1 \Rightarrow M = \left(\begin{smallmatrix} 1 & 0 \\ c & 1 \end{smallmatrix}\right)$ or $M = \left(\begin{smallmatrix} -1 & 0 \\ c & -1 \end{smallmatrix}\right)$. Then,

$$M = B^{-c} \text{ or } M = -I_2 B^c = T^2 B^c = (ABA)^2 B^c.$$

- If $d = 0 \Rightarrow bc = -1 \Rightarrow b = -c = 1$ or $b = -c = -1 \Rightarrow M = \left(\begin{smallmatrix} a & 1 \\ -1 & 0 \end{smallmatrix}\right)$ or $M = \left(\begin{smallmatrix} a & -1 \\ 1 & 0 \end{smallmatrix}\right)$. Then,

$$M = A^{-a} T = A^{-a} ABA \text{ or } M = A^a T^3 = A^a (ABA)^3.$$

- If $b = b(M) \neq 0$ and $d = d(M) \neq 0$, then $AM = \left(\begin{smallmatrix} a+c & b+d \\ c & d \end{smallmatrix}\right)$ and $TM = \left(\begin{smallmatrix} c & d \\ -a & -b \end{smallmatrix}\right)$. So, we have

$$\left.\begin{array}{rcl} b(AM) & = & b(M) + d(M) \\ d(AM) & = & d(M) \end{array}\right\} (1)$$

$$\left.\begin{array}{rcl} b(TM) & = & d(M) \\ d(TM) & = & -b(M) \end{array}\right\} (2)$$

From (1), we observe that exists $n \in \mathbb{Z}$ such that

$$0 \leq |b(A^n M)| \leq |d(A^n M)|.$$

This is because from (1), we have that $d(A^n M) = d(M)$ and if we put $b_n = b(A^n M)$ and $d_n = d(M)$, then we have $b_n = b_{n-1} + d$ (that is, $\{b_n\}$ is an arithmetic progression of difference d). Then, it is easy to see that there exists $n \in \mathbb{Z}$ such that

$$0 \leq |b_n| \leq |d|.$$

Now, using (2), we note that $\pm b$ and $\pm d$ can be exchanged by multiplying the matrix M by T on the left.

So, multiplying M by A or T on the left, as many times as necessary, we arrive to the cases when $b = 0$ or $d = 0$.

b) By a slight abuse of notation, we denote also by $s,\ t$ the classes of $s,\ t$ in H. To see that $\bar{f}$ is injective we will use the presentation $\langle s,\ t \mid s^3 = t^2 = e\rangle$ of H, where e denotes the neutral element of the group.

Any element h of H is a word in s and t, which means that:

$$h = s^{n_1} t^{m_1} s^{n_2} t^{m_2} \dots \text{ where } n_i,\ m_i \in \mathbb{Z}.$$

Then, due to the equality $t^2 = e$, we can assume that t always appears with exponent 1. Since, $s^3 = e$, we can assume that $n_i \in \{0,\ 1,\ -1\}$. So, we can see that any element of H can be written like:

- $\omega = s^{\epsilon_1} t s^{\epsilon_2} t \dots s^{\epsilon_r}$,
- ωt,
- $t\omega$,
- $t\omega t$,
- t and
- e

where $\epsilon_i = \pm 1$.

It will be sufficient to prove that, except the neutral element, none of these elements are in $Ker\left(\bar{f}\right)$. From the definition of $\bar{f}$, we know that $t \notin Ker\left(\bar{f}\right)$.

Note that $t\omega t = t\omega t^{-1}$ is conjugated of ω and $t\omega = t\omega t t^{-1}$ is conjugated of ωt. So, it will be sufficient to verify that $\bar{f}(\omega) \neq 1$ and $\bar{f}(\omega t) \neq 1$.

- Firstly, we want to see $\bar{f}(\omega t) \neq 1$. By definition of ω:

$$\omega t = (s^{\epsilon_1} t) \dots (s^{\epsilon_r} t)$$

From the proof of Lemma 3.6., we have that $s^{-1}t = a$ and $st = \left(t^{-1}s^2\right)^{-1} = b^{-1}$, so we have

$$\bar{f}\left(s^{-1}t\right) = \bar{A} \text{ and } \bar{f}(st) = \bar{B}^{-1}$$

where $\bar{A}$ and $\bar{B}$ are the images of $A, B \in SL_2(\mathbb{Z})$ in $PSL_2(\mathbb{Z})$. Then, $\bar{f}(\omega t)$ is equal to a product of matrices of $\bar{A}$ and $\bar{B}^{-1}$.

Note that no product of A and B^{-1} is equal to $\pm I_2$ because $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $B^{-1} = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$ so, this product will have always positives entries and the sum of the non diagonal entries will increase.

- Secondly, if $\bar{f}(\omega) = 1$, then $\bar{f}(\omega t) = \bar{f}(\omega)\bar{f}(t) = \bar{f}(t) = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. This is impossible as we have just said. □

Theorem 3.8. *The Burau representation ψ_3 is faithful.*

Proof.

We start defining the following group morphism:

$$\begin{array}{rcl} \xi_3 : B_3 & \longrightarrow & SL_2(\mathbb{Z}) \\ \sigma_1 & \longmapsto & a_1 \\ \sigma_2 & \longmapsto & a_2 \end{array}$$

where $a_1 = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$ and $a_2 = \left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right)$.

Note that the matrices a_1 and a_2 are obtained by taking the matrices, V_1 and V_2, defining the reduced Burau representation and putting t equal to -1.

Note that $A = a_1^T$ and $B = a_2^T$, thus $SL_2(\mathbb{Z}) = \left\langle a_1,\ a_2 \mid a_1a_2a_1 = a_2a_1a_2,\ (a_1a_2a_1)^4 = 1 \right\rangle$.

Since $B_3 = \langle \sigma_1,\ \sigma_2 \mid \sigma_1\sigma_2\sigma_1 = \sigma_2\sigma_1\sigma_2 \rangle$, we deduce that ξ_3 is well defined.

Now, we would like to see that $Ker(\xi_3) = \left\langle (\sigma_1\sigma_2\sigma_1)^4 \right\rangle$.

To this end, we observe:

i) It is clear that $(\sigma_1\sigma_2\sigma_1)^4 \in Ker(\xi_3)$.

ii) Due to the fact that the kernel of any group morphism is always a normal subgroup:

$$SNS\left((\sigma_1\sigma_2\sigma_1)^4\right) \subset Ker(\xi_3),$$

where $SNS(\Sigma)$ denotes the smallest normal subgroup which contains Σ. We know that if G is a group and $\Sigma \subset G$, the smallest normal subgroup that contains Σ is:

$$SNS\,(\Sigma) = \{gsg^{-1} |\, g \in G, s \in \Sigma\}.$$

iii) If $p \in Ker\,(\xi_3)$, what is to say $p\,(\sigma_1,\, \sigma_2) \in B_3$ such that $\xi_3\,(p) = 1$. Then $p\,(a_1,\, a_2)$ is the identity element in $SL_2\,(\mathbb{Z})$, thus

$p\,(a_1,\, a_2) \in SNS\left(a_2^{-1}a_1^{-1}a_2^{-1}a_1a_2a_1,\, (a_1a_2a_1)^4\right)$ of the free group generated by $a_1,\, a_2$.

Since in B_3 we already have the relation $\sigma_2^{-1}\sigma_1^{-1}\sigma_2^{-1}\sigma_1\sigma_2\sigma_1 = 1$ we can assume

$$p\,(\sigma_1,\, \sigma_2) \in SNS\left((\sigma_1\sigma_2\sigma_1)^4\right)$$

It is easy to check that the following equalities hold:

$$(\sigma_1\sigma_2\sigma_1)^2\,\sigma_1 = \sigma_1\,(\sigma_1\sigma_2\sigma_1)^2,$$

$$(\sigma_1\sigma_2\sigma_1)^2\,\sigma_2 = \sigma_2\,(\sigma_1\sigma_2\sigma_1)^2.$$

Then, $(\sigma_1\sigma_2\sigma_1)^2$ commutes with all the elements in B_3 and this implies that the smallest normal subgroup that contains $(\sigma_1\sigma_2\sigma_1)^4$ is the cyclic group generated by this element

$$SNS\left((\sigma_1\sigma_2\sigma_1)^4\right) = \left\langle (\sigma_1\sigma_2\sigma_1)^4 \right\rangle.$$

Next, we would like to see that

$$ker\,(\psi_3) \subset ker\,(\psi_3^r) \subset ker\,(\xi_3) = \left\langle (\sigma_1\sigma_2\sigma_1)^4 \right\rangle$$

but, in fact, since $C^{-1}\psi_3 C = \left(\begin{smallmatrix} \psi_3^r & 0 \\ 1 & 1 \end{smallmatrix}\right)$ we have that $ker\,(\psi_3) = ker\,(\psi_3^r)$ and the second inclusion is due to the fact that $a_1,\, a_2$ are obtained from the matrices $V_1,\, V_2$ (defined in Proposition 3.1) putting $t = -1$.

Finally, to prove that ψ_3 is faithful observe

$$V_1V_2V_1 = \begin{pmatrix} 0 & t^2 \\ -t & 0 \end{pmatrix},\ (V_1V_2V_1)^2 = \begin{pmatrix} t^3 & 0 \\ 0 & t^3 \end{pmatrix}$$

where the matrices V_1 and V_2. Thus, if $k \in \mathbb{Z}$, then

$$\psi_3^r\left((\sigma_1\sigma_2\sigma_1)^{4k}\right) = \begin{pmatrix} t^{6k} & 0 \\ 0 & t^{6k} \end{pmatrix} \neq I_2.$$

Consequently, $ker\,(\psi_3^r) = \{e\}$ and $ker\,(\psi_3) = \{e\}$, which means that ψ_3 is faithful. $\square$

The Burau representation is not faithful for $n \geq 5$, this is proved by finding an explicit non-trivial element in its kernel. [1]

Precisely, let $\gamma = \sigma_4\sigma_3^{-1}\sigma_2^{-1}\sigma_1^2\sigma_2^{-1}\sigma_1^{-2}\sigma_2^{-2}\sigma_1^{-1}\sigma_4^{-5}\sigma_2\sigma_3\sigma_4^3\sigma_2\sigma_1^2\sigma_2\sigma_3^{-1}$ be a braid in B_5 and ρ be the commutator [2] of $\gamma\sigma_4\gamma^{-1}$ and $\sigma_4\sigma_3\sigma_2\sigma_1^2\sigma_2\sigma_3\sigma_4$. Then, it can be proved that:

i) $\rho = \left[\gamma\sigma_4\gamma^{-1},\ \sigma_4\sigma_3\sigma_2\sigma_1^2\sigma_2\sigma_3\sigma_4\right] \neq I_5$

ii) $\rho \in ker\,(\psi_5)$

So, we have a braid $\rho \in B_5$, belonging to $ker\,(\psi_5)$ that is different from the group's identity (what can be checked with the help of a computer) and so ψ_5 is not faithful.

Consequently and due to the fact that $B_m \hookrightarrow B_n$ when $n \geq m$, ψ_n is not faithful for any $n \geq 5$.

[1] We have consulted [6, Chapter 3, Theorem 3.3].

[2] $[a, b] = aba^{-1}b^{-1}$.

Chapter 4

Free differential calculus

In this chapter we start defining some concepts related to the free differential calculus, also called Fox calculus, and then we will see how the Burau representation can be described in terms of this calculus.

4.1 The Fox partial derivatives

Let G be a group and $\mathbb{Z}[G]$ its group ring, that is

$$\mathbb{Z}[G] = \{\textstyle\sum_{i=1}^{k} n_i g_i |\ n_i \in \mathbb{Z},\ g_i \in G,\ k \in \mathbb{Z}\}.$$

Definition 4.1. The following homomorphism

$$\begin{array}{rcl} \bullet^{\varepsilon} : \mathbb{Z}[G] & \longrightarrow & \mathbb{Z} \\ \tau = \sum n_i g_i & \longmapsto & \tau^{\varepsilon} = \sum n_i \end{array}$$

is called the *augmentation homomorphism*.

Definition 4.2. A mapping $\Delta : \mathbb{Z}[G] \to \mathbb{Z}[G]$ is called a *derivation* of $\mathbb{Z}[G]$ if

$$\Delta(\xi + \eta) = \Delta(\xi) + \Delta(\eta) \tag{4.1}$$

$$\Delta(\xi\eta) = \Delta(\xi)\,\eta^{\varepsilon} + \xi\Delta(\eta) \tag{4.2}$$

where $\xi,\ \eta \in \mathbb{Z}[G]$. Properly, (4.1) is called additivity and (4.2) the product rule.

Denote $Der(\mathbb{Z}[G])$ the set of derivations of $\mathbb{Z}[G]$.

Lemma 4.3. a) *$Der(\mathbb{Z}[G])$ is a right $\mathbb{Z}[G]$-module under the operations defined by*

$$(\Delta_1 + \Delta_2)(\tau) = \Delta_1(\tau) + \Delta_2(\tau), \tag{4.3}$$

$$(\Delta\gamma)(\tau) = \Delta(\tau)\,\gamma \tag{4.4}$$

where $\Delta_1,\ \Delta_2 \in Der(\mathbb{Z}[G])$ and $\tau,\ \gamma \in \mathbb{Z}[G]$.

b) *Let Δ be a derivation. Then:*

$$\Delta(m) = 0 \textit{ for } m \in \mathbb{Z},$$

$$\Delta\left(g^{-1}\right) = -g^{-1}\Delta(g).$$

Proof.

a) Remember that if A is unitary ring and M is a set with 2 operations $+ : M \times M \to M$ and $\cdot : A \times M \to M$, then M is a right A-module if:

i) $(M, +)$ is an abelian group.

ii) $(x + y) r = xr + yr$

iii) $x (r + s) = xr + xs$

iv) $x (rs) = (xr) s$

v) $x1_A = x$

for all $x,\ y \in M$ and $r,\ s \in A$. In our case, we have $A = \mathbb{Z}[G]$ and $M = Der(\mathbb{Z}[G])$ and i) is easy to check from the definition.

Now, we are going to prove the rest of the properties.

ii) Let Δ_1 and Δ_2 be derivations and $\gamma \in \mathbb{Z}[G]$. Then, if $\tau \in \mathbb{Z}[G]$

$$((\Delta_1 + \Delta_2)\gamma)(\tau) = (\Delta_1 + \Delta_2)(\tau)\gamma = (\Delta_1(\tau) + \Delta_2(\tau))\gamma = (\Delta_1\gamma)(\tau) + (\Delta_2\gamma)(\tau)$$

where in the first and the third equalities we have used (4.4) while in the second we have used (4.3).

iii) Let Δ be a derivation and $\gamma,\ \eta \in \mathbb{Z}[G]$. Then, if $\tau \in \mathbb{Z}[G]$

$$(\Delta(\gamma + \eta))(\tau) = (\Delta\gamma + \Delta\eta)(\tau) = (\Delta\gamma)(\tau) + (\Delta\eta)(\tau)$$

where in all the equalities we have used (4.3).

iv) Let Δ be a derivation and $\gamma,\ \eta \in \mathbb{Z}[G]$. Then, if $\tau \in \mathbb{Z}[G]$

$$(\Delta(\gamma\eta))(\tau) = \Delta(\tau)\gamma\eta = (\Delta(\tau)\gamma)\eta = ((\Delta\gamma)(\tau))\eta = ((\Delta\gamma)\eta)(\tau)$$

where we have used (4.4).

v) Let Δ be a derivation and $1_{\mathbb{Z}[G]}$ be the neutral element of $\mathbb{Z}[G]$. Then, if $\tau \in \mathbb{Z}[G]$

$$\Delta 1_{\mathbb{Z}[G]}\left(\tau\right)=\Delta\left(\tau\right)1_{\mathbb{Z}[G]}=\Delta\left(\tau\right)$$

where in the first equality we have used (4.3) and then the result is obvious.

b) First of all, we notice that $\Delta\left(1\right)=0$. This is easy because:

$$\Delta\left(1\right)=\Delta\left(1\cdot 1\right)=\Delta\left(1\right)+\Delta\left(1\right)=2\Delta\left(1\right)$$

where in the second equality we have used the equation (4.2) and the result is easy to see. This implies that,

$$\Delta\left(m\right)=\Delta\left(1+\ldots+1\right)=\Delta\left(1\right)+\ldots+\Delta\left(1\right)=0$$

where in both additions there are m summands and in the second equality we have used the equation (4.1).

Secondly, we note that

$$0=\Delta\left(1\right)=\Delta\left(g^{-1}g\right)=\Delta\left(g^{-1}\right)+g^{-1}\cdot\Delta\left(g\right)$$

where in the third equality we have used the equation (4.2) and then the result is easy to see. □

Proposition 4.4. *Let $\mathfrak{F}=\langle\{S_i|\ i\in I\}\rangle$ be a free group. There exists a uniquely determined derivation $\Delta:\mathbb{Z}\left[\mathfrak{F}\right]\to\mathbb{Z}\left[\mathfrak{F}\right]$ with $\Delta S_i=\omega_i$ for arbitrary elements $\omega_i\in\mathbb{Z}\left[\mathfrak{F}\right]$ and $i\in I$.*

Proof.

Using additivity and the product rule, uniqueness is assured. We note that,

- $\Delta\left(S_{i_1}^{n_1}S_{i_2}^{n_2}\right)=\Delta S_{i_1}^{n_1}+S_{i_1}^{n_1}\Delta S_{i_2}^{n_2}$
- $\Delta\left(S_{i_1}^{n_1}S_{i_2}^{n_2}S_{i_3}^{n_3}\right)=\Delta\left(S_{i_1}^{n_1}S_{i_2}^{n_2}\right)+S_{i_1}^{n_1}S_{i_2}^{n_2}\Delta S_{i_3}^{n_3}=\Delta S_{i_1}^{n_1}+S_{i_1}^{n_1}\Delta S_{i_2}^{n_2}+S_{i_1}^{n_1}S_{i_2}^{n_2}\Delta S_{i_3}^{n_3}$

where we have used the product rule. Then, we define $\Delta\left(S_{i_1}^{n_1}\ldots S_{i_k}^{n_k}\right)$:

$$\Delta\left(S_{i_1}^{n_1}\ldots S_{i_k}^{n_k}\right)=\Delta S_{i_1}^{n_1}+S_{i_1}^{n_1}\Delta S_{i_2}^{n_2}+\ldots+S_{i_1}^{n_1}\ldots+S_{i_{k-1}}^{n_{k-1}}\Delta S_{i_k}^{n_k}. \qquad (4.5)$$

One can check that Δ is well defined and it is a derivation of $\mathbb{Z}\left[G\right]$.

For example,

$$\Delta\left(uS_i^{\eta}S_i^{-\eta}v\right)=\Delta u+u\Delta S_i^{\eta}+uS_i^{\eta}\Delta S_i^{-\eta}+uS_i^{\eta}S_i^{-\eta}\Delta v=$$
$$\Delta u+u\left(\Delta S_i^{\eta}\Delta S_i^{-\eta}\right)+u\Delta v=\Delta u+u\Delta v=\Delta\left(uv\right)$$

(where we have used: in the first equality the equation (4.5), in the second and the fourth one the product rule and in the third one that $\Delta(1)=0$) thus, Δ is well defined on $\mathfrak{F}$. □

Definition 4.5. The derivations

$$\begin{array}{rcl}\frac{\partial}{\partial S_i}:\mathbb{Z}\left[\mathfrak{F}\right] & \longrightarrow & \mathbb{Z}\left[\mathfrak{F}\right]\\ S_j & \longmapsto & \begin{cases}1 & for\quad i=j\\ 0 & for\quad i\neq j,\end{cases}\end{array}$$

of the group ring of a free group $\mathfrak{F}=\langle\{S_i|\ i\in J\}\rangle$ are called *partial derivations*.

Remark 4.6. It was Fox who, considering the free group $F_n=\langle x_1,\ \ldots,\ x_n\rangle$, showed the existence of derivations in $\mathbb{Z}\left[F_n\right]$ such that

$$\begin{array}{rcl}\frac{\partial}{\partial x_i}:\mathbb{Z}\left[F_n\right] & \longrightarrow & \mathbb{Z}\left[F_n\right]\\ x_j & \longmapsto & \delta_{i,j}=\begin{cases}1 & for\quad i=j\\ 0 & for\quad i\neq j.\end{cases}\end{array}$$

They are called *The Fox partial derivatives* and they are a particular case of Proposition 4.4 and Definition 4.5.

From now and on, we will consider the free group $F_n=\langle x_1,\ \ldots,\ x_n\rangle$ and the Fox partial derivatives.

4.1.1 The Chain Rule

Let $\varrho: F_n\to F_n$ be an automorphism, denoting its extension to $\mathbb{Z}\left[G\right]$ also by ϱ, we have the next lemma.

Lemma 4.7. *If $\omega\in\mathbb{Z}\left[F_n\right]$, then*

$$\tfrac{\partial\varrho(\omega)}{\partial x_j}=\textstyle\sum_{k=1}^{n}\varrho\left(\frac{\partial\omega}{\partial x_k}\right)\cdot\frac{\partial\varrho(x_k)}{\partial x_j}\ \textit{for all}\ j=1,\ \ldots,\ n.$$

Proof.

We consider the set $\mathcal{S}$ of applications

$$\alpha:\mathbb{Z}\left[F_n\right]\longrightarrow\mathbb{Z}\left[F_n\right]$$

such that

$$\alpha\left(\sum n_g g\right) = \sum n_g \alpha\left(g\right)$$

$$\alpha\left(\xi\eta\right) = \alpha\left(\xi\right)\eta^{\varepsilon} + \varrho\left(\xi\right)\alpha\left(\eta\right) \textbf{(*)}.$$

These maps are univocally determined by its value on the generators of F_n.

Any element of F_n can be written as

$$x_{i_1}^{\varepsilon_1}\dots x_{i_{k-1}}^{\varepsilon_{k-1}}x_{i_k}^{\varepsilon_k} \text{ with } \varepsilon_i \in \mathbb{Z}.$$

and we are going to see that $\alpha\left(x_{i_1}^{\varepsilon_1}\dots x_{i_{k-1}}^{\varepsilon_{k-1}}x_{i_k}^{\varepsilon_k}\right)$ is determined by the values of $\alpha\left(x_i\right)$.

Taking $\xi = x_{i_1}^{\varepsilon_1}\dots x_{i_{k-1}}^{\varepsilon_{k-1}}$ and $\eta = x_{i_k}^{\varepsilon_k}$, we note that $\alpha\left(x_{i_1}^{\varepsilon_1}\dots x_{i_{k-1}}^{\varepsilon_{k-1}}x_{i_k}^{\varepsilon_k}\right)$ is determined by $\alpha\left(\xi\right)$ and $\alpha\left(\eta\right)$ because of **(*)**. Then, due to the fact that the length of ξ and η is shorter than the length of $x_{i_1}^{\varepsilon_1}\dots x_{i_{k-1}}^{\varepsilon_{k-1}}x_{i_k}^{\varepsilon_k}$, we can reduce to the case where we only have $\alpha\left(x_i\right)$ and $\alpha\left(x_i^{-1}\right)$ by induction on its length.

But, as we have seen before, $\alpha\left(x_i^{-1}\right) = -\varrho\left(x_i^{-1}\right)\alpha\left(x_i\right)$ and we can affirm that $\alpha\left(x_i^{-1}\right)$ is determined by $\alpha\left(x_i\right)$.

Now, we prove that both sides of the equality in statement of the lemma are in $\mathcal{S}$, additivity is easy to see.

If

$$\alpha\left(\omega\right) = \frac{\partial\varrho(\omega)}{\partial x_j} \text{ ,}$$

we have

$$\alpha\left(\xi\eta\right) = \frac{\partial\varrho(\xi\eta)}{\partial x_j} = \frac{\partial(\varrho(\xi)\varrho(\eta))}{\partial x_j} = \frac{\partial\varrho(\xi)}{\partial x_j}\varrho\left(\eta\right)^{\varepsilon} + \varrho\left(\xi\right)\frac{\partial\varrho(\eta)}{\partial x_j}$$

But, note that $\varrho\left(\eta\right)^{\varepsilon} = \eta^{\varepsilon}$ and we have just obtained that the left hand side of the statement of the lemma is in $\mathcal{S}$.

If

$$\alpha\left(\omega\right) = \sum_{k=1}^{n}\varrho\left(\frac{\partial\omega}{\partial x_k}\right)\frac{\partial\varrho(x_k)}{\partial x_j} \text{ ,}$$

we have

$$\varrho\left(\frac{\partial(\xi\eta)}{\partial x_k}\right)=\varrho\left(\frac{\partial\xi}{\partial x_k}\eta^{\varepsilon}+\xi\frac{\partial\eta}{\partial x_k}\right)=\varrho\left(\frac{\partial\xi}{\partial x_k}\right)\varrho\left(\eta^{\varepsilon}\right)+\varrho\left(\xi\right)\varrho\left(\frac{\partial\eta}{\partial x_k}\right).$$

Then, also using $\varrho\left(\eta\right)^{\varepsilon}=\eta^{\varepsilon}$, we obtain the right hand side (of the statement of the lemma) is also in $\mathcal{S}$.

$$\alpha\left(\xi\eta\right)=\sum_{k=1}^{n}\varrho\left(\frac{\partial\xi}{\partial x_k}\right)\eta^{\varepsilon}\frac{\partial\varrho(x_k)}{\partial x_j}+\sum_{k=1}^{n}\varrho\left(\xi\right)\varrho\left(\frac{\partial\eta}{\partial x_k}\right)\frac{\partial\varrho(x_k)}{\partial x_j}=$$
$$\alpha\left(\xi\right)\eta^{\varepsilon}+\varrho\left(\xi\right)\alpha\left(\eta\right).$$

Finally, we can affirm that it is sufficient to proof the equality if $\omega=x_i$. In this case, we have

$$\sum_{k=1}^{n}\varrho\left(\frac{\partial x_i}{\partial x_k}\right)\frac{\partial\varrho(x_k)}{\partial x_j}=\sum_{k=1}^{n}\varrho\left(\delta_{i,k}\right)\frac{\partial\varrho(x_k)}{\partial x_j}=\frac{\partial\varrho(x_i)}{\partial x_j}.$$

□

4.2 The Fox calculus and the Burau representation

Let B_n be the braid group with n strings and let F_n be the free group with n generators $x_1,\ x_2,\ \ldots,\ x_n$.

It is necessary to remember Theorem 2.25 in chapter 2 where we have defined the following group isomorphism between braids and braid automorphisms:

$$\begin{array}{rcl} B_n & \longrightarrow & \breve{B}_n=Aut\left(F_n\right)\\ \sigma_i & \longmapsto & \breve{\sigma}_i. \end{array}$$

Denoting $\langle t\rangle=\{t^n|\ n\in\mathbb{Z}\}$ the infinite cyclic group generated by a variable t, we define the group morphism:

$$\begin{array}{rcl} F_n & \longrightarrow & \langle t\rangle\\ x_{i_1}^{a_1}\ldots x_{i_k}^{a_k} & \longmapsto & t^{a_{i_1}+\ldots+a_{i_1}} \end{array}$$

which induces a ring morphism:

$$a:\mathbb{Z}\left[F_n\right]\ \longrightarrow\ \mathbb{Z}\left[t,t^{-1}\right].$$

Now, we consider the application:

$$\begin{array}{rcl} \rho:B_n & \longrightarrow & GL\left(n,\mathbb{Z}\left[t,t^{-1}\right]\right)\\ \sigma & \longmapsto & \left(a\left(\dfrac{\partial\breve{\sigma}\left(x_i\right)}{\partial x_j}\right)\right)_{i,j} \end{array}$$

where i denotes the row, j the column and where

$$\frac{\partial}{\partial x_j} : \mathbb{Z}\left[F_n\right] \longrightarrow \mathbb{Z}\left[F_n\right]$$

is the Fox partial derivate.

Theorem 4.8. *ρ is the Burau representation of B_n.*

Proof.

We start with the simplest cases. If we choose $n = 2$, we know that in B_2 there is only one generator σ_1 and we want to see that:

$$\rho\left(\sigma_1\right) \longrightarrow U_1 = \begin{pmatrix} 1-t & t \\ 1 & 0 \end{pmatrix}.$$

Using Definition 2.24 in chapter 2, we note that:

$$\rho\left(\sigma_1\right) \longrightarrow \begin{pmatrix} a\left(\frac{\partial\breve{\sigma}_1(x_1)}{\partial x_1}\right) & a\left(\frac{\partial\breve{\sigma}_1(x_1)}{\partial x_2}\right) \\ a\left(\frac{\partial\breve{\sigma}_1(x_2)}{\partial x_1}\right) & a\left(\frac{\partial\breve{\sigma}_1(x_2)}{\partial x_2}\right) \end{pmatrix} = \begin{pmatrix} a\left(\frac{\partial x_1 x_2 x_1^{-1}}{\partial x_1}\right) & a\left(\frac{\partial x_1 x_2 x_1^{-1}}{\partial x_2}\right) \\ a\left(\frac{\partial x_1}{\partial x_1}\right) & a\left(\frac{\partial x_1}{\partial x_2}\right) \end{pmatrix}.$$

Then, applying the Fox partial derivates (and (4.5)) and the morphism a, we obtain what we were looking for:

$$\rho\left(\sigma_1\right) \longrightarrow \begin{pmatrix} a\left(1-x_2\right) & a\left(x_1\right) \\ a\left(1\right) & a\left(0\right) \end{pmatrix} = \begin{pmatrix} 1-t & t \\ 1 & 0 \end{pmatrix}.$$

If we choose $n = 2$, we know that in B_3 there are 2 generators, σ_1 and σ_2, and we want to see that:

$$\rho\left(\sigma_1\right) \longrightarrow U_1 = \begin{pmatrix} 1-t & t & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$\rho\left(\sigma_2\right) \longrightarrow U_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1-t & t \\ 0 & 1 & 0 \end{pmatrix}.$$

Using Definition 2.24 in chapter 2, we note that:

$$\rho\left(\sigma_1\right) \longrightarrow \begin{pmatrix} a\left(\frac{\partial x_1 x_2 x_1^{-1}}{\partial x_1}\right) & a\left(\frac{\partial x_1 x_2 x_1^{-1}}{\partial x_2}\right) & a\left(\frac{\partial x_1 x_2 x_1^{-1}}{\partial x_3}\right) \\ a\left(\frac{\partial x_1}{\partial x_1}\right) & a\left(\frac{\partial x_1}{\partial x_2}\right) & a\left(\frac{\partial x_1}{\partial x_3}\right) \\ a\left(\frac{\partial x_3}{\partial x_1}\right) & a\left(\frac{\partial x_3}{\partial x_2}\right) & a\left(\frac{\partial x_3}{\partial x_3}\right) \end{pmatrix}$$

and

$$\rho\left(\sigma_2\right) \longrightarrow \left(\begin{array}{ccc} a\left(\frac{\partial x_1}{\partial x_1}\right) & a\left(\frac{\partial x_1}{\partial x_2}\right) & a\left(\frac{\partial x_1}{\partial x_3}\right) \\ a\left(\frac{\partial x_2 x_3 x_2^{-1}}{\partial x_1}\right) & a\left(\frac{\partial x_2 x_3 x_2^{-1}}{\partial x_2}\right) & a\left(\frac{\partial x_2 x_3 x_2^{-1}}{\partial x_3}\right) \\ a\left(\frac{\partial x_2}{\partial x_1}\right) & a\left(\frac{\partial x_2}{\partial x_2}\right) & a\left(\frac{\partial x_2}{\partial x_3}\right) \end{array}\right).$$

Then, applying the Fox partial derivates (and (4.5)) and the morphism a, we obtain U_1 and U_2, respectively:

$$\rho\left(\sigma_1\right) \longrightarrow \left(\begin{array}{ccc} a\left(1-x_2\right) & a\left(x_1\right) & a\left(0\right) \\ a\left(1\right) & a\left(0\right) & a\left(0\right) \\ a\left(0\right) & a\left(0\right) & a\left(1\right) \end{array}\right) = \left(\begin{array}{ccc} 1-t & t & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{array}\right)$$

and

$$\rho\left(\sigma_2\right) \longrightarrow \left(\begin{array}{ccc} a\left(1\right) & a\left(0\right) & a\left(0\right) \\ a\left(0\right) & a\left(1-x_3\right) & a\left(x_2\right) \\ a\left(0\right) & a\left(1\right) & a\left(0\right) \end{array}\right) = \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1-t & t \\ 0 & 1 & 0 \end{array}\right).$$

Now, we want to verify the general case:

$$\rho\left(\sigma_i\right) \longrightarrow U_i = \left(\begin{array}{cccc} I_{i-1} & 0 & 0 & 0 \\ 0 & 1-t & t & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & I_{n-i-1} \end{array}\right) for\ i = 1,\ 2,\ \ldots,\ n-1$$

where I_k denotes the unit $k \times k$ matrix.

Using Definition 2.24 in chapter 2 and applying the Fox partial derivates (and (4.5)), we note that:

$$\left.\begin{array}{lcl} \check{\sigma}_i\left(x_1\right) & = & x_1 \\ \vdots & & \vdots \\ \check{\sigma}_i\left(x_{i-1}\right) & = & x_{i-1} \end{array}\right\} \longrightarrow I_{i-1}$$

$$\left.\begin{array}{lcl} \check{\sigma}_i\left(x_i\right) & = & x_i x_{i+1} x_i^{-1} \\ \check{\sigma}_i\left(x_{i+1}\right) & = & x_i \end{array}\right\} \longrightarrow \textbf{(*)}$$

$$\left.\begin{array}{lcl} \check{\sigma}_i\left(x_{i+2}\right) & = & x_{i+2} \\ \vdots & & \vdots \\ \check{\sigma}_i\left(x_n\right) & = & x_n \end{array}\right\} \longrightarrow I_{n-i-1}$$

(*) → the matrix where all the elements are zero except:

$\frac{\partial\check{\sigma}_i(x_i)}{\partial x_i} = \frac{\partial x_i x_{i+1} x_i^{-1}}{\partial x_i} = 1 - x_{i+1},$

$\frac{\partial\check{\sigma}_i(x_i)}{\partial x_{i+1}} = \frac{\partial x_i x_{i+1} x_i^{-1}}{\partial x_{i+1}} = x_i$ and

$\frac{\partial\check{\sigma}_i(x_{i+1})}{\partial x_i} = \frac{\partial x_i}{\partial x_i} = 1.$

Finally, applying the morphism a, we obtain U_i:

$$\rho\left(\sigma_i\right) \quad \longrightarrow \quad \begin{pmatrix} I_{i-1} & 0 & 0 & 0 \\ 0 & a\left(1 - x_{i+1}\right) & a\left(x_i\right) & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & I_{n-i-1} \end{pmatrix} = \begin{pmatrix} I_{i-1} & 0 & 0 & 0 \\ 0 & 1-t & t & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & I_{n-i-1} \end{pmatrix}.$$

Now, we verify that the following equality holds:

$$\rho\left(\eta\tau\right) = \rho\left(\eta\right)\rho\left(\tau\right). \tag{4.6}$$

We first note that

$$a\left(\check{\tau}\left(\omega\right)\right) = a\left(\omega\right)$$

for any $\omega \in \mathbb{Z}\left[F_n\right]$. This equality holds because, by induction on the length, we can reduce to the case where $\omega = x_k \in F_n$ and $\tau = \sigma_i$ and then, the result is trivial.

If we fix $x_i,\ x_j$ (generators of F_n) and we apply the chain rule proved in Lemma 4.7, we have:

$$\frac{\partial\check{\tau}(\check{\eta}(x_i))}{\partial x_j} = \textstyle\sum_{k=1}^{n} \check{\tau}\left(\frac{\partial\check{\eta}(x_i)}{\partial x_k}\right)\frac{\partial\check{\tau}(x_k)}{\partial x_j}\ .$$

Then, applying the ring morphism a, we obtain:

$$a\left(\frac{\partial\check{\tau}\left(\check{\eta}\left(x_i\right)\right)}{\partial x_j}\right) = \sum_{k=1}^{n} a\left(\check{\tau}\left(\frac{\partial\check{\eta}\left(x_i\right)}{\partial x_k}\right)\right) a\left(\frac{\partial\check{\tau}\left(x_k\right)}{\partial x_j}\right) = \sum_{k=1}^{n} a\left(\frac{\partial\check{\eta}\left(x_i\right)}{\partial x_k}\right) a\left(\frac{\partial\check{\tau}\left(x_k\right)}{\partial x_j}\right) \tag{4.7}$$

On the left most term of (4.7) we have the entries of the matrix $\rho\left(\eta\tau\right)$, on the right most term there is the product of rows of $\rho\left(\eta\right)$ by columns of $\rho\left(\tau\right)$, thus we have

$$\rho\left(\eta\tau\right) = \rho\left(\eta\right)\rho\left(\tau\right).$$

Finally, we already know that any braid of B_n can be written as a product of generators $\sigma_1, \ldots, \sigma_{n-1}$, the determinant of any $\rho(\sigma_i)$ is equal to $-t$ and taking into account (4.6), we can affirm that ρ is well-defined, namely that $\rho(\omega)$ is invertible for all $\omega \in \mathbb{Z}[F_n]$, and, again by (4.6), it is a group morphism.

Thus, ρ is the Burau representation of B_n. □

Conclusions

We have seen what is the braid group and how it can be defined in geometrical terms, in algebraic terms through a group presentation by generators and relations (the Artin braid group) and in terms of automorphisms of the free group.

Taking this into account, we have defined an important representation of the braid group, the Burau representation. We have proved that for all $n \geq 2$ it is reducible and we have given its reduced form. Then, we have added the detailed study of its faithfulness, arriving to affirm that for $n = 1$, 2 and 3 the representation is faithful. When $n \geq 5$, we have indicated how it can be seen that it is not faithful. [1]

In addition, we have also proved that the Burau representation can also be defined through the Fox calculus.

From my point of view, there are three possible paths to follow after this work, from which, in the future, I would like to explore the first and the third ones:

1. Go in depth in braid theory: understanding the proof of the fact that two braid diagrams present isotopic braids if and only if these diagrams are R-equivalent (Theorem 2.9) and of the fact that two polygonal n-braids are isotopic if and only if one can be transformed into the other one by a finite number of elementary moves (Theorem 2.11) and, also, the proof of the fact that the n-braid group is isomorphic to the Artin braid group (Theorem 2.22).

2. Make the necessary calculations and computations to check that ψ_5 is not faithful.

3. Study other representations of the braid group and their properties.

[1] We do not mention the case when $n = 4$ because, as we have said in the introduction, it have not been solved yet.

Bibliography

[1] Colin C. Adams, *The knot book*, American Mathematical Society, Providence, RI, 2004. An elementary introduction to the mathematical theory of knots, Revised reprint of the 1994 original. MR2079925

[2] Joan S. Birman, *Braids, links, and mapping class groups*, Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1974. Annals of Mathematics Studies, No. 82. MR0375281

[3] Gerhard Burde, Heiner Zieschang, and Michael Heusener, *Knots*, extended, De Gruyter Studies in Mathematics, vol. 5, De Gruyter, Berlin, 2014. MR3156509

[4] Richard H. Crowell and Ralph H. Fox, *Introduction to knot theory*, Springer-Verlag, New York-Heidelberg, 1977. Reprint of the 1963 original, Graduate Texts in Mathematics, No. 57. MR0445489

[5] V. F. R. Jones, *Hecke algebra representations of braid groups and link polynomials*, Ann. of Math. (2) **126** (1987), no. 2, 335–388. MR908150

[6] Christian Kassel and Vladimir Turaev, *Braid groups*, Graduate Texts in Mathematics, vol. 247, Springer, New York, 2008. With the graphical assistance of Olivier Dodane. MR2435235

[7] Emmanuel Kowalski, *An introduction to the representation theory of groups*, Graduate Studies in Mathematics, vol. 155, American Mathematical Society, Providence, RI, 2014. MR3236265

[8] Kunio Murasugi and Bohdan I. Kurpita, *A study of braids*, Mathematics and its Applications, vol. 484, Kluwer Academic Publishers, Dordrecht, 1999. MR1731872