UNIVERSITAT DE BARCELONA

Facultat de Matemàtiques
i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

BURNSIDE'S $p^a q^b$ THEOREM

Autor: Víctor Santiago Blanco

Director: Dr. Santiago Zarzuela Armengou
Realitzat a: Departament de Matemàtiques
i Informàtica

Barcelona, 12 de juny de 2022

Abstract

The main goal of this work is to prove the Burnside's $p^a q^b$-theorem, which states that every group G of order $p^a q^b$ is solvable. In order to justify the importance of this well-known theorem, a first chapter about solvable groups will be included, in which will be analyzed some properties of solvable groups. The provided proof will use Representation and Character Theory which will be studied in depth.

Resum

El principal objectiu d'aquest treball és provar el teorema $p^a q^b$ de Burnside, que afirma que tot grup G d'ordre $p^a q^b$ és resoluble. Per justificar la importància d'aquest conegut teorema, s'inclou un primer capítol sobre grups resolubles, en el qual s'analitzaran algunes propietats d'aquest tipus de grups. La prova que es donarà utilitza teoria de representacions i teoria de caràcters que seran estudiades en profunditat.

Acknowledgements

I would like to express my gratitude to my advisor, Dr. Santiago Zarzuela Armengou, for guiding me throughout the past few months and providing me a very exhaustive monitoring. Sharing with him the exploration of this field of Mathematics has been a great pleasure for me. I would also like to thank my close family for making possible for me to pursue my studies and always supporting me in my successes and my failures. I would also like to thank my very best mathematician friends, Sixte Oriol Llenas and Pol Vallés, for being there always helping each other in our struggles with Mathematics. Finally, I would like to thank my girlfriend, Paula Hernandez, for being always there for me and giving me her support. It could not have been possible without all of you.

2020 Mathematics Subject Classification. 20C15, 20C05, 20F19

Contents

Introduction

This project aims to prove the Burnside's $p^\alpha q^\beta$ theorem, that states that *If G is a group of order $p^\alpha q^\beta$ with p and q primes, then G is solvable.* The proof I will provide follows the same lines as the proof of Burnside himself. It is not a pure group-theoretical proof as it uses character theory.

William Burnside (2 July 1852 – 21 August 1927) was an English mathematician mainly known for his contributions to finite group theory. He studied Mathematics at the University of Cambridge. He lectured at Cambridge for ten years and later he became professor of Mathematics at Royal Naval College, at Greenwich. He dedicated most of his research to finite group theory. This was not a widely studied subject in Britain in the late 19th century, and it took some years for his research in this area to gain recognition.

Burnside's most important contribution to group theory was the development of the area of group representations, where he helped to grow some of the foundational theory, complementing with Ferdinang Georg Frobenius. In 1904 he published the proof to the famous $p^\alpha q^\beta$ theorem, which was purely character theory related. He had published the classic work *Theory of Groups of Finite Order* in 1897 but later he released a new version in 1911 where he included character theory. This edition was for many decades the standard work in the field.

After Burnside's proof, many other mathematicians have struggled to give a "pure group-theoretic" proof of it. It was not until sixty years later, when mathematician John Griggs Thompson pointed out that a character-free proof of Burnside's $p^\alpha q^\beta$ theorem may be extracted from his work about the classification of the n-groups, done between 1968 and 1974, and also from his proof of the famous Feit-Thompson Theorem (1963) which states that *every group of finite odd order is solvable*. However, he did not provide any proof. David M. Goldschmidt gave an explicit proof of it in 1970 [3], restricting himself to the case in which p and q are odd. It was not until 1972, when Helmut Bender gave a complete character-free proof of the whole theorem [1]. The problem was that it was widely longer and more complicated than the original by Burnside. Nevertheless, in 1973, Japanese mathematician Hiroshi Matsuyama joined Goldschmidt and proved the case of $p^\alpha 2^\beta$ [6] and gave a purely group-theoretic proof of Burnside's theorem that was much simpler than Bender's.

My interests on group theory began when I attended to the curricular course of

“Estructures Algebraiques” (Algebraic Structures), where I saw this concept for the first time. In this course we studied the very basics of group theory, as well as ring theory. But it was not until I studied Galois Theory in the course “Equacions Algebraiques” (Algebraic Equations) when I became really interested in group theory. Seeing the Galois groups encouraged me to investigate on some finite group behaviours and it opened a wide world to me: the finite group theory.

I became really excited about finite group theory and my first idea was to dedicate this project to the Classification of Finite Groups theorem, which I found really interesting. Later on, I realised the immensity of this field of study and I decided to change the subject to a more specific result that I could understand knowing only the basics of group theory from the degree.

Santiago Zarzuela Armengou proposed me a few results regarding finite groups, that could be suitable for such project, and then I decided to work on Burnside’s Theorem. Not only because it is a theorem about finite groups, but also because the proof required a long investigation and learning process about representations and characters which I assume is very useful in the study of group theory in general. Therefore, it was necessary that I learned the basics of this essential tool that is representation theory and also character theory.

Both representation and character theory are completely new to me, but I had an advantage. Representation theory is intimately related to module theory, as it will be seen later, and I was lucky enough to be able to choose an optional course of the degree, called “Introducció a l’Àlgebra Commutativa” (Introduction to Commutative Algebra), in where we studied deeply the basics of modules over commutative rings. Although here we will take non-commutative rings, the theory seen in the course classes was really helpful and useful in most of the cases.

This project will contain the basic study of representation theory, restricting it to the results that will apply to the particular case we are interested in. As I said, it is intimately related to module theory, hence there will be results and definitions that I will assume already known. Moreover, I will study the basic concepts of character theory, and those necessary results to prove our main theorem. Finally, in the last chapter, I will give the proof to the theorem that was given by Burnside.

Furthermore, as this theorem states a characterization of solvable groups, I will study in depth some pure group-theoretic concepts of solvable groups, that have not been seen in any of the courses I have taken. This will illustrate the importance of a group to be solvable and hence, the relevance of the theorem we are aiming to prove.

Chapter 1

Solvable groups

This chapter will start from basic definitions such as the main definition of solvable group, followed by some important results about solvability. The main objective is to understand and being able to manage solvable groups and composition series. There will also be stated and proved based on [7] the famous Zassernhaus Lemma, known also as Butterfly Lemma.

Furthermore, I will add the Jordan-Hölder Theorem and I will present the "group extension problem" to illustrate how important are solvable groups. This will be based on [7] and also on [4]. To end the sections I will add the definitions of "commutator", "derived series" and finally there will be stated some relevant results about them.

1.1 Normal and composition series. Solvable groups

Suppose we have a group G and we want to study this group's properties from its subgroup's properties. We already know a technique to do so: the Sylow Theorems and the p-Sylow groups. Here we will define another way to study a group from its subgroups that will consist on keep simplifying the group into its normal subgroups and quotients.

Take any group G and we will try to get a series of subgroups $G = H_0 \supseteq H_1 \supseteq H_2 \supseteq \cdots$ such that they are normal in each "parent" group, i.e., $H_{i+1} \lhd H_i$ and hence we will be able to compute the quotient group $Q_i = H_i / H_{i+1}$. We will try to find such series of H_i that make these quotients simple. This will be defined later as *composition series*. This way, we will be able to recover information from these "factors" to the main group.

Nevertheless, there are two problems that need to be taken care of carefully. The first one is the fact that this series may not be finite, so this technique mentioned will be useful if the group allows to find finite series. The other main problem is that, supposing one can find this finite series, we are trying to recover information of a group G by its simple subgroup quotients. Hence, if we knew the simple groups, we would have information of G. But simple groups are not that simple as they seem. Nonetheless, if we add the condition that these quotients should be also abelian, then it would be very easy to study them. This is the particular case of *solvable* groups.

Definition 1.1.1 (Normal series)**.** A *normal series*[1] of a group G is a sequence of subgroups

$$G = G_0 \supseteq G_1 \supseteq G_2 \supseteq \cdots \supseteq G_t$$

with each G_{i+1} a normal subgroup of G_i; the *factor groups* of this series are the quotient groups

$$G_0/G_1,\ G_1/G_2,\ \ldots,\ G_{t-1}/G_t.$$

and the *length* of the series is the number of non trivial factor groups.

Proposition 1.1.2. *Note that if G is a finite group if and only if all the factor groups are finite. In this case, we can compute $|G| = \prod |G_t/G_{t+1}|$.*

Proof. Let $G = G_0 \supsetneq G_1 \supsetneq \cdots G_t$ a normal series of G. Then $|G| = |G_1| \cdot |G/G_1|$. But now, applying this to G_1 we obtain $|G_1| = |G_2| \cdot |G_1/G_2|$. In general, we get $|G_{t-1}| = |G_t| \cdot |G_t/G_{t-1}|$ hence we get the formula. □

Once seen the definition of a normal series and a sequence of subgroups as the one in 1.1.1, we might consider the case in which the factor groups G_i/G_{i+1} are simple, as this will give more information about the main group, because one can study easily the simple groups. The idea of the decomposition is to make it as simple as possible to get more information of the group.

Definition 1.1.3 (Composition series)**.** Let G be a group. A *composition series* is a normal series that eventually ends at $\{1\}$ all of whose nontrivial factor groups are simple. The nontrivial factor groups of a composition series are called *composition factors* of G. The *length* of a composition series is the number of nontrivial factors.

Note that a group need not to have a composition series, as it might have a series with factor groups non simple. For example, the abelian group $\mathbb{Z}$ with respect to the sum operation has not simple subgroups. Hence, one cannot find a composition series.

Proposition 1.1.4. *Every finite simple group G has a composition series.*

Proof. Take a group G with the minimum order possible such that G does not have a composition series. Now, G is not simple, otherwise $G \supsetneq \{1\}$ is a composition series. Hence, G has a proper normal subgroup H. Since G is finite, we may assume that H is a maximal normal subgroup, so that G/H is a simple group. But $|H| < |G|$, so that H has a composition series: say $H = H_0 \supsetneq H_1 \supsetneq \cdots \supsetneq \{1\}$. Hence, $G \supsetneq H \supsetneq H_1 \supsetneq \cdots \supsetneq \{1\}$ is indeed a composition series for G, a contradiction. □

[1] This terminology is not quite standard. We know that normality is not transitive; that is, if $H \subseteq K$ are subgroups of a group G, then $H \lhd K$ and $K \lhd G$ do not force that $H \lhd G$. A subgroup $H \subseteq G$ is called a *subnormal subgroup* if there is a chain $G = G_0 \supseteq G_1 \supseteq \cdots \supseteq G_t = H$ with $G_i \lhd G_{i-1}$ for all $i \geq 1$. Normal series as defined in the text are called subnormal series by some authors; they reserve the name *normal series* for those series in which each G_i is a normal subgroup of the big group G.

Now that we have seen what a composition series is, and why are they important, we will add more conditions to the factor groups of the composition series so that the series provides more information about the group. In the composition series we imposed the finiteness of the series and the simplicity of its factor groups. Now we add the condition of the abelianity of the factor groups. A group G which can have this series will be called *solvable*.

Definition 1.1.5 (Solvable)**.** A group G is called *solvable*[2] if it has a composition series with abelian composition factors.

Example 1.1.6. Let's look at some examples.

(1) Every finite abelian group is solvable. Indeed, as in 1.1.4 we have seen that a finite group does have a composition series and if it is abelian, so will the composition factors be.

(2) One can easily prove that S_3 and S_4 are solvable groups, by giving explicitly a chain of subgroups. Nonetheless, for $n \geq 5$, S_n is not solvable.

(3) Another examples of solvable groups are the quaternion group H_8, the dihedral groups D_{2n} or the p-groups.

In general, it seems quite hard to check whether a group has a normal series, in order to prove it is a solvable group. In the following result, there will be given characterizations of solvable groups that will allow us to check if a group is solvable by studying relative characteristics of it.

Theorem 1.1.7. *(1) Every quotient of a solvable group is itself solvable.*

(2) Every subgroup of a solvable group is itself solvable.

(3) Given G a group, if $H \lhd G$ and G/H are both solvable groups, then G is also solvable.

The proof of these three results is not given as it was seen in "Estructures Algebraiques". Nevertheless, at the end of this chapter, there will be introduced a new notion which will give us another equivalent definition of solvable groups. Using this new definition, the proof for this theorem will be shorter and straightforward.

Example 1.1.8. Let's look at some examples and how given these properties can a solvable group be characterized in an easier way.

(1) A non-abelian simple group G is not solvable, for its only proper normal subgroup is $\{1\}$, and $G/\{1\} \cong G$ is not abelian. We see that, then, following our definition, all its factor groups are G himself and it is non-abelian.

[2] The name "solvable" comes from Galois Theory, as the "solvable groups" where those generated by the roots of polynomials "solvable by radicals". That is, Galois discovered solvable groups even before they were defined.

(2) For $n \geq 5$, S_5 is not solvable. If it was, then all its subgroups would be solvable as well, following the theorem 1.1.7. But S_5 has A_5 as a subgroup and we already know that A_5 is simple, so it cannot be solvable by the previous example, ergo S_5 cannot be solvable neither. As $A_5 \subseteq S_5 \subseteq S_n$, $n \geq 5$, we have that S_n is not solvable. This is very relevant to Galois Theory, as it will be needed to prove that a polynomial of degree $n \geq 5$ cannot be solved by radicals.

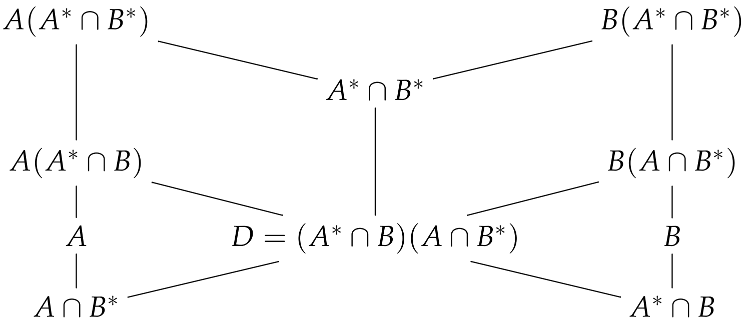
(3) If H and K are solvable groups, then $H \times K$ is solvable. Indeed, as $(H \times K)/H \cong K$.

To end the section, we will prove a more general result that does not stand only in the solvable groups, but also in more general groups. The theorem is known as the Butterfly Lemma or the Zassenhaus, in honor to Hans Zassenhaus[3], and it is called Butterfly because it has a diagram that recalls a butterfly. Although [7] does not agree with that.

Lemma 1.1.9 (Zassenhaus)**.** *Given four subgroups $A \lhd A^*$ and $B \lhd B^*$ of a group G, then $A(A^* \cap G) \lhd A(A^* \cap B)$, $B(B^* \cap A) \lhd B(B^* \cap A^*)$, and there is an isomorphism*

$$\frac{A(A^* \cap B^*)}{A(A^* \cap B)} \cong \frac{B(B^* \cap A^*)}{B(B^* \cap A)}$$

This isomorphism is symmetric in the sense that the right side is obtained from the left side by changing the symbols A and B. The following diagram is what gives this lemma the name of *butterfly lemma*.



Proof. Let's begin with the proof of the Zassenhaus lemma. We begin by proving $(A \cap B^*) \lhd (A^* \cap B)$ and the fact $(A^* \cap B) \lhd (A^* \cap B^*)$ is analogous so we omit it. Then we will prove that there is an isomorphism

$$\frac{A(A^* \cap B^*)}{A(A^* \cap B)} \longrightarrow \frac{A^* \cap B^*}{D}$$

which by the symmetry commented before will prove the isomorphism wanted.

- Claim that $(A \cap B^*) \lhd (A^* \cap B^*)$: that is, if $c \in A \cap B^*$ and $x \in A^* \cap B^*$ then $xcx^{-1} \in A \cap B^*$. Now, $xcx^{-1} \in A$ because we have chosen $c \in A$ and also because

[3] Hans Julius Zassenhaus (28 May 1912 – 21 November 1991) was a German mathematician, known for work in many parts of abstract algebra, and as a pioneer of computer algebra. Note from [11]

$x \in A^*$ and then, as $A \lhd A^*$, $xcx^{-1} \in A$. Easier is to see that $xcx^{-1} \in B^*$, as $x, c \in A \cap B^* \subset B^*$. Hence, $(A \cap B^*) \lhd (A^* \cap B^*)$. Similarly, we have that $(A^* \cap B) \lhd (A^* \cap B^*)$.

Therefore, the subset $D = (A^* \cap B)(A \cap B^*)$ is a normal subgroup of $A^* \cap B^*$ because it is generated by two normal subgroups[4].

- Now we prove that $A(A^* \cap B^*)/A(A^* \cap B) \cong (A^* \cap B^*)/D$ and by symmetry we will have also proved that $B(A^* \cap B^*)/B(A \cap B^*) \cong (A^* \cap B^*)/D$ and hence we will have the isomorphism wanted.

  Define $\varphi : A(A^* \cap B^*) \to (A^* \cap B^*)/D$ by $\varphi(ax) = [x]$, where $a \in A$ and $x \in A^* \cap B^*$ and $[x]$ represents the left or right coset, i.e. the equivalence class by the quotient D. We need to check that φ is well defined, that is a morphism, and then the isomorphism.

  Now take $ax = a'x'$, where $a, a' \in A$ and $x, x' \in A^* \cap B^*$. We can arrange that to obtain $(a')^{-1}a = x'x^{-1}$. As $a, a' \in A$, $(a')^{-1}a \in A$ and as $x, x' \in A^* \cap B^*$, $x'x^{-1} \in A^* \cap B^*$. Then everything belongs to $A \cap (A^* \cap B^*) = A \cap B^* \subseteq D$.

  Also φ is a homomorphism: $axa'x' = a''xx'$ where $a'' = a(xa'x^{-1}) \in A$ because $A \lhd A^*$ and so $\varphi(axa'x') = \varphi(a''xx') = [xx'] = [x][x'] = \varphi(ax)\varphi(a'x')$.

  It is easy to see that φ is surjective, it gets values from a set which contains $A^* \cap B^*$. Also it is easy to check that $\operatorname{Ker}\varphi = A(A^* \cap B)$ as $ax \in \operatorname{Ker}\varphi$ if and only if $\varphi(ax) = [x] = 0$, i.e. if and only if $ax \in D$ and as $a \in A$ and $x \in (A^* \cap B^*)$, this only happens if $x \in (A^* \cap B)$, i.e. if and only if $ax \in A(A^* \cap B)$. By the First Isomorphism Theorem we have the proof.

□

Note that the Second Isomorphism Theorem is a particular case of the Zassenhaus Lemma, i.e. the Zassenhaus Lemma implies the Second Isomorphism Theorem: if S and T are subgroups of a group G with $T \lhd G$, then $TS/T \cong S/(S \cap T)$; set $A^* = G$, $A = T$, $B^* = S$ and $B = S \cap T$.

1.2 The Jordan-Hölder Theorem

The next goal is to understand and prove the Jordan-Hölder[5] Theorem, which states that every two composition series of a group G are equivalent. To understand and make the proof we will need some previous motivation and results.

[4] Take G a group and $K \lhd G$ and $H \lhd G$. Then $HK \lhd G$. Indeed, we want to check that for any $x \in G$ and $y \in HK$ we have $xyx^{-1} \in HK$. As $y \in HK$ we can write $y = hk$ where $h \in H$ and $k \in K$. Then, $xyx^{-1} = xhkx^{-1} \in HK$ because $xh \in H$ and $kx^{-1} \in K$ for H and K being subgroups of G and $x \in G$.

[5] This important theorem is called after Camille Jordan (Lyon, 1838 - 1922) and Otto Hölder (Stuttgart, 1859 - 1937). Jordan is known for his fundational work in Group Theory, and for his influential *Cours d'analyse* (source: [10]). Not to be confused with the geodesist Wilhelm Jordan, from Gauss-Jordan elimination. Hölder is noted for many theorems, from abstract algebra to differential equations (source [12]).

Firstly, note that if G is a solvable group, then there is a composition series whose factor groups are abelian. For the definition of composition series, these factor groups are also simple. Hence, the only simple abelian groups are those cyclic of prime order, and hence we have that each factor group of a solvable group is finite.

Example 1.2.1. Let's see all these definitions in a particular case to see them more clearly. Take $G = \langle x \rangle$ of order 30. As G is abelian, we have automatically the normality of all subgroups of G, hence I will just assume that. We can calculate easily these two composition series of G:

$$G = \langle x \rangle \supsetneq \langle x^2 \rangle \supsetneq \langle x^{10} \rangle \supsetneq \{1\}; \tag{1.1}$$

$$G = \langle x \rangle \supsetneq \langle x^5 \rangle \supsetneq \langle x^{15} \rangle \supsetneq \{1\}; \tag{1.2}$$

The factor groups of the series (1.1) are $\langle x \rangle / \langle x^2 \rangle \cong \mathbb{Z}_2$, $\langle x^2 \rangle / \langle x^{10} \rangle \cong \mathbb{Z}_5$, and $\langle x^{10} \rangle / \{1\} \cong \mathbb{Z}_3$; while the factor groups of the second series (1.2) are $\langle x \rangle / \langle x^5 \rangle \cong \mathbb{Z}_5$, $\langle x^5 \rangle / \langle x^{15} \rangle \cong \mathbb{Z}_3$ and $\langle x^{15} \rangle / \{1\} \cong \mathbb{Z}_2$. This gives a clear example of how the factor groups are the same but they appear in a different order. This is the essence of the Jordan-Hölder Theorem.

In order to prove the Jordan-Hölder Theorem and to realise its implications we are going to give a brief definition that will make clearer its statement. Once seen the definition of a normal series, one could ask him or herself if a group could have more than one normal series, and when would they be equivalent, in some way. It is easy to see that there can be more than one normal series for a group G, but we are going to see when are they equivalent. In addition, we will be giving a definition of a "simpler normal series" which we will call *refinement* and we will prove that any two normal series have equivalent refinements. This last statement is known as *Schreier Refinement Theorem* 1.2.4 and it will be the key to the Jordan-Hölder Theorem.

Definition 1.2.2 (Equivalent normal series)**.** Two normal series of a group G are *equivalent* if there is a bijection between the list of nontrivial factor groups of each so that corresponding factor groups are isomorphic.

In view of this definition, we can say that the two composition series of the previous example 1.2.1 are equivalent. The Jordan-Hölder Theorem states that any two composition series of a group G are equivalent. Nevertheless, there is a more general result, due to Schreier, which I will do first and then Jordan-Hölder Theorem will be a corollary of it. We need a previous definition.

Definition 1.2.3 (Refinement)**.** A *refinement* of a normal series of a group G is a normal series $G = N_0 \supseteq \ldots \supseteq \ldots$ having the original series as subsequence.

In other words, a refinement of a normal series is a normal series obtained by the original one by inserting more subgroups.

Notice that if we have a composition series then all refinements are insignificant; one can merely repeat terms (if G_i/G_{i+1} is simple, there will be no more proper nontrivial normal subgroups and, hence, there is no intermediate subgroup L with $G_i \supsetneq L \supsetneq G_{i+1}$ and $L \lhd G_i$). Therefore, any refinements of a composition series is equivalent to the original composition series.

Theorem 1.2.4 (Schreier Refinement Theorem)**.** *Any two normal series ending at* $\{1\}$

$$G = G_0 \supseteq G_1 \supseteq \cdots \supseteq G_n = \{1\}$$

and

$$G = N_0 \supseteq N_1 \supseteq \cdots \supseteq N_k = \{1\}$$

of a group G have equivalent refinements.

Proof. The idea consists on inserting a copy of the second series between each pair of adjacent terms in the first series. In more detail, for each $i > 0$, define

$$G_{ij} := G_{i+1}(G_i \cap N_j)$$

which is a subgroup because $G_{i+1} \lhd G_i$. Since $N_0 = G$ we have now

$$G_{i0} = G_{i+1}(G_i \cap N_0) = G_{i+1}G_i = G_i$$

and since $N_k = \{1\}$ we have

$$G_{ik} = G_{i+1}(G_i \cap N_k) = G_{i+1}\{1\} = G_{i+1}$$

Therefore, we can create a series of G_i and G_{i+1} like this:

$$G_i = G_{i0} \supseteq G_{i1} \supseteq G_{i2} \supseteq \cdots \supseteq G_{ik} = G_{i+1}$$

and if we do this for every i, it gets interspersed in our first original series, so we created another series with a total of nk elements having the first original series as a subsequence. Now, we can do the analogous with $N_{ji} = N_{j+1}(N_j \cap G_i)$ and create another series which the second original series as a sub-sequence, and with exactly nk elements. Careful because nobody says that these are normal series, i.e. nobody says that these are refinements of the original series respectively. Lucky for us, we have Zassenhauss Lemma. For each i, j, Zassenhaus Lemma (1.1.9), for the subgroups $G_{i+1} \lhd G_i$ and $N_{j+1} \lhd N_j$, says both sub-sequences are normal series, hence are refinements because the Lemma gives the isomorphism

$$\frac{G_{i+1}(G_i \cap N_j)}{G_{i+1}(G_i \cap N_{j+1})} \cong \frac{N_{j+1}(N_j \cap G_i)}{N_{j+1}(N_j \cap G_{i+1})};$$

that is, with our double-index notation,

$$G_{i,j}/G_{i,j+1} \cong N_{i,j}/N_{i,j+1}$$

and as the association $G_{i,j}/G_{i,j+1} \to N_{i,j}/N_{i,j+1}$ is a bijection, by the definition of refinement and equivalence we have that the two refinements are equivalent. □

With this result, we are able now to prove Jordan-Hölder Theorem in just a few lines.

Corollary 1.2.5 (Jordan-Hölder Theorem). *Any two composition series of a group G are equivalent. In particular, the length of a composition series, if one exists, is an invariant of G.*

Proof. As we remarked earlier, this is a particular case of the Schreier Theorem (1.2.4). Any refinement of a composition series is equivalent to the original composition series, as every factor group is simple, hence it follows that any two composition series are equivalent by the Theorem. □

A direct implication of this theorem is the Fundamental Theorem of Arithmetic. Rotman in [7] gives a proof based on this last theorem and although it is not our main purpose, it is a clear example of how important and relevant is Jordan-Hölder Theorem in other contexts rather than pure Group Theory.

Theorem 1.2.6 (Fundamental Theorem of Arithmetic). *Every integer $n \geq 2$ has a factorization into primes, and the prime factors and their multiplicities are uniquely determined by n.*

Proof. Since the group $\mathbb{Z}_n$ is finite, it has a composition series. Let $S_1, \ldots, S_t$ be the factor groups. Now, an abelian group is simple if and only if is of prime order. Since $n = |\mathbb{Z}_n|$ is the product of the orders of the factor groups (by 1.1.2), we have proved that n is a product of primes. Moreover, the Jordan-Hölder Theorem (1.2.5) gives the uniqueness of the (prime) orders of the factor groups and their multiplicities. □

Furthermore, one can observe that non isomorphic groups can have the same composition factors. For example, $\mathbb{Z}_4$ and $\mathbb{V} = \mathbb{Z}_2 \times \mathbb{Z}_2$ have the same factor groups, which are $\mathbb{Z}_2$ and $\mathbb{Z}_2$, but they are not isomorphic groups.

The next example will illustrate how to use Jordan-Hölder Theorem to prove that a particular group is not solvable.

Example 1.2.7. Let $G = \mathrm{GL}(2, \mathbb{F}_4)$ be the general group of all 2×2 non-singular matrices with entries in the field $\mathbb{F}_4$ with four elements. Now, $\det : G \to (\mathbb{F}_4)^*$, where $(\mathbb{F}_4)^* \cong \mathbb{Z}_3$ is the multiplicative group of non-zero elements of $\mathbb{F}_4$. Since $\operatorname{Ker}\det = \mathrm{SL}(2, \mathbb{F}_4)$, the special linear group consisting of those matrices of determinant 1, there is a normal series

$$G = \mathrm{GL}(2, \mathbb{F}_4) \supsetneq \mathrm{SL}(2, \mathbb{F}_4) \supsetneq \{1\}$$

The factor group of this normal series are $\mathbb{Z}_3$ and $\mathrm{SL}(2, \mathbb{F}_4)$. It is true that $\mathrm{SL}(2, \mathbb{F}_4)$ is a non-abelian simple group (there exists the isomorphism $\mathrm{SL}(2, \mathbb{F}_4) \cong A_5$) and so this series is a composition series. Observe that there is a factor group with order not prime, as $|A_5| = 5!/2 = 60$. We cannot yet conclude that G is not solvable, for the definition of solvability requires that there be some composition series, not necessarily this one, having factor groups of prime order. However, the Jordan-Holder Theorem (1.2.5) says that if one composition series of G has all its factor groups of prime order, then so does every other composition series. We may now conclude that $\mathrm{GL}(2, \mathbb{F}_4)$ is not a solvable group.

1.3 Group extensions and the extension problem

In this section we are going to give a little change of view to discuss more deeply the significance of Jordan-Hölder Theorem and solvability. We will talk about "group extensions" and we will see an important mathematical problem called the "extension problem" for groups. Then we will see that this is not that problematic when the groups treated are solvable, and hence the importance of solvability again. To get started there will be defined the notion of *group sequence* and then there will be given the definition of the notion of *group extension*, together with some examples. Moreover, one will see the definition of *equivalent group extensions* and then there will be proposed the "Extension Problem"

Definition 1.3.1 (Group sequence)**.** A *Group sequence* is a sequence of groups and group homomorphisms

$$G_0 \xrightarrow{f_0} G_1 \xrightarrow{f_1} \cdots \xrightarrow{f_r} G_{r+1} \xrightarrow{f_{r+1}} \cdots$$

and it is said to be *exact* at G_i if $\mathrm{Im}(f_i) = \mathrm{Ker}(f_{i+1})$. A group sequence is said to be exact if it is exact in each step (i.e. for each $i \geq 0$, $\mathrm{im}(f_i) = \mathrm{Ker}(f_{i+1})$). Finally, a group sequence is called *short exact sequence* if it is an exact sequence of the form

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

Then, following the above definition, we have here that $\mathrm{im}(f) = \mathrm{Ker}(g)$ and that f is a monomorphism and g an epimorphism. Furthermore, the First Isomorphism Theorem gives us the isomorphism $B/\mathrm{im}(f) \cong C$, while the fact that $\mathrm{Ker}\, g = \mathrm{im} f$ gives us that $f(A) \lhd B$

Definition 1.3.2 (Group extensions)**.** Given two groups K and H, a group G is called extension of K by H if there exists a small exact sequence

$$0 \to K \xrightarrow{\iota} G \xrightarrow{\pi} H \to 0.$$

Note that $\iota(K) \lhd G$ and also that $G/\iota(K) \cong H$ for the First Isomorphism Theorem. The fact that ι is a monomorphism gives us the isomorphism $\iota(K) \cong K$.

Example 1.3.3. Let's study some examples given the definition.

(i) The direct product $K \times Q$ is an extension of K by Q. Indeed, we can consider the following short exact sequence

$$0 \longrightarrow K \xrightarrow{\iota} K \times Q \xrightarrow{\pi} Q \longrightarrow 0.$$

This is a well defined sequence. We can define $\iota : K \to K \times Q$ as the inclusion, and then we get a monomorphism. Also we can define the $\pi : K \times Q \to Q$ as for any $(k, q) \in K \times Q$, $\pi(k, q) = q$ and we get trivially an epimorphism. Then we check that $\mathrm{Ker}\, \pi = \mathrm{im} \iota$ which is trivial. We have, hence, an exact short series and then $K \times Q$ is indeed an extension of K by Q.

(ii) Another example, which is indeed an example from the last example, could be the following. Given the groups $\mathbb{Z}_3$ and $\mathbb{Z}_2$ we want to look for an extension of $\mathbb{Z}_3$ by $\mathbb{Z}_2$. We know that $\mathbb{Z}_3 \times \mathbb{Z}_2 \cong \mathbb{Z}_6$ is one, so we can take it. But one may notice that S_3 is another extension of $\mathbb{Z}_3$ by $\mathbb{Z}_2$.

Following the last example, one may wonder some questions about the extensions. We already saw in the first example that we can always find an extension of K by Q, for any given two groups K, Q, which is $K \times Q$. Hence, an extension of K by Q may be viewed as a "product" of K and Q. But is this product the only possibility? Are there even non-isomorphic groups to $K \times Q$ that can also be extensions of K by Q? Is there some notion of "extension equivalence"? These questions are treated more deeply in "Chapter 7: Extensions and Cohomology" chapter in [8].

There exists a notion of equivalence regarding group extensions, and it is defined as follows.

Definition 1.3.4 (Equivalent group extensions)**.** We say that extensions

$$0 \to K \xrightarrow{\iota} G \xrightarrow{\pi} H \to 0 \qquad \text{and} \qquad 0 \to K \xrightarrow{\iota'} G' \xrightarrow{\pi'} H \to 0$$

are equivalent[6] if there exists a group isomorphism $f : G \to G'$ that makes commutative the following diagram

$$\begin{tikzcd}
0 \arrow[r] & K \arrow[r, "\iota"] \arrow[d, "="] & G \arrow[r, "\pi"] \arrow[d, "f"] & H \arrow[r] \arrow[d, "="] & 0 \\
0 \arrow[r] & K \arrow[r, "\iota'"] & G' \arrow[r, "\pi'"] & H \arrow[r] & 0
\end{tikzcd}$$

We can take then equivalence classes and the set of all the equivalent classes of extensions of K by Q are given by the set $\mathbf{Ext}^1_{\mathbb{Z}}(K, Q)$. This set is, in fact, a group, but we will not discuss this as it is getting very deep into cohomology of groups and it diverges from the purpose of this notes.

This explanation about group extensions was made only to illustrate how useful can be the solvable groups into this matter: the group extension problem is the problem of classification of all the possible group extensions modulus extension equivalence. I.e., given two groups K and Q we want to find all possible extension groups of K from Q modulus extension equivalence. Thus, if we could solve this problem, we would be able to "recover" the information of a group G given a normal subgroup $N \lhd G$ and its quotient G/N. This would be particularly useful in the context of normal series and solvability. For example, let's suppose that the group G has the ending normal series

$$G = K_0 \supseteq K_1 \supseteq \cdots \supseteq K_{n-1} \supseteq K_n = \{1\}$$

with factor groups $Q_1, \ldots, Q_n$, where $Q_i = K_{i-1}/K_i$, for all $i \geq 1$. Then, if we knew all these factors, we could recover the information of G recursively like this: $K_{n-1} = Q_n$,

[6] Sometimes they are also called *congruent extensions*.

but K_{n-2} is an extension of K_{n-1} by Q_{n-1}, for $K_{n-1} \lhd K_{n-2}$ and $K_{n-2}/K_{n-1} = Q_{n-1}$. Then, going backwards we obtain that G is an extension of K_1 by Q_1.

This means that we could recover G from all its factor groups by just making the "product" of them (see 1.3.3). But then we face two problems:

1. A group could have more than one non equivalent normal series, and
2. we don't know if there are more non equivalent group extensions, apart from the "product" already seen.

Regarding the first problem, Jordan-Hölder Theorem 1.2.5 gives a solution for it, if our series is a composition series. Indeed, as it is a unique factorization theorem: the factors in this product, namely, the composition factors of G, are uniquely determined by G. Therefore, we could survey all finite groups if we knew the finite simple groups and we could solve the extension problem. In fact, all the finite simple groups are known and proved in the *Classification Theorem of Finite Simple Groups*. This shows the importance of composition series and the Jordan-Hölder Theorem.

Regarding the second problem, it is unsolved in the sense that no one knows a way, given K and Q, to compute the exact number of non-isomorphic extensions of K by Q.

As it was said at the beginning, this was like a subsection talking about the importance and relevance of the Jordan-Hölder Theorem, but it is not the main object of the notes, so we have to keep on track.

1.4 Commutators and derived series

In this section we will see a new notion related to group series, which is the *derived series* that will be defined from a particular operation between groups: the *commutator*. This is particularly useful as it will give us some results that will characterize when a group is abelian, not simple or solvable. Firstly, there will be stated the basic definition of commutator of a group and given some properties and characterizations. Secondly, there will be given the definition of *derived series* and its implications regarding solvability of a group. Finally there will be given alternative proofs to our "main properties" of solvable groups already seen on the first section of this chapter.

Definition 1.4.1 (Commutator)**.** If G is a group and $x, y \in G$, then their *commutator* is the element

$$[x, y] := xyx^{-1}y^{-1}$$

Moreover, if X and Y are subgroups of a group G, then we define their *commutator* by

$$[X, Y] := \langle [x, y] \ : \ x \in X, \ y \in Y \rangle$$

In particular, the *commutator subgroup* G' of a group is

$$G' = [G, G]$$

the subgroup generated by all the commutators.

Note that for any subgroups X, Y of a group G, the set $[X, Y]$ defined here needs not to be closed under products, i.e. it may not be a subgroup of G. Nevertheless, the commutator subgroup G' *is* a subgroup because it is defined as the subgroup generated by all the commutators of G.

It is clear that two elements x and y in a group G commute if and only if their commutator $[x, y]$ is 1. The next proposition generalizes this observation.

Proposition 1.4.2. *The commutator subgroup G' is a normal subgroup of G and G/G' is abelian. Moreover, if $H \lhd G$ and G/H is abelian, then $G' \subseteq H$.*

Proof. First we may notice that, for $x, y \in G$,

$$[x,y]^{-1} = (xyx^{-1}y^{-1})^{-1} = yxy^{-1}x^{-1} = [y,x].$$

Therefore, each element of G' is a product of commutators. But any conjugate of a commutator is another commutator. Indeed, given $a \in G$ and $x, y \in G$ we have

$$a[x,y]a^{-1} = axyx^{-1}y^{-1}a^{-1} = (axa^{-1})(aya^{-1})(ax^{-1}a^{-1})(ay^{-1}a^{-1}) = [axa^{-1}, aya^{-1}].$$

This implies that $G' \lhd G$. Finally, if we consider cosets in G/G', $[a], [b] \in G/G'$, then we may observe that

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

where the outer brackets symbolize the coset and the inner brackets symbolize the commutator. Also $[1]$ denotes the neutral element of G/G'.

To end the proof, let's see that if $H \lhd G$ such that G/H is abelian, then $G' \subseteq H$. Indeed, if $a, b \in G$, then $[a][b] = [b][a]$, where the brackets denote cosets in G/H. This implies that $b^{-1}a^{-1}ba \in H$ and as every commutator has this form, we have $G' \subset H$. □

This tells us that G' is the smallest normal subgroup of G such that G/G' is abelian. Let's see some examples of the usability of G'.

Remark 1.4.3. *A group G is abelian if and only if $G' = \{1\}$.*

Remark 1.4.4. *If G is a simple group, then $G' = \{1\}$ or $G' = G$.*

Proof. Suppose G is simple. Then it has no proper non-trivial normal subgroups except from G and $\{1\}$. □

An example of this may be the group A_n. We have that $A_n' = A_n$, for all $n \geq 5$. A group G with $G' = G$ is called *perfect*. Thus, every non-abelian simple group is perfect.

Example 1.4.5. Let's study what is S_n'. Since $S_n/A_n \cong \mathbb{Z}_2$ is abelian, the proposition we just saw tells us that $S_n' \subseteq A_n$. For the reverse inclusion, first note that $S_n' \cap A_n \lhd A_n$; hence, if $n \geq 5$, the simplicity of A_n implies that this intersection is trivial or A_n. But $S_n' \cap A_n \neq \{(1)\}$, so $S_n' \cap A_n = A_n$ and $A_n \subseteq S_n'$. This proves that $S_n' = A_n$ for all $n \geq 5$. A further exercise shows that the equality also holds for $n = 2, 3, 4$.

Definition 1.4.6 (Derived series)**.** The *derived series* of a group G is

$$G = G^{(0)} \supseteq G^{(1)} \supseteq G^{(2)} \supseteq \cdots \supseteq G^{(i)} \supseteq G^{(i+1)} \supseteq \cdots,$$

where $G^{(0)} = G$, $G^{(1)} = G'$ and $G^{(i+1)} = (G^{(i)})' = [G^{(i)}, G^{(i)}]$ for all $i \geq 0$.

Remark 1.4.7. *The derived series is a normal series.*

Proposition 1.4.8. *Let G be a finite group. Then the following are equivalent*

(1) G is solvable.

(2) For some n, $G^{(n)} = \{1\}$.

Proof. Assume that G is solvable, so there is a normal series

$$G = G_0 \supseteq G_1 \supseteq \cdots \supseteq G_n = \{1\}$$

whose factor groups G_i/G_{i+1} are abelian. We show, by induction on $i \geq 0$, that $G^{(i)} \subseteq G_i$. Since $G^{(0)} = G = G_0$, first step is already covered. For the inductive step, as G_i/G_{i+1} is abelian, the proposition (1.4.2) gives $G_i' \subseteq G_{i+1}$. On the other hand, the inductive hypothesis gives $G^{(i)} \subseteq G_i$, which implies that

$$G^{(i+1)} = (G^{(i)})' \subseteq (G_i)' \subseteq G_{i+1}$$

In particular, $G^{(n)} \subseteq G_n = \{1\}$ which is what we wished to show.

Conversely, if $G^{(n)} = \{1\}$, then the derived series is a normal series with abelian factor groups and then it gives what we want. □

This characterization of solvable groups allows us to re-define the concept. In fact, some algebra books such as [4], define "solvable group" as a group G whose derived series eventually ends in $\{1\}$, i.e., for some $n \geq 0$, $G^{(n)} = \{1\}$. This allows us to extend the definition of "solvable" to groups of any order, even infinite. With the previous definition, we only could define a solvable group if it was finite, because if G was not finite, one could not assure that the factor groups were abelian finite.

To end this section it is proposed to re-prove all the characterizations of solvable groups stated in theorem 1.1.7, but with this definition.

Theorem 1.4.9. *(1) Every quotient of a solvable group is itself solvable.*

(2) Every subgroup of a solvable group is itself solvable.

(3) Given G a group, if $H \lhd G$ and G/H are both solvable groups, then G is also solvable.

Proof. (1) If $f : G \to K$ is surjective, then $f(G^{(i)}) = f(G)^{(i)}$ for all i[7]. Therefore, $G^{(n)} = 1$ implies that $1 = f(G^{(n)}) = f(G)^{(n)}$, so that taking $f = \pi : G \to G/H$ (where H is some normal subgroup), we get that G/H is solvable.

[7] To prove this, we just need to prove that $f(G') = f(G)'$. This is true as a generator of G' is $xyx^{-1}y^{-1}$, for $x, y \in G$. Then $f(xyx^{-1}y^{-1}) = f(x)f(y)f(x)^{-1}f(y)^{-1}$ which is a generator of $f(G)'$. The other inclusion is the same backwards.

(2) If H is a subgroup of G, then $H^{(i)}$ is a subgroup of $G^{(i)}$ for all i; hence $G^{(n)} = 1$ implies $H^{(n)} = 1$ for some n.

(3) Let $\pi : G \to G/H$ be the natural map. Then $(G/H)^{(n)} = 1$ for some n by hypothesis, which implies that $\pi(G^{(n)}) = 1$ and hence $G^{(n)}$ is a subgroup of H. As H is also solvable, there exists m (not necessarily equal to n) such that $H^{(m)} = 1$ and then $(G^{(n)})^{(m)} = 1$. Finally one proves by induction that $G^{(n+m)}$ is a subgroup of $(G^{(n)})^{(m)}$ for all m, and so G is solvable.

□

Then with this final definition of solvable groups, the abelian groups are always solvable, regardless they are finite or infinite.

Chapter 2

Representation theory

Every mathematician that studies Group Theory comes across Representation Theory at the very beginning. Thus, this theory is used to give proofs of well-known theorems although they are not directly involved in this area. However they are shorter and more intuitive

Thus, Burnside's Theorem, is proven using Character Theory, which will be discussed in the next chapter, but it is very related and based on Representation Theory. I will give the basic definitions, as well as prove important theorems regarding them.

As we may see, Representation Theory is intimately related to Module Theory. The first chapter will contain basic definitions of our main object: a linear representation. Moreover, I will discuss the group ring, i.e., a ring described by a finite group, and the modules over them. Then we will see the relation between modules over group rings and linear representations, and it will seen that in fact we will be able to change from one theory to another when we need to. The last two sections will aim to prove two important theorems. The first one is Maschke's Theorem which is a purely Representation Theory theorem and the second one is a more general Module Theory theorem, called Wedderburn's Theorem. This last theorem will require a bit more background on Module Theory that may be essential for it.

2.1 Linear representations

To start, let V be a vector space over the field algebraically closed. Unless specifically said, we will fix V a finite dimensional vector space over an algebraically closed field k. Now, remember that we called $\mathbf{GL}(V)$ the *linear group*[1] to the group of isomorphisms of V onto itself. An element $a \in \mathbf{GL}(V)$ is, by definition, a linear mapping of V into V which has an inverse, denoted a^{-1} which is also linear. When we can establish a finite basis $\{e_1, \dots, e_n\}$ of V, each linear map $a : V \to V$ is defined obtaining the images of the basis, in terms of this basis. All the images can be expressed in an $n \times n$ matrix $(a_{ij})_{ij}$,

[1] The GL stands for its name in French "Groupe Linéaire", and it is also denoted LG for its name in English. I will always write GL as I'm also used to this for it is the same in Spanish and Catalan.

where a_{ij} are the coefficients in the field k given by

$$a(e_j) = \sum_{i=1}^{n} a_{ij} e_i$$

This means that every element $a \in \mathbf{GL}(V)$ has an $n \times n$ matrix associated, which we can call $M_a := (a_{ij})_{ij}$ and if $v \in V$, the image of v by the linear map $a : V \to V$ is given by the product $a(v) = M_a v$.

Saying that a is an isomorphism from V onto V is the same as saying that the matrix representing $a : V \to V$ has determinant $\det(a) \neq 0$. Hence, the group $\mathbf{GL}(V)$ is identifiable with the group of *invertible square matrices of order n*.

Definition 2.1.1 (Linear representation). Let $(G, \cdot)$ be a group with neutral element 1, and V be a $\mathbb{C}$-linear space. A *linear representation* of G in V is a homomorphism $\rho : G \to \mathbf{GL}(V)$.

In other words, we associate each element $x \in G$ an element $\rho(x) \in \mathbf{GL}(V)$ in such way that we have the equality

$$\rho(xy) = \rho(x)\rho(y) \qquad \forall x, y \in G$$

and also $\rho(1) = 1 \in \mathbf{GL}(V)$ and $\rho(x^{-1}) = \rho(x)^{-1}$. To denote $\rho(x)$ we will often write ρ_x, for any $x \in G$. When this ρ is given, we say that V is a *representation space* of G (or even simply, by abuse of language, a *representation* of G). The fact that we fixed a vector space of finite dimension is not a severe restriction: for most applications, one is interested in dealing with a *finite number of elements x_i* of V, and can always find a *sub-representation* of V (we will define later this concept) of finite dimension, which contain the x_i; just take the vector subspace generated by the images $\rho_s(x_i)$ of the x_i, for $s \in G$.

Suppose then that V has a finite dimension, and let n be its dimension. We say also that n is the *degree of the representation* under consideration. Let $\{e_1, \ldots, e_n\}$ be a basis of V, and if $s \in G$ suppose given the representation $\rho_s \in \mathbf{GL}(V)$. As we said early, this is a linear map with a matrix associated to it with respect the given basis, which we will call R_s. Now, by the properties we discussed before, we have

$$\det(R_s) \neq 0, \qquad R_{st} = R_s \cdot R_t$$

for any $s, t \in G$ with the product $\cdot$ here being the usual matrix product. If we denote $r_{ij}(s)$ to the (i, j)-th element of R_{ij} the second formula becomes

$$r_{ik}(st) = \sum_{j=1}^{n} r_{ij}(s) \cdot r_{jk}(t)$$

Conversely, given matrices $R_s = (r_{ij}(s))_{ij}$ satisfying the preceding identities, there is a corresponding linear representation ρ of G in V: this is what it means to give a representation "in matrix form".

2.2 Group rings

Before going further in the study of the invariants of such representations, and giving some examples of them, we should introduce and emphasize before the concept of "group ring" that will be particularly useful and will bring us a more general way of understanding representation theory.

Suppose we have a commutative ring R with unity and $1 \neq 0$. Suppose we also have a finite group, say $G = \{g_1, \dots, g_n\}$, with the operation expressed multiplicative. Then we define the group ring as follows.

Definition 2.2.1 (Group ring)**.** We define $R[G]$ as the set of all formal sums of the form

$$a_1 g_1 + \cdots + a_n g_n, \qquad a_i \in R$$

Now, if e is the neutral element (or identity) of G we shall write $ae = a$, for any $a \in R$. In the same way, we write $1g = g$ for any $g \in G$. We define the addition of elements of $R[G]$ as follows

$$\sum_{i=1}^{n} a_i g_i + \sum_{i=1}^{n} b_i g_i = \sum_{i=1}^{n} (a_i + b_i) g_i.$$

It can be understood as "componentwise". To define the multiplication, first for $a, b \in R$ we define $(ag_i)(bg_j) = (ab)g_k$, where the product ab is performed in R and $g_i g_j = g_k$ in the group G. Then we extend this definition with the "distributive" laws:

$$\left(\sum_{i=1}^{n} a_i g_i\right)\left(\sum_{i=1}^{n} b_i g_i\right) = \sum_{k=1}^{n}\left(\sum_{\substack{i,j \\ g_i g_j = g_k}} a_i b_j\right) g_k$$

Example 2.2.2. Let's see a few examples to understand the concept more clear.

(1) Take the group $G = D_{2\cdot 4}$ (the dihedral group of order 8) often denoted by $D_{2\cdot 4} = \langle \sigma, \rho \ : \ \sigma^2 = \rho^4 = \text{id},\ \sigma\rho^{-1} = \rho\sigma \rangle$. Take as the ring $R = \mathbb{Z}$ and then we study the group ring $\mathbb{Z}[D_{2\cdot 4}]$ often called "integral ring of $D_{2\cdot 4}$". Some elements of this ring might be

$$\alpha = \rho + \rho^2 - 2\sigma, \qquad \beta = -3\rho^2 + \rho\sigma$$

and we can compute the sum and the product and we get

$$\alpha + \beta = \cdots = \rho - 2\rho^2 - 2\sigma + \sigma\rho$$

and

$$\alpha\beta = (\rho + \rho^2 - 2\sigma)(-3\rho^2 + \rho\sigma) = \cdots = -3 - 5\rho^3 + 7\rho^2\sigma + \rho^3\sigma$$

(2) Take now the ring $\mathbb{Q}$ and the group S_3, the permutation group. The elements $r = 5(1,2) - 7(1,2,3)$ and $s = -4(1,2,3) + 12(1,3,2)$ are some examples of members of $\mathbb{Q}[S_3]$. We can compute the sum and product of them to see how they interact.

$$r + s = 5(1,2) - 11(1,2,3) + 12(1,3,2)$$

$$rs = -20(2,3) + 28(1,3,2) + 60(1,3) - 84$$

(3) Take now $G = \langle g \rangle$ a cyclic group of order $n \in \mathbb{Z}^+$ and $R = k$ to be a field. Then the elements of $k[G]$ are of the form

$$\sum_{i=0}^{n-1} \alpha_i g^i$$

where $\alpha_i \in k$. The map $k[x] \to k[\langle g \rangle]$ which sends x^k to g^k for all $k \geq 0$ extends by k-linearity to a surjective ring homomorphism with Ker equal to the ideal generated by $x^n - 1$ (as g has order n). Then, by the First Isomorphism Theorem we have an isomorphism of k-algebras (i.e. a ring isomorphism which is k-linear)

$$k[\langle g \rangle] \cong k[x]/(x^n - 1)$$

Let now $r = 1 + g + g^2 + \cdots + g^{n-1}$, so $r(1-g) = 0$ and thus $k[\langle g \rangle]$ contains zero divisors (supposing $n > 1$). This is a particular case of what I will discuss in a brief moment.

To end the section, let's discuss a few more properties of these group rings.

Remark 2.2.3. *The ring R appears in $R[G]$ as the elements of R multiplied by the neutral element e of G are in R. Note that the definition of the addition and multiplication in $R[G]$ restricted to these elements is just the addition and multiplication in R).*

The group G also appears in $R[G]$, as the element g_i appears in $R[G]$ as $1g_i$, for all $i = 1, \dots, n$. In particular, each element of G has a multiplicative inverse in the ring $R[G]$ (namely, its inverse in G). This means that G is a subgroup of the group of units of $R[G]$.

Remark 2.2.4. *If $|G| > 1$, $R[G]$ always has zero divisors, as for example, take g an element of order $m > 1$. Then*

$$(1-g)(1+g+\cdots+g^{m-1}) = 1 - g^m = 1 - 1 = 0$$

so $1 - g$ is a zero divisor as, by the definition of $R[G]$ we gave, neither of the formal sums in the above product is zero.

Remark 2.2.5. *If S is a subring of R, then $S[G]$ is a subring of $R[G]$. For instance, $\mathbb{Z}[G]$ (the integral group ring of G) is a subring of $\mathbb{Q}[G]$ (the rational group ring of G).*

On the other hand, if H is a subgroup of G, then $R[H]$ is a subring of $R[G]$.

2.3 Correspondence between representations of G and $k[G]$-modules

Now we will study the representations taking into account what we have discussed of group rings. Take any field k and then, in particular, k is a ring. Take now a finite group $G = \{g_1, \dots, g_n\}$ and take the group ring $k[G]$. Then, as we discussed in the previous section, the group G appears in $k[G]$ (identifying g_i with $1g_i$) and the field k appears in $k[G]$ (identifying β with βg_1, where g_1 is the identity of G). Under these identifications

$$\beta \left(\sum_{i=1}^{n} \alpha_i g_i \right) = \sum_{i=1}^{n} (\beta \alpha_i) g_i, \qquad \forall \beta \in k.$$

In this way, $k[G]$ *is a vector space over k with the elements of G as a basis*. In particular, $k[G]$ is a vector space over k of dimension equal to $|G|$. The elements of k commute with all elements of $k[G]$, i.e. k is contained in the *center* of $k[G]$. When we wish to emphasize the latter two properties we shall say that $k[G]$ is an k-algebra[2].

In this chapter we will discuss then the relation between the representations of G on the k-vector space V and the $k[G]$-modules, but first let's recall what is a module over a ring.

Definition 2.3.1 (Module over a ring)**.** Let R be a ring (not necessarily commutative nor with 1). A *left R-module* or a *left module over R* is a set M together with

(1) a binary operation $+$ on M under which M is an abelian group, and

(2) an action of R on M (that is, a map $R \times M \to M$) denoted by rm, for all $r \in R$ and for all $m \in M$ that satisfies

   (a) $(r+s)m = rm + sm$, for all $r, s \in R$, $m \in M$,

   (b) $(rs)m = r(sm)$, for all $r, s \in R$, $m \in M$, and

   (c) $r(m+n) = rm + rn$, for all $r \in R$, $m, n \in M$.

   (d) If the ring R has a 1 we impose the additional axiom: $1m = m$, for all $m \in M$.

This definition is given in [2]. The descriptor "left" in the definition indicates that the ring elements appear on the left; "right" R-modules can be defined analogously. If the ring R is commutative and M is a left R-module, we can make M into a right R-module by defining $mr = rm$ for $m \in M$ and $r \in R$. But if R is not commutative, axiom 2(b) generally will not hold with this definition of right R-module, so not every left R-module is also a right R-module. Unless explicitly mentioned otherwise the term "module" will always mean "left module". Finally, if the ring is a field k, the axioms for an R-module are precisely the same as those for a vector space over k, so that *modules over a field k and vector spaces over k are the same*.

We now go back to representations and let's see what does this modules have to do with them. Take a finite group G and a field k and take $\sigma : G \to \mathbf{GL}(V)$ to be a linear representation over an k-vector space V. Write $G = \{g_1, \ldots, g_n\}$ and then, for some $i \in \{1, \ldots, n\}$, $\sigma_{g_i} := \sigma(g_i)$ is a linear transformation from V into itself: $\sigma_{g_i} : V \to V$. Now we will build V as an $k[G]$-module. Define the following action of any element of $k[G]$ on an element $v \in V$:

$$\left(\sum_{i=1}^{n} \alpha_i g_i \right) v := \sum_{i=1}^{n} \alpha_i \sigma_{g_i}(v)$$

Proposition 2.3.2. *This action defines V (the vector space representation of G) as an $k[G]$-module.*

[2] In general, an k-algebra is a ring R which contains k in its center, so R is both a ring and an k-vector space.

Proof. It is clear that satisfies the first condition of the definition of module over a ring (2.3.1). Now let's see the second. We have that $\alpha_i \in k$ for the definition of group ring (2.2.1), thus $\alpha_i \sigma_{g_i}(v)$ is an element of V (as it is a vector space over k) hence all the sum is. Then it is a well defined action. Now let's see if it satisfies the conditions.

(a) It is easy to see that this one holds:

$$\left(\sum_{i=1}^{n} \alpha_i g_i + \sum_{i=1}^{n} \beta_i g_i\right) v = \left(\sum_{i=1}^{n} (\alpha_i + \beta_i) g_i\right) v = \sum_{i=1}^{n} (\alpha_i + \beta_i) \sigma_{g_i}(v) =$$

$$= \sum_{i=1}^{n} (\alpha_i \sigma_{g_i}(v) + \beta_i \sigma_{g_i}(v)) = \left(\sum_{i=1}^{n} \alpha_i g_i\right) v + \left(\sum_{i=1}^{n} \beta_i g_i\right) v$$

The first equality holds for the definition of the addition in $k[G]$, the second one is the definition of the action and then it is just play with the properties of sums.

(b) We verify a special case of this axiom which shows precisely where the fact that σ is a group homomorphism is needed:

$$\begin{aligned} (g_i g_j) v &= \sigma_{g_i g_j}(v) = \\ &= (\sigma_{g_i} \circ \sigma_{g_j})(v) = \\ &= \sigma_{g_i}(\sigma_{g_j}(v)) = \\ &= g_i(g_j v) \end{aligned}$$

and then the argument extends to the general case of an arbitrary element of $k[G]$. The first and last equality hold for the definition of the action, while the second one since σ is a group homomorphism.

(c) This case is very easy to prove, and it is similar to case (a).

(d) This case is trivial.

Hence, V is an $k[G]$ module with this action we defined. □

Now we will see a result that states the correspondence we want: the $k[G]$-modules are in bijective correspondence with the representations of G over an k-vector space.

Proposition 2.3.3. *There is a bijection between $k[G]$-modules and pairs (V, σ), where V is an k-vector space and σ a representation of G over V, i.e.*

$$\{V \textit{ an } k[G]\textit{-module}\} \longleftrightarrow \left\{ \begin{array}{c} V \textit{ a vector space over } k \\ \textit{and} \\ \sigma : G \to \mathbf{GL}(V) \textit{ a representation} \end{array} \right\}$$

Giving a representation $\sigma : G \to \mathbf{GL}(V)$ on a vector space V over k is therefore equivalent to giving an $k[G]$-module V. Under this correspondence we shall say that the module V affords the representation φ of G.

Proof. This proof comes from a discussion made in [2] that I transform in a more formal proof.

Note that k is a subring of $k[G]$ (we already discussed that in the general case at the beginning of the previous section) and the action of the field element α on a vector is the same as the action of the ring element $\alpha 1$ on a vector, i.e- the $k[G]$-module action extends the k action on V. This proves the arrow from right to left.

Conversely, suppose now that we are given an $k[G]$-module V. We obtain an associated vector space over k and representation of G as follows. We already saw that since V is an $k[G]$-module, it is also an k-vector space, so there we got it. Also, for $g \in G$, we obtain a map from V to V, denoted by σ_g, defined by

$$\sigma_g(v) = gv \qquad \forall v \in V$$

where gv is the given action of the ring element g on the element v of V. Since the elements of k commute with each $g \in G$, it follows by the axioms for a module that for all $v, w \in V$ and all $\alpha, \beta \in k$ we have

$$\begin{aligned}\sigma_g(\alpha v + \beta w) &= g(\alpha v + \beta w) = \\ &= g(\alpha v) + g(\beta w) = \\ &= \alpha(gv) + \beta(gw) = \\ &= \alpha\sigma_g(v) + \beta\sigma_g(w)\end{aligned}$$

that is, for each $g \in G$, σ_g is a linear transformation. Furthermore it follows from axiom 2(b) of a module that

$$\sigma_{g_i g_j}(v) = (\sigma_{g_i} \circ \sigma_{g_j})(v)$$

which proves that σ is a group homomorphism (in particular $\sigma_{g^{-1}} = \sigma_g^{-1}$) so every element of G maps to a non-singular linear transformation, i.e. a representation $\sigma : G \to \mathbf{GL}(V)$. □

Remark 2.3.4. *If V is an $k[G]$-module and its correspondent representation is φ, then a subspace U of V is called G-invariant or G-stable if $g \cdot u \in U$ for all $g \in G$ and $u \in U$ (it is the applied definition of stability in group action context). By the correspondence we have $\varphi_g(u) \in U$ for all $g \in G$ and all $u \in U$. It follows easily that the $k[G]$-submodules of V are precisely the G-stable subspaces of V.*

Example 2.3.5. Let's see some examples of representations.

(1) Let V be a 1-dimensional vector space over k and make V into an $k[G]$-module by letting $gv = v$ for all $g \in G$ and $v \in V$. This module affords the representation $\varphi : G \to \mathbf{GL}(V)$ defined by $\varphi(g) = \text{id}$ for all $g \in G$, where id represents the identity transformation. The corresponding matrix with respect to any basis of V is the identity matrix. We shall henceforth refer to this as the *trivial representation* of G. It has degree 1.

(2) Let g be the order of G, i.e. $|G| = g$. Let V be a vector space of dimension g, with basis $(e_t)_{t\in G}$, indexed by elements t of G. For $s \in G$, let ρ_s be the linear map of V into V which sends e_t into e_{st}. This defines a linear representation which is called the *regular representation* of G. Its degree is equal to the order of G.

Let's try to imagine how would the matrix representation of $\rho(s)$ be. As $\rho(s) : V \to V$ sends e_g to e_{sg} the only thing that does is to permute the columns of the identity matrix, as $\forall g \in G$, $sg \in G$.

(3) Let $n \in \mathbb{Z}_{>0}$ and $G = S_n$ the permutation group. Let V be an k-vector space of dimension n with basis $\{e_1, \dots, e_n\}$. Then we define $\forall \sigma \in S_n$, the action of S_n over V as

$$\sigma \cdot e_i := e_{\sigma(i)}, \quad \forall i = 1, \dots, n$$

and hence we have a representation $\varphi : S_n \hookrightarrow \mathbf{GL}(V)$ that is a representation of S_n into an $k[S_n]$-module. Then, call $j = \sigma(i)$ and we get that $\varphi(\sigma)$ is a linear transformation

$$\begin{aligned} \varphi(\sigma) = \varphi_\sigma : k[G] &\longrightarrow k[G] \\ e_i &\longmapsto e_j \end{aligned}$$

and hence each column of the matrix representation of φ_σ is full of zeroes except in position j that there is a 1.

Let's examine a practical example. Suppose $n = 3$ and take $\sigma = (1, 2)$ and $\tau = (1, 2, 3)$. Then the matrices of φ_σ and φ_τ would be

$$\varphi_\sigma = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \varphi_\tau = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

(4) Take now $G = D_{2\cdot n} = \langle \sigma, \rho \ : \ \sigma^2 = \rho^n = \text{id}, \ \rho\sigma = \sigma\rho^{-1} \rangle$. Now, if R and S are matrices that satisfy $R^n = S^2 = \text{id}$ and $RS = SR^{-1}$, then the map $\rho \mapsto R$ and $\sigma \mapsto S$ extends uniquely to a homomorphism from D_{2n} to the matrix group generated by R and S, hence gives a representation of D_{2n}.

If we recall what is the dihedral, ρ represents a rotation of $2\pi/n$ degrees and σ a symmetry, of a n-gon in the plane. Then R is a rotation matrix and S is a symmetry matrix, both in the plane, and these matrices are well-known:

$$R = \begin{pmatrix} \cos \frac{2\pi}{n} & -\sin \frac{2\pi}{n} \\ \sin \frac{2\pi}{n} & \cos \frac{2\pi}{n} \end{pmatrix}$$

Hence the maps $\rho \mapsto R$ and $\sigma \mapsto S$ extend uniquely to a degree 2 representation of D_{2n} into $\mathbf{GL}(V)$ where V is a 2-degree $\mathbb{R}$-vector space.

This is, in fact, a very trivial example, because it is equivalent to giving the matrix representations of the elements of $D_{2\cdot n}$ which are exactly the isometries.

Definition 2.3.6 (Equivalent representations)**.** Two representations of G are *equivalent* or *similar* if the $k[G]$-modules that correspond to them are isomorphic modules. Representations which are not equivalent are called *inequivalent*.

We stop a moment to recall what does it mean to have isomorphic modules. If R is any ring, then a R-morphism of modules M_1 and M_2 is a map $f : M_1 \mapsto M_2$ such that $f(x+y) = f(x) + f(y)$ and $f(\lambda x) = \lambda f(x)$ for all $x, y \in M_1$ and all $\lambda \in R$. In our case, the ring is $k[G]$ and the modules are the ones corresponding to each representation. For example, suppose that we have $\varphi : G \to \mathbf{GL}(V)$ and $\psi : G \to \mathbf{GL}(W)$ two representations of a group G. Then V is a $k[G]$-module, as well as W, and φ and ψ will be equivalent representations if $V \cong W$.

In this situation, let $T : V \to W$ be a $k[G]$-module isomorphism between them. Since T is, in particular, a k-isomorphism, T is a vector space isomorphism, so V and W must have the same dimension as k-vector spaces. Furthermore, for all $g \in G$, $v \in V$, we have $T(g\cdot v) = g\cdot (T(v))$, since T is an isomorphism of $k[G]$-modules. By definition of the action of ring elements, this means that $T(\varphi_g(v)) = \psi_g(T(v))$, that is

$$T \circ \varphi_g = \psi_g \circ T \qquad \forall g \in G$$

then we have the following commutative diagram

$$\begin{tikzcd} V \arrow[r, "\varphi_g"] \arrow[d, "T"'] & V \arrow[d, "T"] \\ W \arrow[r, "\psi_g"'] & W \end{tikzcd}$$

In particular, if we identify V and W as vector spaces, then two representations φ and ψ on the vector space V are equivalent if, and only if, there is some $T \in \mathbf{GL}(V)$ such that $T \circ \varphi_g \circ T^{-1} = \psi_g$, for all $g \in G$. This T is a *simultaneous* change of basis for all φ_g, $g \in G$.

All this translated to matrix terminology would be saying that two representations φ and ψ are equivalent if there is a fixed invertible matrix P such that

$$PM_{\varphi_g}P^{-1} = M_{\psi_g} \qquad \forall g \in G$$

The linear transformation T or the matrix P in each case is said to *intertwine* the representations φ and ψ (it gives the "rule" for changing φ into ψ).

Regarding this definition and these observations, we can now see some properties of modules and submodules to transfer them to representations. The following definitions are in general affecting R-modules for any ring R, and further we will aply them to our particular case.

Definition 2.3.7. Let R be a ring and M a non-zero R-module.

(1) The module M is said to be *irreducible* if its only submodules are 0 and M. Otherwise M is called *reducible*.

(2) The module M is said to be *indecomposable* if M cannot be written as $M_1 \oplus M_2$ for any nonzero submodules M_1 and M_2. Otherwise M is called *decomposable*.

(3) The module M is called *completely reducible* if it is a direct sum of irreducible submodules.

(4) If M is a completely reducible R-module, any direct summand of M is called *constituent* of M (i.e. N is a constituent of M if there is a submodule N' of M such that $M = N \oplus N'$).

An irreducible module is, by definition, both indecomposable and completely reducible. We shall shortly give examples of indecomposable modules that are not irreducible. Now I translate these definitions to representations as follows.

Definition 2.3.8. A representation is called *irreducible, reducible, indecomposable, decomposable* or *completely reducible* according to whether the corresponding $k[G]$-module has the corresponding property.

If $R = k[G]$, an irreducible $k[G]$-module V is a nonzero k-vector space with no nontrivial proper G-invariant subspaces. For example, if $\dim_k V = 1$, then V is necessarily irreducible (its only subspaces are 0 and V).

Suppose V is a finite dimensional $k[G]$-module and V is reducible. Let U be a G-stable subspace. Take a basis of U and enlarge it to obtain a basis of V. Then, for each $g \in G$, the matrix of φ_g, call it M_{φ_g}, of g acting on V with respect to this basis is of the form

$$M_{\varphi_g} = \begin{pmatrix} \varphi_1(g) & \psi(g) \\ 0 & \varphi_2(g) \end{pmatrix}$$

where $\varphi_1 = \varphi_{|U}$ with respect to the chosen basis of U and φ_2 is the representation of G on V/U, and ψ is not necessarily a homomorphism (not need to be a square matrix). I denote $\varphi_i(g)$ instead of $(\varphi_i)_g$ for cleaner notation. So, reducible representations are those with a corresponding matrix representation whose matrices are in block upper triangular form.

Furthermore, assume that the $k[G]$-module V is decomposable, and write $V = U \oplus U'$. Take for basis of V the union of a basis of U and a basis of U'. With this choice of basis, the matrix for each $g \in G$ is of the form

$$\varphi_g = \begin{pmatrix} \varphi_1(g) & 0 \\ 0 & \varphi_2(g) \end{pmatrix}$$

i.e. $\psi_g = 0$ for all $g \in G$. Thus, decomposable representations are those with corresponding matrix representation whose matrices are in block diagonal form.

2.4 Maschke's Theorem

In this section there is proved the first fundamental result on Representation Theory of Finite Groups, known as Maschke's Theorem[3].

Theorem 2.4.1 (Maschke's Theorem)**.** *Let G be a finite group and let k be a field whose characteristic does not divide $|G|$. If V is a $k[G]$-module and U is any submodule of V, then V has a submodule W such that $V = U \oplus W$. In other words, every submodule is a direct summand.*

The key to prove Maschke's Theorem relays on "producing" a $k[G]$-module epimorphism (a $k[G]$-module homomorphism that is onto) $\pi : V \to U$ which is a projection onto U, i.e. that satisfies the following two properties:

(i) $\pi(u) = u$ for all $u \in U$,

(ii) $\pi(\pi(v)) = \pi(v)$ for all $v \in V$, i.e. $\pi^2 = \pi$.

Once we have proven the existence of such projection, we will have easily the proof for Maschke's Theorem.

Proposition 2.4.2. *If k is any field and G a finite group such that the characteristic of k does not divide $|G|$, then consider V any $k[G]$-module and U any submodule of V. Then, there exists a $k[G]$-module homomorphism*

$$\pi : V \to U$$

such that satisfies the following two properties:

(i) $\pi(u) = u$ for all $u \in U$

(ii) $\pi(\pi(v)) = \pi(v)$ for all $v \in V$, so $\pi^2 = \pi$.

These two properties indicate that π can be understood as a projection map from V to U.

Proof. Let's try to see that this π exists. Thinking U as a subspace of V, we can take a complement W_0, i.e. $V = U \oplus W_0$ as subspaces. This can be done in the following way: take $\mathcal{B}_1 = \{u_1, \ldots, u_r\}$ a basis of U and take a basis of V extending it, i.e. $\mathcal{B} = \{u_1, \ldots, u_r, v_{r+1}, \ldots, v_n\}$. Then, $\mathcal{B} \setminus \mathcal{B}_1 = \{v_{r+1}, \ldots, v_n\}$ is a basis for W_0.

Now, the problem is that W_0 need not be a $k[G]$-module. We are going to build a $k[G]$-module about it. Take $v \in V$ as $v = u + w_0$ with $u \in U$ and $w_0 \in W_0$ and define the following application between vector spaces over k:

$$\pi_0 : V \to U$$

that takes $v = u + w_0$ and sends it to $\pi_0(v) = u$. This can easily seen as a projection from V onto U as vector spaces. Now, take for all $g \in G$ the following map

$$g\pi_0 g^{-1} : V \to U$$

[3] Heinrich Mascke (Breslau, German Empire (now Wroclaw, Poland), 1853 - Chicago, USA, 1908) was a german matematician principally known for this theorem.

which sends $v \in V$ to $g\pi_0 g^{-1}(v) = g\cdot \pi_0(g^{-1}\cdot v)$. Here we are extending ourselves to $k[G]$ because the "$\cdot$" denotes the action of elements in $k[G]$. Since we have that π_0 maps V into U and U is stable under the action of g, we have that $g\pi_0 g^{-1}$ maps V into U. Also, g and g^{-1} both act as k-linear transformations so $g\pi_0 g^{-1}$ is also a k-linear transformations. Furthermore, if u is in the G-stable subspace U, then so is $g^{-1}u$, and by definition of π_0 we have $\pi_0(g^{-1}u) = g^{-1}u$, hence $g\pi_0 g^{-1}(u) = gg^{-1}u = u$ for all $u \in U$. This shows that $g\pi_0 g^{-1}$ is also a projection of V onto U.

Consider now $n = |G|$ and see n as an element of k, i.e. $n = 1 + 1 + \cdots + 1$ n times. By hypothesis, the characteristic of k does not divide n, so n can be inverted. Define now

$$\pi := \frac{1}{n}\sum_{g\in G} g\pi_0 g^{-1}$$

We will prove now that π is the one that satisfies all the properties. Since π is a scalar multiple of a sum of linear transformations from V to U, it is himself a linear transformation from V to U. Furthermore, each term in the sum defining π restricts to the identity map on the subspace U and so $\pi|_U$ is $1/n$ times the sum of n copies of the identity. These observations prove that $\pi : V \to U$ is a linear transformation that satisfies (i) and (ii).

It remains to show that π is a $k[G]$-module homomorphism (i.e. that is $k[G]$-linear). It suffices to prove that for all $h \in G$, $\pi(hv) = h\pi(v)$ for $v \in V$. In this case

$$\pi(hv) = \frac{1}{n}\sum_{g\in G} g\pi_0(g^{-1}hv) = \frac{1}{n}\sum_{g\in G} h(h^{-1}g)\pi_0((g^{-1}h)v) =$$

$$= \frac{1}{n}\sum_{\substack{k=h^{-1}g \\ g\in G}} h(k\pi_0(k^{-1}v)) = h\pi(v)$$

as g runs over all elements of G, so does $k = h^{-1}g$ and the module element h may be brought outside the summation by the distributive law in modules). This establishes the existence of the $k[G]$-module projection π and so completes the proof. □

Now we can give the proof to Maschke's Theorem 2.4.1 as follows.

Proof of Maschke's Theorem 2.4.1. Suppose $\pi : V \to U$ with the above properties. Then define $W = \operatorname{Ker}\pi$. Since π is a module homomorphism, W is a submodule. Now it only remains to see that W is a direct sum complement of U, i.e. that $U \cap W = 0$ and $U + W = V$.

- If $v \in U \cap W$, then by the first property of π we have $v = \pi(v)$ but by definition of $W = \operatorname{Ker}\pi$ it must be $\pi(v) = 0$. Hence $U \cap W = 0$.

- Let v be an arbitrary element of V and write $v = \pi(v) + (v - \pi(v))$. By definition, $\pi(v) \in U$ and it remains to show that $v - \pi(v) \in W$. By the second property of π, and the fact that π is a $k[G]$-module homomorphism, we have

$$\pi(v - \pi(v)) = \pi(v) - \pi(\pi(v)) = \pi(v) - \pi(v) = 0$$

which shows that $v - \pi(v) \in \operatorname{Ker} \pi = W$ as we wanted. This shows that any v is expressed as a sum of an element of U and an element of W, hence $V = U + W$

These two points show that $V = U \oplus W$ as we wanted. □

A direct result of this theorem is the following corollary.

Corollary 2.4.3. *Every representation of a finite group G over a field k whose characteristic does not divide the order of G is a direct sum of irreducible representations.*

Proof. Given the definitions of irreducible representations 2.3.8 and the correspondence theorem 2.3.3 the proof is straight forward as the Maschke's Theorem states that if V is the correspondent $k[G]$-module of the representation $\varphi : G \to \mathbf{GL}(V)$, then any submodule U is a direct summand, i.e., there exists another submodule W such that $V = U \oplus W$. Now, if U and W are both irreducible as $k[G]$ modules, the correspondent representations will be also irreducible and hence we found an irreducible decomposition of φ. If we suppose that U is not irreducible, we can apply again Maschke's Theorem 2.4.1 and obtain another decomposition $U = U_1 \oplus U_2$ and so on. □

We have seen so far that every representation of a finite group G over a field k whose characteristic does not divide $|G|$ decomposes in direct sum of irreducible representations. Therefore, we are interested now in the study of these irreducible representations. In the next section we will give some results about the structure of these irreducible modules and its correspondent irreducible representations. We will study the structure of $k[G]$ as a ring and we will see that it belongs to a special category of rings: the *semisimple* rings. We will see also that $k[G]$ can also be seen as a $k[G]$-module and it is itself reducible into a direct sum of irreducible modules.

2.5 Wedderburn's Theorem and consequences

In this section I will state the Wedderburn's Theorem which is a general theorem in ring theory that gives the structure of semisimple rings. Unlike Maschke's Theorem, this theorem will characterize a more generic type of rings, which are semisimple rings, and we wil say that $k[G]$ is one of them.

I will state the theorem and after we will discuss some of its consequences and analyze it from the point of view of our $k[G]$-modules. In "Introducción al Álgebra Conmutativa" we have already seen and studied the concepts of *projective* and *injective* modules, which are used in the following results. Hence, I will assume these concepts as known.

The theorem by Wedderburn stated in [2] is as follows:

Theorem 2.5.1 (Wedderburn's). *Let R be a nonzero ring with 1 (not necessarily commutative). Then, the following are equivalent:*

(1) every R-module is projective

(2) every R-module is injective

(3) every R-module is completely reducible

(4) the ring R considered as a left R-module is a direct sum:

$$R = L_1 \oplus L_2 \oplus \cdots \oplus L_n,$$

where each L_i is a simple module (i.e., a simple left ideal) with $L_i = Re_i$, for some $e_i \in R$ with

(i) $e_i e_j = 0$ if $i \neq j$

(ii) $e_i^2 = e_i$ for all i

(iii) $\sum_{i=1}^{n} e_i = 1$

(5) as rings, R is isomorphic to a direct product of matrix rings over division rings[4]*, i.e., $R = R_1 \times R_2 \times \cdots \times R_r$ where R_j is a two-sided ideal of R and R_j is isomorphic to the ring of all $n_j \times n_j$ matrices with entries in a division ring Δ_j, $j = 1, 2, \ldots, r$. The integer r, the integers n_j, and the division rings Δ_j (up to isomorphism) are uniquely determined by R.*

A ring R satisfying any of the (equivalent) properties of the theorem is called *semisimple with minimum condition*. These rings also satisfy the *minimum condition* or the *descending chain condition* (D.C.C.) on left ideals: if $I_1 \supseteq I_2 \cdots$ is a descending chain of left ideals of R, then there is an $N \in \mathbb{Z}_{\geq 0}$ such that $I_k = I_N$ for all $k \geq N$, which explains the use of this term above in the definition.

The rings we are dealing with will all have this minimum condition. For example, group algebras always have this property since in any strictly descending chain of ideals, the vector space dimensions of the ideals (which are k-subspaces of $k[G]$) are strictly decreasing, hence the length of a strictly descending chain is at most the dimension of $k[G]$ (i.e. $|G|$). We shall therefore use the term “semisimple” to mean “semisimple with the minimum condition”. The rings R_i in conclusion (5) of Wedderburn’s Theorem 2.5.1 are called *Wedderburn components* of R and the direct product decomposition of R is called its *Wedderburn decomposition*.

We now apply Wedderburn’s Theorem to our group algebra $k[G]$. First of all, let’s recall that it is mandatory that our characteristic of k does not divide $|G|$, the order of the group. In fact, since we shall be dealing with numerical data in the sections to come it will be convenient to have the characteristic of k equal to 0. Secondly, it will simplify matters if we force all the division rings which will appear in the Wedderburn

[4] A division ring Δ is a ring all whose elements different from zero are invertible, but that does not need to be commutative. A commutative division ring is what we usually call field. Sometimes fields are refereed to as commutative division rings, and sometimes there are references that call fields to the division rings, without the commutativity. In these notes, a division ring will be a ring all whose non-zero elements are invertible. but without being commutative.

There is another well known “Wedderburn Theorem” which states that any division ring being finite is also commutative.

decomposition of $k[G]$ to equal the field k (later on we will prove that imposing the condition that k be algebraically closed is sufficient to ensure this). In the book [2] the authors argue that it is useful for the reader to take $k = \mathbb{C}$.

Recalling the definitions of injective and projective modules, Maschke's Theorem 2.4.1 implies the following corollary.

Corollary 2.5.2. *If G is a finite group and k a field whose characteristic does not divide $|G|$, then the group ring $k[G]$ is a semisimple ring.*

Proof. Maschke's Theorem states that if K is a submodule of a $k[G]$-module M, then there exists L such that $M = K \oplus L$. This is equivalent to say that $k[G]$-modules are injective. Hence, Wedderburn's Theorem gives that $k[G]$ is semisimple. □

In particular, this implies that there is the following decomposition:

$$k[G] \cong R_1 \times \cdots \times R_r$$

where R_i is the ring of $n_i \times n_i$ dimension matrices over some division ring Δ_i.

Our next step will be to see a result that gives, in particular, the structure of the simple $k[G]$-modules.

Proposition 2.5.3. *Let $R = R_1 \times \cdots \times R_r$, where R_i is the ring of $n_i \times n_i$ matrices over the division ring Δ_i, $i = 1, \ldots, r$. It is a very extense proposition, based on [2], and the proof is to much extended and requires too many tools, hence it will not be proven here.*

(1) Identify R_i with the i-th component of the direct product. Let z_i be the r-tuple with the identity of R_i at position i and zero in all other positions, i.e. $z_i := (0, \ldots, 0, \underset{i)}{1_{R_i}}, 0, \ldots, 0)$. Then $R_i = z_i R$ and for any $a \in R_i$, $z_i a = a$ and $z_j a = 0$ for $j \neq i$. The elements $z_1, \ldots, z_r$ are all of the primitive central idempotents[5] of R. They are pairwise orthogonal and $\sum_{i=1}^{r} z_i = 1$ in R.

(2) Let N be any left R-module and let $z_i N = \{z_i x \ : \ x \in N\}$, $1 \leq i \leq r$. Then $z_i N$ is a left R-submodule and it is an R_i-module on which R_j acts trivially for all $j \neq i$. And also

$$N = z_1 N \oplus z_2 N \oplus \cdots \oplus z_r N$$

(3) The simple R-modules are the simple R_i-modules on which R_j acts trivially for $j \neq i$ in the following sense. Let M_i be the unique simple R_i-module. We may consider M_i as an R-module by letting R_j act trivially for all $i \neq j$. Then $M_1, \ldots, M_r$ are pairwise non-isomorphic simple R-modules and any simple R-module is isomorphic to one of $M_1, \ldots, M_r$. Explicitly, the R-module M_i is isomorphic to the simple left ideal $(0, \ldots, 0, \underset{i)}{L_i}, 0, \ldots, 0)$ of all elements of R whose i-th component, L_i, consists of matrices with arbitrary entries in the first column and zeros elsewhere.

[5] The *central* elements are those who belong to the center. The *primitive* elements are those that cannot be decomposed as a sum of central orthogonal elements.

(4) *For any R-module N the R-submodule z_iN is a direct sum of simple R-modules, each of which is isomorphic to the module in (3). In particular, if M is a simple R-module, then there is a unique i such that $z_iM = M$ and for this index i we have $M_i \cong M$ and for all $j \neq i$, $z_jM = 0$.*

(5) *If each Δ_i equals the field k, then R is a vector space over k of dimension $\sum_{i=1}^{r} n_i^2$ and $\dim_k \mathcal{Z}(R) = r$.*

This big proposition gives a characterization of the structure of the simple $k[G]$-module. Note that, in particular, the division ring Δ_i is a vector space over k of dimension $n_i \leq n$. Now, the following proposition will show that this implies $\Delta_i = k$.

Proposition 2.5.4. *If Δ is a division ring that is a finite dimensional vector space over an algebraically closed field k and $k \subseteq \mathcal{Z}(\Delta)$ (its center), then $\Delta = k$.*

Proof. Suppose $k \subseteq \mathcal{Z}(\Delta)$. Then, for each $\alpha \in \Delta$, the division ring generated by α and k is a field. Also, since Δ is finite dimensional over k, the field extension $k(\alpha)$ of k will be also finite. But as k is algebraically closed, it cannot have finite non-trivial extensions, hence $k(\alpha) = k$ for all $\alpha \in \Delta$, ergo $\Delta = k$. □

With this proposition we get that each R_i in Wedderburn's decomposition of $k[G]$, where k is an algebraically closed field, is a matrix ring over k, i.e. $M_{n_i}(k)$.

Proposition 2.5.5. *The number of matrix rings in $k[G]$'s Wedderburn decomposition is equal to the number of conjugacy classes of G.*

Proof. Let $\mathcal{K}_1, \ldots, \mathcal{K}_s$ be the conjugacy classes of G. Recall that they are disjoint, and form a partition of G. Now, for each i define

$$X_i := \sum_{g \in \mathcal{K}_i} g \qquad \in k[G]$$

As the conjugacy classes are disjoint, we have that $\mathcal{K}_i$ and $\mathcal{K}_j$ have distinct elements and hence these X_i's are linearly independent elements of $k[G]$. Furthermore, since conjugation by a group element permutes the elements of each class, $h^{-1}X_ih = X_i$, i.e. X_i commutes with all group elements. This implies that $X_i \in \mathcal{Z}(k[G])$.

Finally we show that X_i's form a basis of $\mathcal{Z}(k[G])$ and we will have $s = r = \dim_k \mathcal{Z}(k[G])$. We already discussed that they are linearly independent. We now show that they span $\mathcal{Z}(k[G])$. Let $X = \sum_{g \in G} \alpha_g g$ be an arbitrary element of $\mathcal{Z}(k[G])$. Since $h^{-1}Xh = X$,

$$\sum_{g \in G} \alpha_g h^{-1}gh = \sum_{g \in G} \alpha_g g.$$

since the elements of G form a basis of $k[G]$, the coefficients of g in the above two sums are equal, i.e. $\alpha_{hgh^{-1}} = \alpha_g$. Since h was arbitrary, every element in the same conjugacy class of a fixed group element g has the same coefficient in X, hence X can be written as a linear combination of the X_i's. □

Before going through some particular examples, let me state this theorem that summarizes everything we have seen during these last two sections and then a corollary directly from it. This theorem does not need a proof, as it collects all the results seen since now, that have already been proved.

Theorem 2.5.6. *Let G be a finite group and k an algebraically closed field whose characteristic does not divide $|G|$.*

(1) $k[G] \cong M_{n_1}(k) \times M_{n_2}(k) \times \cdots \times M_{n_r}(k)$.

(2) $k[G]$ has exactly r distinct isomorphism types of irreducible modules and these have dimensions (as k-vector spaces) $n_1, \ldots, n_r$ respectively (and so G has exactly r inequivalent irreducible representations over k of the corresponding degrees).

(3) $\sum_{i=1}^{r} n_i^2 = |G|$.

(4) r equals the number of conjugacy classes in G.

Corollary 2.5.7. *Let G be a finite abelian group. Every irreducible representation over an algebraically closed field k is 1 dimensional (i.e., a homomorphism $G \to k^*$) and G has $|G|$ inequivalent irreducible representations over k. Furthermore, every finite dimensional matrix representation over k of G is equivalent to a representation into a group of diagonal matrices.*

Proof. If G is abelian, $k[G]$ is a commutative ring. Since a $m \times m$ matrix is not commutative whenever $m > 1$, we must have each $n_i = 1$. Thus, $r = |G|$ (and equals also the number of conjugacy classes of G). Since every $k[G]$-module is a direct sum of irreducible submodules, there is a basis such that the matrices are diagonal with respect to this basis. □

Example 2.5.8. To end the section, we are going to see some examples. For these examples, to make it easier, we are going to take $k = \mathbb{C}$, the complex field.

(1) Take G an abelian finite group. Then, by the theorem of decomposition for abelian groups. We have $G \cong C_1 \times \cdots \times C_n$, where $C_i = \langle x_i \rangle$ are cyclic groups. Let $d_i := |C_i| = |\langle x_i \rangle$ be its degree. Now we will describe the irreducible $\mathbb{C}$-representations of G, i.e. the homomorphisms $G \to \mathbb{C}^*$.

If $f : C_i \to \mathbb{C}^*$ is a homomorphism, it should preserve the fact that $x_i^{d_i} = 1$ and as f sends the neutral to the neutral, it must be $f(x_i)^{d_i} = f(x_i^{d_i}) = f(1) = 1$, hence $f(x_i)^{d_i} = 1$ meaning that $f(x_i)$ must be one of the d_i-th roots of unity in $\mathbb{C}^*$.

So, fixing an $i \in \{1, \ldots, n\}$, take $C_i = \langle x_i \rangle$ one of these cyclics of order d_i, and let $\{\xi_1, \ldots, \xi_{d_i}\}$ be the d_i d_i-th roots of 1 in $\mathbb{C}^*$. We have then d_i possible homomorphisms

$$\begin{aligned} \varphi_j : G &\longrightarrow \mathbb{C}^* \\ x_i &\longmapsto \xi_j \end{aligned} \qquad j = 1, \ldots, d_i$$

Then, for each $i = 1, \ldots, n$ we have d_i choices for φ_j, then for all G we will have in total $d_1 \cdot \cdots \cdot d_n$ choices of representations, i.e. we have $|G|$ possible representations and all of them are like these.

(2) Let $G = S_3$. By 2.5.5, the number of $\mathbb{C}$-representations of G equals the number of conjugacy classes of S_3, in this case 3. Then we must have only three representations (up to equivalence) and let them be

$$\varphi_i : G \longrightarrow \mathbf{GL}\left(\mathbb{C}\right) \cong \mathbb{C}^*, \qquad i = 1,2,3.$$

We also have the fact that $\sum_{i=1}^{r} n_i^2 = |G|$, where in our case $r = 3$ and n_i is the degree of the representation φ_i, for $i = 1,2,3$. Hence, we have $n_1^2 + n_2^2 + n_3^2 = 6$, for $n_1, n_2, n_3 \in \mathbb{Z}_{\geq 1}$. So, up to permutations of the sub-indexes, the only possibility is to have $n_1 = 1$, $n_2 = 1$ and $n_3 = 2$. Now, let's try to find each representation.

- $\underline{\varphi_1 \text{ of degree } 1}$. This can be the trivial representation, the one that sends every $\sigma \in S_3$ to 1 in $\mathbb{C}^*$.
- $\underline{\varphi_2 \text{ of degree } 1}$. This can be the following

$$\varphi_2(\sigma) = \begin{cases} 1 & \text{if } \sigma \text{ is an even permutation} \\ -1 & \text{if } \sigma \text{ is an odd permutation} \end{cases}$$

  Recall that the parity of a permutation was $(-1)^s$, where s was the number of transpositions in which the permutation decomposed.
- $\underline{\varphi_3 \text{ of degree } 2}$. This is the tricky one.

  Consider the correspondence with $k[G]$-modules, in this case $k[S_3]$-modules, and the representations (see 2.3.3). Let V be a 3-dimensional vector space with basis $\{e_1, e_2, e_3\}$ and consider the action on V by $G = S_3$, i.e. permuting the indexes of the basis. Then the vector $t = e_1 + e_2 + e_3$ is S_3-invariant and thus it spans a S_3-invariant 1-dimensional vector subspace. Call it U. Then, by Maschke's Theorem 2.4.1 there is a 2-dimensional S_3-invariant vector subspace W.

  Let's now see that this W cannot be reduced in $W_1 \oplus W_2$ (i.e. it is irreducible). If it was reducible, then the affording representation would have as a matrix a diagonal 2×2 matrix, hence φ_3 would have a diagonal matrix too. This cannot happen as the permutations would commute on their action, but this is impossible since G is not abelian.

  Thus, W affords the representation of degree 2 irreducible. We can see W as

$$W = \{w \in V \ : \ w = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 \text{ with } \alpha_1 + \alpha_2 + \alpha_3 = 0\}.$$

  Clearly $e_1' := e_1 - e_2$ and $e_2' := e_2 - e_3$ are linearly independent vectors in W, hence they form a basis in W, $\mathcal{B}_2 = \{e_1', e_2'\}$. With this basis one can find the matrix representations for given elements in S_3. For example, let's take $(1,2,3) \in S_3$ and try to find the matrix representation of this element. We have to see how is the action in the basis elements of W:

$$(1,2,3)e_1' = (1,2,3)(e_1 - e_2) = (e_2 - e_3) = e_2'$$

and

$$(1,2,3)e'_2 = (1,2,3)(e_2 - e_3) = (e_3 - e_1) = -(e_1 - e_2 + e_2 - e_3) = -e'_1 - e'_2$$

hence, the matrix representation would be

$$(1,2,3) \longmapsto \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$$

and for the element $(1,2) \in S_3$ one can follow the same calculations and get the matrix representation

$$(1,2) \longmapsto \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Chapter 3

Character Theory

This chapter contains all the basic knowledge about Character Theory, as well as the results needed to prove Burnside's Theorem at the end. It will be based on the reference [2] as well as on [5].

In general, for groups of large order the representations are difficult to compute and unwieldy if not impossible to write down. For example, a matrix representation of degree 100 involves matrices with over 10,000 entries, and a number of 100×100 matrices may be required to describe the representation, even on a set of generators for the group.

In this chapter we will be attaching, for each representation $\varphi : G \to \mathrm{GL}_n(k)$, an element of k to each matrix $\varphi(g)$ and we shall see that this number can, in many instances, be computing without really knowing the matrix $\varphi(g)$. Moreover, we shall see that these invariants are independent of the similarity class of φ (i.e. they are the same for a fixed $g \in G$ if the representation φ is replaced by an equivalent representation[1] and that they, in some sense, characterize the similarity classes of representations of G.

3.1 Character theory: basic knowledge

Throughout this section G is a finite group and k will be an arbitrary algebraically closed field. All the representations considered are assumed to be finite dimensional.

In this section I will give the definition of character of a group, and then prove that two representations are equivalent if, and only if, they have the same character. Moreover I will give an important general proposition regarding ring decompositions and then I will apply it to our case to show how to find "irreducible characters".

Definition 3.1.1 (Class function)**.** Let G be a finite group and k an arbitrary field. A *class function* is any function from G into k which is constant on the conjugacy classes of G, i.e. $f : G \to k$ such that $f(g^{-1}xg) = f(x)$ for all $g, x \in G$.

[1] see 2.3.6

Definition 3.1.2 (Character)**.** If φ is a representation of G afforded by the $k[G]$-module V, the *character* of φ is the function

$$\begin{aligned} \chi : G &\longrightarrow k \\ g &\longmapsto \chi(g) := \mathrm{tr}\varphi(g) \end{aligned}$$

where $\mathrm{tr}\varphi(g)$ means the trace of the matrix $\varphi(g)$ with respect to some basis of V (i.e. the sum of the diagonal entries of the matrix). The *degree* of a character is the degree of any representation affording it.

Definition 3.1.3 (Irreducible character)**.** A character is called *reducible* or *irreducible* according to whether the representation affording it is reducible or irreducible, respectively.

In the notation we shall also refer to χ as the character afforded by the $k[G]$-module V. In general, a character is *not* a homomorphism from a group into either the additive or multiplicative group of the field.

Example 3.1.4. Let's see some examples before continuing.

(1) The character of the trivial representation is the function $\chi(g) = 1$ for all $g \in G$. This is called the *principal character* of G.

(2) If φ is a degree 1 representation, then the character is usually identified with φ by identifying the entry of the 1×1 matrix representation. Thus, for abelian groups, irreducible complex representations and their characters are the same.

(3) Let $\Pi : G \to S_n$ be a permutation representation and let φ be the resulting linear representation on the basis $e_1, \ldots, e_n$ of the vector space:

$$\varphi_g(e_i) = e_{\Pi_g(i)}$$

(see previous examples). With respect to this basis, the matrix of $\varphi_g = \varphi(g)$ has a 1 in the diagonal entry (i,i) if $\Pi(g)$ fixes i; otherwise the matrix of φ_g has a zero in position (i,i). Thus, if π is the character of φ, then $\pi(g)$ equals the number of fixed points of g on $\{1, \ldots, n\}$.

(4) Recall the definition of *regular representation* in 2.5.8 number (2). Now we will define the *regular character*, which is a special case of the previous example, when Π is the regular permutation representation of G. If we call ρ the character, then we have

$$\rho(g) = \begin{cases} 0 & \text{if } g \neq 1 \\ |G| & \text{if } g = 1 \end{cases}$$

The reason of this character definition follows from the definition of the matrix representation of the regular representation.

Proposition 3.1.5. *Characters are also to be considered as functions defined on the group ring $k[G]$.*

Proof. If V is a $k[G]$-module whose corresponding representation has character χ, then each element of the group ring $k[G]$ acts as a linear transformation $V \to V$. Thus, for each $\sum_{g\in G} \alpha_g g \in k[G]$ there is a trace associated when it is considered as a linear transformation from V to V. The trace of $g \in G$ acting on V is, by definition, $\chi(g)$. Since the trace of any linear combination of matrices is the linear combination of the traces, the trace of $\sum_{g\in G} \alpha_g g$ acting on V would be $\sum_{g\in G} \alpha_g \chi(g)$.

To end the proof, given any map $f : G \to k$, this can be extended by linearity and obtain $k[G] \to k$ homomorphism □

Now we will see two important results regarding characters, that will help us work with them. The first one states that the character of a representation is the sum of the characters of the constituents appearing in a direct sum decomposition. The second one will state that two representations are equivalent if, and only if, they have the same character.

Theorem 3.1.6. *The character of a representation is the sum of the characters of the constituent appearing in a direct sum decomposition.*

Proof. Consider $M = M_1 \oplus M_2$ and φ the representation afforded by the module M. We choose a basis of M consisting on one of M_1 extended adding the basis of M_2. Then, the matrix representation with respect to this basis is of the form

$$\varphi(g) = \begin{pmatrix} \varphi_1(g) & 0 \\ 0 & \varphi_2(g) \end{pmatrix}$$

where φ_i is the representation afforded by M_i, for $i = 1, 2$. From this point it is immediate that if ψ is the character of φ and ψ_1 and ψ_2 are the characters of φ_1 and φ_2 respectively, then $\psi(g) = \psi_1(g) + \psi_2(g)$, for all $g \in G$. Then, by induction we obtain the result. □

This last theorem states that if ψ is the character afforded by the module in (2.5.3) above, this gives

$$\psi = a_1\chi_1 + \cdots + a_r\chi_r$$

Thus, every character over k is a non-negative sum of irreducible characters over k. Conversely, by taking direct sum of modules, one sees that every such sum of characters is the character of some complex representation of G.

We now focus our attention on the next theorem that, as it may be expected, it will state that two representations are equivalent if and only if they have the same character. For this, though, we have to go through a deep proposition regarding ring theory, that we skipped before, but now we can not avoid.

Theorem 3.1.7. *Two representations are equivalent if and only if they have the same character.*

Proof. One implication is easy: if ψ and φ are equivalent representations, then the matrices $M_{\varphi(g)}$ and $M_{\psi(g)}$ of $\varphi(g)$ and $\psi(g)$ satisfy that there exists a fixed invertible matrix P such that $P^{-1}M_{\psi(g)}P = M_{\varphi(g)}$ for all $g \in G$. Hence, it is easy to see that the traces are the same.

Conversely, let $z_1, \dots, z_r$ be the primitive central idempotents of $k[G]$ described in 2.5.3. Since they are orthogonal, it is trivial that they are k-linearly independent elements of $k[G]$. As above in 3.1.5, each irreducible character χ_i is a function on $k[G]$. By proposition 2.5.3, part (3), we have

(a) If $j \neq i$, then $z_j M_i = 0$, i.e. z_j acts as the zero matrix on M_i, hence $\chi_i(z_j) = 0$.

(b) z_i acts as the identity on M_i, hence $\chi_i(z_i) = n_i$.

Thus, $\chi_1, \dots, \chi_r$ are multiples of the dual basis to the independent set $z_1, \dots, z_r$, hence are linearly independent functions[2]. Now, if the $k[G]$-module M described in 2.5.3 above can be decomposed in a different fashion into irreducibles, say

$$M \cong b_1 M_1 \oplus \cdots \oplus b_r M_r$$

then we would obtain a relation

$$a_1\chi_1 + \cdots + a_r\chi_r = b_1\chi_1 + \cdots + b_r\chi_r$$

and by linear independence of the irreducible characters, we would get $b_i = a_i$ for all $i \in \{1, \dots, r\}$. Thus, in any decomposition of M into a direct sum of irreducibles, the multiplicity of the irreducible M_i is the same, for all $1 \leq i \leq r$ and in particular we get the result. □

3.2 Orthogonality relations between characters

In this section we are going to build an Hermitian inner product structure on the space of class functions and prove that the irreducible characters form an orthonormal basis with respect to this inner product. We first check that they already form a basis of the complex class functions.

For the rest of the chapter, we will take $k = \mathbb{C}$ as our field, as we are attaching numerical data to representations over k. Hence, we restrict ourselves to the case of

[2] Here is where we see the importance of extending a character from k to all $k[G]$ ring. From this point of view, a character can be thought as a linear form from $k[G]$ to k, i.e. an element of the dual space of $k[G]$ as a vector space over k.Then, $z_1, \dots, z_r$ are elements linearly independent of $k[G]$ (but not necessarily a basis, as they may not span all $k[G]$). Then we consider the subspace W of $k[G]$ generated by $z_1, \dots, z_r$ and then they define the dual basis $z_1^*, \dots, z_r^*$ of W^*. Now, we can restrict the character to this vector subspace and think it as a dual form of W^*. Hence, it can be expressed as a linear combination of irreducible characters χ_i we discussed before. But these irreducible characters we know how they act on $z_1, \dots, z_r$ on W, therefore, we can deduce from here that any irreducible character is a multiple of the z_i^* corresponding and then every character is the linear combination of the dual basis $a_1 z_1^* + \cdots + a_r z_r^*$, where $a_1, \dots, a_r$ are uniquely determined.

complex representations and complex characters, as we will be able to define an explicit inner product.

Proposition 3.2.1. *The irreducible characters are a basis for the space of all complex class functions.*

Proof. The vector space of all complex valued class functions on G has a basis consisting of the functions which are 1 on a given class and zero on all the other classes[3]. Therefore, there are r of these, where r is the number of conjugacy classes of G, so the dimension of the complex vector space of class functions is r. Since the number of complex irreducible characters of G equals the number of conjugacy classes and these are linearly independent class functions, we see the proposition. □

Definition 3.2.2 (Inner product)**.** Let θ and ψ be class functions. We define

$$(\theta,\psi) = \frac{1}{|G|}\sum_{g\in G}\theta(g)\overline{\psi(g)}$$

where the bar denotes the complex conjugation.

Proposition 3.2.3. *The product defined in 3.2.2 is an Hermitian inner product.*

Proof. For $\alpha,\beta\in\mathbb{C}$ and $\theta_1,\theta_2,\psi_1,\psi_2,\theta$ and ψ class functions, we have

(a) $(\alpha\theta_1+\beta\theta_2,\psi) = \alpha(\theta_1,\psi)+\beta(\theta_2,\psi)$,

(b) $(\theta,\alpha\psi_1+\beta\psi_2) = \overline{\alpha}(\theta,\psi_1)+\overline{\beta}(\theta,\psi_2)$,

(c) $(\theta,\psi) = \overline{(\psi,\theta)}$,

and they can be proved easily by playing with sums and conjugations in complex numbers. □

Our principal aim now is to show that the irreducible characters form an orthonormal basis for the space of complex class functions $f : G \to \mathbb{C}$ with respect to this Hermitian form. We already saw in 3.2.1 that they form a basis, so it just remain to show the orthonormality. This fact will follow from the orthogonality of the primitive central idempotents, once we have explicitly determined these in the next proposition.

Proposition 3.2.4. *Let $z_1,\dots,z_r$ be the orthogonal primitive central idempotents in $\mathbb{C}[G]$ labelled in such a way that z_i acts as the identity on the irreducible $\mathbb{C}[G]$-module M_i and as zero in all other M_j, for $i\neq j$, and let χ_i be the character afforded by M_i. Then*

$$z_i = \frac{\chi_i(1)}{|G|}\sum_{g\in G}\chi_i(g^{-1})g.$$

[3] Consider any complex class function $f : G\to\mathbb{C}$, then it is in particular $f : \mathbb{C}[G]\to\mathbb{C}$ such that $f(g^{-1}xg) = f(x)$ for all $x\in G$. Then, all the class functions seen as dual applications, form a vector space with generators $f_i(x_i) = 1$ and $f_i(x_j) = 0\ \forall j\neq i$, $i = 1,\dots,|G|$. Then if we take $\mathcal{K}_1,\dots,\mathcal{K}_r$ all the conjugacy classes of G with representatives $g_1,\dots,g_r$ respectively, then every f can be expressed as $f = \sum_{i=1}^{r}\lambda_i f_i$, where $\lambda_i\in\mathcal{C}$. Then f_i generate the vector space of complex class functions of G and, as they can be associated with characters, as we saw before, they are linearly independent, hence basis.

Proof. Let $z = z_i$ and write

$$z = \sum_{g \in G} \alpha_g g.$$

Now, recall from example 4 the definition of the regular character. Then, by theorem 2.5.6 we have that $\mathbb{C}[G] \cong M_{n_1}(\mathbb{C}) \times \cdots \times M_{n_r}(\mathbb{C})$ and by proposition 3.2.4 each $M_{n_i}(\mathbb{C})$ decomposes further as a direct sum of n_i isomorphic simple ideals. These ideals give a complete set of isomorphic classes of irreducible $\mathbb{C}[G]$-modules. Thus, the regular representation over $\mathbb{C}$ of G decomposes as the direct sum of all irreducible representations of G, each appearing with multiplicity equal to the degree of that (irreducible) representation. Taking now the characters, this implies that the regular character decomposes as follows:

$$\rho = \sum_{i=1}^{r} \chi_i(1)\chi_i$$

due to the fact that $\chi_i(1)$ is the degree of the corresponding irreducible representation.

To find the coefficient α_g, apply ρ to zg^{-1} and using linearity of ρ together with its definition, we obtain that $\rho(zg^{-1}) = \alpha_g|G|$. Computing $\rho(zg^{-1})$ and using ρ's decomposition, we finde

$$\sum_{j=1}^{r} \chi_j(1)\chi_j(zg^{-1}) = \alpha_g|G|$$

Let φ_j be the irreducible representation afforded by M_j, for $1 \leq j \leq r$. Since we may consider φ_j as an algebra homomorphism from $\mathbb{C}[G]$ into $\mathrm{End}(M_j)$, we obtain $\varphi_j(zg^{-1}) = \varphi_j(z)\varphi_j(g^{-1})$. Also, we have already observed that $\varphi_j(z)$ is 0 if $i \neq j$ and $\varphi_i(z)$ is the identity endomorphism on M_i. Thus

$$\varphi_j(zg^{-1}) = \begin{cases} 0 & \text{if } j \neq i \\ \varphi_i(g^{-1}) & \text{if } j = i \end{cases}$$

This proves $\chi_j(zg^{-1}) = \chi_i(g^{-1})\delta_{ij}$, where δ_{ij} is zero if $i \neq j$ and 1 if $i = 1$. Substituing this into the last equation gives $\alpha_g = \frac{1}{|G|}\chi_i(1)\chi_i(g^{-1})$. This is the coefficient of g in the statement of the proposition. □

We already can prove the orthonormality of irreducible characters by just doing some little calculations, which are done in 3.2.6. But there remains one little step at the end of the proof, which states that $\chi_j(g^{-1}) = \overline{\chi_j(g)}$. This little step is proven in the next proposition.

Proposition 3.2.5. *If ψ is any character of G then $\psi(x)$ is a sum of roots of 1 in $\mathbb{C}$ and $\psi(x^{-1}) = \overline{\psi(x)}$ for all $x \in G$.*

Proof. Let φ be a representation whose character is ψ, fix an element $x \in G$ and let $|x| = m$. Since the minimal polynomial of $\varphi(x)$ divides $X^m - 1$, hence has distinct roots, there is a basis of the underlying vector space such that the matrix of $\varphi(x)$ with respect to this basis is a diagonal matrix with m-th roots of 1 on the diagonal. Since $\psi(x)$ is the sum of the diagonal entries, $\psi(x)$ is a sum of roots of 1. Moreover, if ξ is a root of 1, it is clear that $\xi^{-1} = \overline{\xi}$. Thus, the inverse of a diagonal matrix with roots of 1 on the diagonal is the diagonal matrix with the complex conjugates of those roots of 1 on the diagonal Since the complex conjugate of a sum is the sum of complex conjugates, $\psi(x^{-1}) = \text{trace}(\varphi(x^{-1})) = \overline{\text{trace}(\varphi(x))} = \overline{\psi(x)}$. □

It is important to notice that in this proof we first fixed a group element x and then chose a basis of the representation space so that $\varphi(x)$ was a diagonal matrix. It is always possible to diagonalize a single element but it is possible to simultaneously diagonalize all $\varphi(x)$'s if and only if φ is similar to a sum of degree 1 representations.

Finally we prove our result.

Proposition 3.2.6. *The irreducible characters are orthonormal.*

Proof. It will follow directly from the orthogonality of the central primitive idempotents via the following calculations. Let δ_{ij} be the "Kronecker Delta", valued $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ for $i \neq j$. Then it is clear that $z_i\delta_{ij} = z_iz_j$. Indeed, as if $i = j$ then we have $z_iz_j = z_i^2 = z_i$ for being idempotent, and if $j \neq i$ then we have $z_iz_j = 0$ because they are orthogonal, and $\delta_{ij} = 0$ so it is consistent. Now, we calculate z_iz_j with the given values in 3.2.4 and get

$$z_i\delta_{ij} = z_iz_j = \frac{\chi_i(1)}{|G|}\frac{\chi_j(1)}{|G|}\sum_{g,h\in G}\chi_i(g^{-1})\chi_j(h^{-1})gh$$

Now, substituting $y = gh$ and $x = h$, we have that $g = x^{-1}$ and hence, the last calculations remain

$$\cdots = \frac{\chi_i(1)}{|G|}\frac{\chi_j(1)}{|G|}\sum_{y\in G}\left(\sum_{x\in G}\chi_i(xy^{-1})\chi_j(x^{-1})\right)y$$

Now, in the other side of the equation, we had $z_i\delta_{ij}$ which can be rewritten, taking into account proposition 3.2.4, as $\delta_{ij}\frac{\chi_i(1)}{|G|}\sum_{g\in G}\chi_i(g^{-1})g$ hence the equality remains

$$\delta_{ij}\frac{\chi_i(1)}{|G|}\sum_{g\in G}\chi_i(g^{-1})g = \frac{\chi_i(1)\chi_j(1)}{|G|^2}\sum_{y\in G}\left(\sum_{x\in G}\chi_i(xy^{-1})\chi_j(x^{-1})\right)y$$

Since the elements of G are a basis of $\mathbb{C}[G]$, we may equate coefficients with those of z_i in the other side, i.e.

$$\delta_{ij}\frac{\chi_i(1)}{|G|}\chi_i(g^{-1}) = \frac{\chi_i(1)\chi_j(1)}{|G|^2}\sum_{x\in G}\chi_i(xg^{-1})\chi_j(x^{-1}).$$

Simplifying and replacing g by g^{-1} gives

$$\delta_{ij}\frac{\chi_i(g)}{\chi_j(1)} = \frac{1}{|G|}\sum_{g\in G}\chi_i(xg)\chi_j(x^{-1})$$

for all $g \in G$. Taking now $g = 1$ we have that

$$\delta_{ij} = \frac{1}{|G|}\sum_{x\in G}\chi_i(x)\chi_j(x^{-1})$$

Now we have that $\chi_j(x^{-1}) = \overline{\chi_j(x)}$ for all $x \in G$ by the proposition 3.2.5. □

Combining the last two propositions we can state our final main theorems.

Theorem 3.2.7 (The First Orthogonality Relation for Group Characters)**.** *Let G be a finite group and let $\chi_1, \dots, \chi_r$ be the irreducible characters of G over $\mathbb{C}$. Then with respect to the inner product $(\cdot, \cdot)$ defined in 3.2.2, we have*

$$(\chi_i, \chi_j) = \delta_{ij}$$

and the irreducible characters are an orthonormal basis for the space of class functions. In particular, if θ is any class function,

$$\theta = \sum_{i=1}^{r} (\theta, \chi_i)\chi_i.$$

Proof. We just proved that the irreducible characters form an orthonormal basis for the space of class functions. If θ is any class function, $\theta = \sum_{i=1}^{r} a_i \chi_i$ for some $a_i \in \mathbb{C}$. It follows from linearity of Hermitian product that $a_i = (\theta, \chi_i)$, as stated. □

3.3 Characters of some groups of small order

As an example of character calculations we find here some character tables that are listed in [2].

The *character table* of a finite group is a table of the character values formatted as follows: list representative of the r conjugacy classes along the top row and list the irreducible characters down the first column. The entry in the table in row χ_i and column g_j is $\chi_i(g_j)$. The character table of a finite group is unique up to a permutation of its rows and columns. It is customary to make the principal character the first row and the identity the first column and to list the characters in increasing order by degrees.

A large number of character tables is given in the *Atlas of Finite Groups* by Conway, Curtis, Norton, Parker and Wilson, Clarendon Press, 1985. These include the character table of the monster simple group, M. This group has 194 irreducible characters. çThe smallest degree of a nonprincipal irreducible character of M is 196883 and the largest degree is on the order of 2×10^{26}. Nonetheless, it is possible to compute the values of all these characters on all conjugacy classes of M.

Example 3.3.1. Consider $G = \langle x \rangle$ be the cyclic group of order 2. Then G has 2 conjugacy classes and two irreducible characters. Indeed, as any homomorphism $f : \langle x \rangle \to \mathbb{C}^*$ should preserve the fact that $f(1) = 1$ so $f(x)^2 = f(x^2) = f(1) = 1$ meaning that f must be one of the 2 roots of 1 in $\mathbb{C}$. Then we have two possible representations:

$$\psi_1(x) = 1 \qquad \text{and} \qquad \psi_2(x) = -1$$

(and they send 1 to 1 trivially) of degree obviously 1. Then the table is this:

classes:	1	x
sizes:	1	1
χ_1	1	1
χ_2	1	-1

Table 3.1: Character Table of $\mathbb{Z}_2$

Example 3.3.2. Consider now $G = \langle x \rangle$ of degree 3. By the same reasoning as in the previous example, we have that every representation must send each element to a 3-th root of 1 in $\mathbb{C}$. Call $\zeta \in \mathbb{C}$ one of the three third roots of 1 Then, we have three options, and they let us the next table:

classes:	1	x	x^2
sizes:	1	1	1
χ_1	1	1	1
χ_2	1	ζ	ζ^2
χ_3	1	ζ_2	ζ

Table 3.2: Character Table of $\mathbb{Z}_3$

Again, the representations will be of degree 1 because we have the theorem stating that the sum of the squares of the degrees of the representations must be equal to $|G|$ which in this case is three, ergo we have $n_1^2 + n_2^2 + n_3^2 = 3$ and the only possibility is $n_1 = n_2 = n_3 = 1$. Hence, each representation matrix is a 1×1 matrix over $\mathbb{C}^*$, i.e., a complex nonzero number, and then the character is this number itself.

Example 3.3.3. Let $G = S_3$. In 2.5.8 we computed all the possible representations, and we found one representation of degree 2 and two of degree 1. We can use this to compute the table. I will compute manually some of them to have an example. Take φ_3 of this example, which was the degree 2 representation. We had that

$$\varphi_3((1,2,3)) = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}, \qquad \varphi_3((1,2)) = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \varphi_3(\mathrm{id}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

hence we just have to calculate the trace of each matrix to obtain $\chi_3((1,2,3)) = -1$, $\chi_3((1,2)) = 0$ and $\chi_3(\mathrm{id}) = 2$. The other representations were of degree 1 so they are only 1×1 matrices. The table is then:

classes:	1	$(1,2)$	$(1,2,3)$
sizes:	1	3	2
χ_1	1	1	1
χ_2	1	-1	1
χ_3	2	0	-1

Table 3.3: Character Table of S_3

Chapter 4

Burnside's Theorem

In this chapter we will finally prove the Burnside Theorem 4.1.1 that will use all the theory seen in previous chapters about characters and representations.

4.1 Burnside's proof

The following result was proved by Burnside in 1904. Although purely group-theoretic proofs of it were discovered, the original proof by Burnside and the one we will be giving is very accessible and lies over character theory. The theorem states the following.

Theorem 4.1.1 (Burnside)**.** *For p and q primes, every group of order $p^a q^b$ is solvable.*

To move to the proof we will need to establish some previous results. The first propositions I will state now are to prove corollary 4.1.9, which states that the degrees of the irreducible characters of any finite group divide its order. Then there will come the final lemmas 4.1.10 and 4.1.11 which will lead directly to the proof of Burnside's Theorem.

Firstly, I will recall the definition of algebraic integer and state a proposition that characterizes them. This has all been seen in class of "Introducció a l'Àlgebra Commutativa".

Definition 4.1.2 (Algebraic integer)**.** An element $\alpha \in \mathbb{C}$ is called an *algebraic integer* if it is a root of a monic polynomial with coefficients from $\mathbb{Z}$.

Proposition 4.1.3. *Let $\alpha \in \mathbb{C}$.*

(1) The following are equivalent:

(i) α is an algebraic integer.

(ii) α is algebraic over $\mathbb{Q}$ and the minimal polynomial of α over $\mathbb{Q}$ has integer coefficients, and

(iii) *$\mathbb{Z}[\alpha]$ is a finitely generated $\mathbb{Z}$-module (where $\mathbb{Z}[\alpha]$ is the sub-ring of $\mathbb{C}$ generated by $\mathbb{Z}$ and α, i.e. all the $\mathbb{Z}$-linear combinations of non-negative powers of α).*

(2) *The algebraic integers in $\mathbb{C}$ form a ring and the algebraic integers in $\mathbb{Q}$ are the elements of $\mathbb{Z}$.*

Corollary 4.1.4. *For every character ψ of the finite group G, $\psi(x)$ is an algebraic integer, for all $x \in G$.*

Proof. By proposition 3.2.5 of the previous chapter, $\psi(x)$ is a sum of roots of 1. Each root of 1 is an algebraic integer, so the result follows immediately from proposition 4.1.3. □

Next proposition is more character related, and we will see how characters and algebraic integers are related. We will use the following notation, which is the usual one we have been using all along. Take an arbitrary finite group G and $\chi_1, \dots, \chi_r$ are the distinct irreducible complex characters of G, $\mathcal{K}_1, \dots, \mathcal{K}_r$ are the conjugacy classes of G and φ_i is an irreducible matrix representation whose character is χ_i for each i.

Before going into it, I state the Schur's Lemma, which is proposed as an exercise in [2] and used in the proof. I will state the general Schur's Lemma I found in [9], which is a pure representation theory lemma, and then I will state a corollary, which is the exercise of [2] and the one we will use directly.

Definition 4.1.5 (G-linear)**.** Let k be an algebraically closed field and G a finite group. Let V and W be k-vector spaces and $f : V \to W$ a map. It is said to be *G-linear* if it is a linear map $f : V \to W$ as $k[G]$-modules.

Lemma 4.1.6 (Schur)**.** *Let V and W be vector spaces and ρ_V and ρ_W be irreducible representations of a finite group G in V and W respectively.*

(1) *If V and W are not isomorphic, then there are no nontrivial G-linear maps $f : V \to W$.*

(2) *If $V = W$ are finite dimensional over an algebraically closed field (e.g. $\mathbb{C}$ in our particular case) and $\rho_V = \rho_W$, then the only nontrivial G-linear maps $f : V \to V$ are the identity and homothecy.*

Proof. (1) Suppose $f : V \to W$ is nonzero G-linear map. We will prove that $V \cong W$, i.e., that f is an isomorphism.

For each $g \in G$, choose any $x \in \operatorname{Ker} f$. Then

$$f(\rho_V(g)(x)) = \rho_W(g)(f(x)) = \rho_W(g)(0) = 0$$

where $\rho_V(g)(x)$ is the value by the linear map $\rho_V(g) : V \to W$ of x. Then, this means that $f(\rho_V(g)(x)) = 0$ meaning that $\rho_V(g)(x) \in \operatorname{Ker} f$. That is, for each $x \in \operatorname{Ker} f$, $\rho_V(g)(x) \in \operatorname{Ker} f$ which implies that $\operatorname{Ker} f$ is G-invariant, hence a subrepresentation. But we assumed that the representations where irreducible, hence $\operatorname{Ker} f = 0$ (as f is nonzero, so we assume $\operatorname{Ker} f$ cannot be the whole V) and we have injectivity.

For the surjectivity the arguments are pretty similar. As we have that $f(\rho_V(g)(x)) = \rho_W(g)(f(x))$, we can conclude that for an arbitrary choice of $f(x)$ in $\operatorname{Im} f$, $\rho_W(g)$ sends $f(x)$ to somewhere else in $\operatorname{Im} f$, in particular to the image of $\rho_V(g)(x)$. So $\operatorname{Im} f$ is a subspace of W which is G-stable, and hence it is a subrepresentation. But again, as the representations where assumed to be irreducible, this must be $\operatorname{Im} f = 0$ or $\operatorname{Im} f = V$. The first one cannot be true as we assumed f nonzero, hence $\operatorname{Im} f = W$ and it is surjective.

(2) If $V = W$ and $f : V \to V$ is G-linear and V is a vector space over an algebraically closed field (we can assume $\mathbb{C}$, as the rest of the chapter is done over $\mathbb{C}$), there exists an eigenvalue $\lambda \in \mathbb{C}^*$. Let $f' = f - \lambda \mathrm{Id}$ and x be an eigenvector of eigenvalue λ. Then $f'(x) = 0$. We can also see that f' is trivially G-linear. Then, as before, $\operatorname{Ker} f'$ is G-stable, because

$$f'(\rho_V(g)(x)) = \rho_V(g)(f'(x)) = \rho_V(g)(0) = 0 \Longrightarrow \rho_V(g)(x) \in \operatorname{Ker} f'$$

and hence $\operatorname{Ker} f'$ is a subrepresentation. As V is irreducible, $\operatorname{Ker} f' = 0$ or V, but cannot be zero, as $x \in \operatorname{Ker} f'$ and we know $x \neq 0$ because it is the eigenvector of an existing eigenvalue $\lambda \in \mathbb{C}^*$. Hence, $\operatorname{Ker} f' = V$ being then $f' = 0$, i.e., $f = \lambda \mathrm{Id}$.

□

Corollary 4.1.7. *If $\varphi : G \to \mathrm{GL}_n(\mathbb{C})$ is an irreducible matrix representation and A is an $n \times n$ matrix commuting with $\varphi(g)$ for all $g \in G$, then A is a scalar matrix, i.e., $A = \lambda I$, with $\lambda \in \mathbb{C}^*$ and I is the identity $n \times n$ matrix.*

Proof. Using Schur's Lemma 4.1.6 part (2) it is very easy. As A can be seen as the matrix representation of a G-linear form of a $\mathbb{C}$-vector space V, i.e., $f : V \to V$ with matrix A, then this f must be $f = \lambda \mathrm{Id}$ and hence the matrix is also a scalar matrix. It cannot be zero matrix, because we know $\lambda \in \mathbb{C}^*$ for V being a nonzero G-linear vector space over $\mathbb{C}$, an algebraically closed field, and hence the existence of nontrivial eigenvalues is proved. □

Proposition 4.1.8. *Define the complex valued functions ω_i on $\{\mathcal{K}_1, \dots, \mathcal{K}_r\}$ for each i by*

$$\omega_i(\mathcal{K}_j) = \frac{|\mathcal{K}_j|\chi_i(g)}{\chi_i(1)}$$

where g is any element of $\mathcal{K}_j$. Then $\omega_i(\mathcal{K}_j)$ is an algebraic integer for all i and j.

Proof. This proof is quite hard so I will do it step by step. First we prove that if I is the identity matrix, then

$$\sum_{g \in \mathcal{K}_j} \varphi_i(g) = \omega_i(\mathcal{K}_j) I \tag{4.1}$$

Then, if a_{ijs} is the number of ordered pairs g_i, g_j, with $g_i \in \mathcal{K}_i$ and $g_j \in \mathcal{K}_j$ and $g_i g_j = 0$, we will prove that for all $i, j, t \in \{1, \dots, r\}$

$$\omega_t(\mathcal{K}_i)\omega_t(\mathcal{K}_j) = \sum_{s=1}^{r} a_{ijs}\omega_t(\mathcal{K}_s). \tag{4.2}$$

Finally we will see that this last equality implies that the subring of $\mathbb{Z}$ generated by $\mathbb{Z}$ and $\omega_t(\mathcal{K}_1),\dots,\omega_t(\mathcal{K}_r)$ is a finitely generated $\mathbb{Z}$-module for each $t \in \{1,\dots,r\}$ and this will imply the result, using proposition 4.1.3.

1. As we saw in the previous chapter, each $x \in G$ acting by conjugation permutes the elements of $\mathcal{K}_j$, and so $\sum_{g\in\mathcal{K}_j} \varphi_i(g)$ commutes with $\varphi_i(g)$ for all g. By Schur's Lemma 4.1.7 we have then that $\sum_{g\in\mathcal{K}_j} \varphi_i(g) = \alpha I$ for some $\alpha \in \mathcal{C}$. It remains to show that $\alpha = \omega_i(\mathcal{K}_j)$. But

   $$\operatorname{tr} \sum_{g\in\mathcal{K}_j} \varphi_i(g) = \sum_{g\in\mathcal{K}_j} \operatorname{tr}\varphi_i(g) = \sum_{g\in\mathcal{K}_j} \chi_i(g) = |\mathcal{K}_j|\chi_i(g),$$

   thus $\alpha\chi_i(1) = \operatorname{tr}\sum_{g\in\mathcal{K}_j} \varphi_i(g) = |\mathcal{K}_j|\chi_i(g)$ as needed.

2. Let g be a fixed element of $\mathcal{K}_s$ and let a_{ijs} be the number of ordered pairs g_i, g_j such that $g_i \in \mathcal{K}_i$, $g_j \in \mathcal{K}_j$ and $g_i g_j = g$. Notice that a_{ijs} is an integer. It is independent of the choice of g in $\mathcal{K}_s$ because if $x^{-1}gx$ is a conjugacy of g, then every ordered pair g_i, g_j whose product is g gives rise to another ordered pair $x^{-1}g_ix, x^{-1}g_jx$ whose product is $x^{-1}gx$ and vice versa.

   Now we prove (4.2). To see this, note that by 4.1 we have that, adding I matrix in each term of the left hand side of 4.2, we get

   $$\omega_t(\mathcal{K}_i)I\omega_t(\mathcal{K}_j)I$$

   and applying 4.1 we get then

   $$= \left(\sum_{g\in\mathcal{K}_i} \varphi_t(g)\right)\left(\sum_{g\in\mathcal{K}_j} \varphi_t(g)\right) = \sum_{g_i\in\mathcal{K}_i}\sum_{g_j\in\mathcal{K}_j} \varphi_t(g_ig_j)$$

   by multiplicativity of φ and doing some distributive. Now, taking into account a_{ijs} is the number of combinations g_i, g_j such that $g_ig_j = g$, we get that

   $$= \sum_{s=1}^{r}\sum_{g\in\mathcal{K}_s} a_{ijs}\varphi_t(g) = \sum_{s=1}^{r} a_{ijs} \sum_{g\in\mathcal{K}_s} \varphi_t(g) = \sum_{s=1}^{r} a_{ijs}\omega_t(\mathcal{K}_s)I$$

   where I applied the fact that a_{ijs} is independent of $g \in \mathcal{K}_s$ and in the last equality I applied 4.1. Then we got an equality of scalar matrices, so we compare its entries and we get then the equality 4.2.

3. Now we want to see that $\mathbb{Z}[\omega_t(\mathcal{K}_i)]$ is a finitely generated $\mathbb{Z}$-module. To see this, we will see that the set generated by $\mathbb{Z}$ and $\omega_t(\mathcal{K}_1),\dots,\omega_t(\mathcal{K}_r)$ is a finitely generated $\mathbb{Z}$-module as follows:

   If we take any element x of $\mathbb{Z}[\omega_t(\mathcal{K}_1),\dots,\omega_t(\mathcal{K}_r)]$ we want to see that this element is indeed a linear combination of the elements $\omega_t(\mathcal{K}_i)$. But it is, because of the equality 4.2, any product of $\omega_t(\mathcal{K}_i)\omega_t(\mathcal{K}_j)$ gives a linear combination of elements of this type. So we have that the subring of $\mathcal{C}$ generated by $\mathbb{Z}$ and $\omega_t(\mathcal{K}_i)$, for $i = 1,\dots,r$, generate a finitely generated $\mathbb{Z}$-module, and as $\mathbb{Z}$ is a Principal Ideal Domain, $\mathbb{Z}[\omega_t(\mathcal{K}_i)]$ is also a finitely generated $\mathbb{Z}$-module, for all $i = 1,\dots,r$.

Then, by proposition 4.1.3 we have that $\omega_t(\mathcal{K}_i)$ is an algebraic integer, as we wanted. □

Corollary 4.1.9. *The degree of each complex irreducible representations of a finite group G divides the order of G, i.e.,* $\chi_i(1) \mid |G|$ *for* $i = 1, \dots, r$.

Proof. Under the notation of the previous proposition 4.1.8 and with $g_j \in \mathcal{K}_j$ we have, where (ψ_1, ψ_2) was the hermitian product defined previously,

$$\frac{|G|}{\chi_i(1)} = \frac{|G|}{\chi_i(1)} \operatorname{gcd}(\chi_i, \chi_i) = \sum_{j=1}^{r} \frac{|\mathcal{K}_j| \chi_i(g_j) \overline{\chi_i(g_j)}}{\chi_i(1)} = \sum_{j=1}^{r} \omega_i(\mathcal{K}_j) \overline{\chi_i(g_j)}.$$

where I am using that $(\chi_i, \chi_i) = 1$ and the definition of the hermitian product we defined previously. The right hand side of this equality is an algebraic integer and the left hand side is obviously a rational number, hence is an integer. This completes the proof. □

The next two lemmas are the final results needed to complete the proof of Burnside's Theorem.

Lemma 4.1.10. *Let G any group and* $\mathcal{K}$ *a conjugacy class and* φ *an irreducible representation with character* ψ *such that* $\operatorname{gcd}(|\mathcal{K}, \chi(1)) = 1$. *Then, for* $g \in \mathcal{K}$ *either* $\chi(g) = 0$ *of* $\varphi(g)$ *is a scalar matrix.*

Proof. By hypothesis and Bézout's Identity, there exists $\lambda, \mu \in \mathbb{Z}$ such that

$$\lambda |\mathcal{K}| + \mu \chi(1) = 1$$

Multiplying every side by $\chi(g)$ we get

$$\lambda |\mathcal{K}| \chi(g) + \mu \chi(1) \chi(g) = \chi(g)$$

and now dividing by $\chi(1)$ we get

$$\lambda \frac{|\mathcal{K}| \chi(g)}{\chi(1)} + \mu \chi(g) = \frac{\chi(g)}{\chi(1)}$$

By corollary 4.1.4 $\chi(g)$ is an algebraic integer, and by proposition 4.1.8 $\frac{|\mathcal{K}|\chi(g)}{\chi(1)}$ is also an algebraic integer. Hence, $\frac{\chi(g)}{\chi(1)}$ is a linear combination of algebraic integers, i.e., it is an algebraic integer.

Let $a_1 = \frac{\chi(g)}{\chi(1)}$ and let $a_2, \dots, a_n$ be its conjugates in $\mathbb{Q}$, i.e., the other roots of the minimal polynomial of a_1 over $\mathbb{Q}$. We know by 3.2.5 that $\chi(g)$ is a sum of roots of 1, then a_1 will be the sum of $\chi(1)$ roots of 1 divided by $\chi(1)$, i.e., $\chi(g) = \xi_1 + \cdots + \xi_n$, where ξ_i is a root of 1 over $\mathbb{C}$, then

$$a_1 = \frac{\chi(g)}{\chi(1)} = \frac{\xi_1 + \cdots + \xi_n}{\chi(1)}$$

and computing the complex absolute value we get

$$|a_1| = \left| \frac{\chi(g)}{\chi(1)} \right| = \frac{|\xi_1 + \cdots + \xi_n|}{|\chi(1)|} \leq \frac{|\xi_1| + \cdots + |\xi_n|}{n} = 1$$

so $\forall i$, $|a_i| \leq 1$. Now define

$$b := \prod_{i=1}^{n} a_i \in \mathbb{Q}$$

and we see that b is an algebraic integer. Indeed, as the minimal polynomial of a_1 is

$$(x - a_1)(x - a_2) \cdots (x - a_n) = \prod_{i=1}^{n} a_i + b_1 x + b_2 x^2 + \cdots + b_n x^n$$

where b_i are the coefficients (integers) but we are not interested in them. The constant term is exactly b, so b is the product of algebraic integers and by a result regarding algebraic integers, we obtain that b is an algebraic integer. Even more, as the polynomial described above is the minimal polynomial over $\mathbb{Q}$ of a_1, but we know a_1 is an algebraic integer, is then the minimal polynomial of a_1 over $\mathbb{Z}$, hence all the coefficients are integer, and in particular $b \in \mathbb{Z}$.

Then, taking modules we get

$$|b| = \prod_{i=1}^{n} |a_i| \leq 1$$

and as $b \in \mathbb{Z}$, we obtain $b = 0$ or $b = 1$. If $b = 0$ then there must be some $i \in \{1, \ldots, n\}$ such that $a_i = 0$. Let without loss of generality, $a_1 = 0$. Then $\chi(g) = 0$, which was one option of the lemma. Take for instance $b = 1$. Then it must be $|a_i| = 1$ for all i and hence $|\chi(g)| = |\chi(1)| = \chi(1)$.

Let φ_1 be a matrix representation equivalent to φ in which is a diagonal matrix (by doing some translation). Then it is like

$$\varphi_1(g) = \begin{pmatrix} \epsilon_1 & 0 & \cdots & 0 \\ 0 & \epsilon_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \epsilon_n \end{pmatrix}$$

thus $\chi(g) = \epsilon_1 + \epsilon_2 + \cdots + \epsilon_n$. By triangle inequality, if $\epsilon_i \neq \epsilon_j$ we get $|\chi(g)| = |\epsilon_1 + \cdots + \epsilon_n| < n = \chi(1)$. But we had equality, so it must be $\epsilon_i = \epsilon$ for all $i = 1, \ldots, n$ and hence $\varphi_1 = \epsilon I$ is a scalar matrix. Since scalar matrices are similar only to themselves, it must be $\varphi(g) = \varphi_1(g) = \epsilon I$ as well. Then this gives the other option of the lemma, and completes the proof. □

Lemma 4.1.11. *If $|\mathcal{K}|$ is a power of a prime for some non identity conjugacy class $\mathcal{K}$ of G, then G is not a non-abelian simple group.*

Proof. Suppose that G is a non-abelian simple group. Let $|\mathcal{K}| = p^c$ where p is prime and $c \in \mathcal{Z}_{\geq 0}$. Then take $g \in \mathcal{K}$. If $c = 0$ then we would have $\mathcal{K} = \{g\}$ and if $g \neq 1$ (as $\mathcal{K}$ is not the identity class by hypothesis) this means that for every $x \in G$, $xgx^{-1} = g$, i.e., $g \in \mathcal{Z}(G)$ having center not trivial. But we have that always $\mathcal{Z}(G) \lhd G$ contradicting that G is simple and not abelian (if it was abelian then $\mathcal{Z}(G) = G$ and this would not contradict simplicity). Then, for the rest of the proof we will consider $c \neq 0$.

Let $\chi_1, \ldots, \chi_r$ be all the irreducible characters of G, with χ_1 the principal character (i.e., $\chi_1(g) = 1$ for all $g \in G$) and let ρ be the regular character of G[1]. Then, we decompose the regular character with the decomposition given by the First Orthogonality Relation Theorem 3.2.7 and we get

$$\rho(g) = \sum_{i=1}^{r} (\rho(g), \chi_i(g))\chi_i(g)$$

but the computation of the hermitian product goes as follows:

$$(\rho, \chi_i) = \frac{1}{|G|} \sum_{g \in G} \rho(g)\overline{\chi_i(g)} = \frac{1}{|G|}\rho(1)\overline{\chi_i(1)} = \chi_i(1)$$

for all i, because $\rho(g) = 0$ for all $g \neq 1$ and $\rho(1) = |G|$. Also because $\chi_i(1)$ is the degree of χ_i, i.e. it is an integer and its complex conjugate is itself. Then the decomposition follows

$$\rho(g) = \sum_{i=1}^{r} \chi_i(1)\chi_i(g) = 1 + \sum_{i=2}^{r} \chi_i(1)\chi_i(g)$$

as $\chi_1(g) = 1$ for all $g \in G$. Also, if we suppose $g \neq 1$ we know $\rho(g) = 0$, hence

$$1 + \sum_{i=2}^{r} \chi_i(1)\chi_i(g) = 0 \tag{4.3}$$

Now, if $p \mid \chi_j(1)$ for all $j > 1$ with $\chi_j(g) \neq 0$, then we write $\chi_j(1) = pd_j$. In this case, the equality in 4.3 becomes

$$1 + p\sum_{j=2}^{r} d_j\chi_j(g)$$

hence $\sum_{j=2}^{r} = -\frac{1}{p}$ is an algebraic integer because it is a linear combination of characters (whose are algebraic integers by 4.1.4). But this is a contradiction, as we know that $\mathbb{Z}$ is integrally closed over $\mathbb{Q}$. This proves there must exist some j such that p does not divide $\chi_j(1)$ and $\chi_j(g) \neq 0$. Then take φ the representation whose character is χ_j, then φ is faithful (because G is simple by hypothesis) and, by lemma 4.1.10, as $\gcd(p^c, \chi_j(1)) = 1$, we get that $\varphi(g)$ is a scalar matrix. A scalar matrix always commutes with all matrices, hence $\varphi(g) \in \mathcal{Z}(\varphi(G))$ and this forces $g \in \mathcal{Z}(G)$ contradicting again the simplicity of G. So either way we get a contradiction and this completes the proof. □

[1] Recall that the regular character of a group was the character of the regular representation of G. That is, take $k[G]$ as a $k[G]$-module and take a basis of this module formed by the elements of G, e.g., $\{g_1, \ldots, g_n\}$. Then for every $g \in G$, the action is defined as a permutation of the element in G, for example: $gg_i = g_j$ and then the matrix representation of g has 1 in row i and column j if $gg_j = g_i$ and zero in all the other entries. Hence, the only element $g \in G$ such that the diagonal has a nonzero element will be $g = 1$, as $gg_i = g_i$ if and only if $g = 1$ in this representation. This way the principal character is defined as follows:

$$\rho(g) = \begin{cases} 1 & \text{if } g = 1 \\ 0 & \text{otherwise} \end{cases}$$

We are now prepared to complete the proof of Burnside's Theorem. I will state again the theorem and write the proof below.

Theorem 4.1.12 (Burnside)**.** *For p and q primes, every group of order $p^a q^b$ is solvable.*

Proof. Let G be a group of order $p^a q^b$ for some primes p and q. We first discard some more trivial cases.

If $p = q$, or if either exponent is zero, then G is nilpotent hence solvable. Then let's assume this is not the case, i.e. $p \neq q$ and $a, b > 0$.

Suppose p, q, a, b be the minimum possible such that G is not solvable (i.e., for lower p, q, a, b combinations suppose G is solvable and the we do some kind of induction). If G has a proper non trivial normal subgroup N, then both N and G/N are solvable, and then G is solvable (due to the proposition 1.1.7), so this is another trivial case. Thus we may assume G is a non-abelian simple group.

Suppose G is a non-abelian simple group and let P be a p-Sylow subgroup of G. Then, there exists $g \in \mathcal{Z}(g), g \neq 1$, by consequences of Sylow groups we saw on the courses of the degree. Since P is a subgroup of $C_G(g)$, the centralizer of g, then the order of conjugacy class of g which equals to $[G : C_G(g)]$ is prime to p. As it must divide the order of $|G| = p^a q^b$ it must hence divide q, i.e., it is a power of q. But then we get a nontrivial conjugacy class $\mathcal{K}$ (the conjugacy class of g) whose order is a power of prime and hence we violate lemma 4.1.11. Hence, G must be solvable. □

Conclusions

In this project, we have worked through group representations and character theory to reach the Burnside's Theorem proof like Burnside himself stated. Firstly, we have studied in depth some properties of solvable groups, illustrating how the solvability is an important property for a group. As an example, we have seen that the extension problem could be solved easier if the group is solvable. We have also worked on new concepts such as the commutator and derived series, and we have learnt a new approach to the concept of solvable groups.

Moreover, we have dealt with the basics of representation theory. It has been established the definition of linear representation of a finite group G over a k-vector space and we have proven some properties, for instance the bijective correspondence with modules over the ring $k[G]$. This last property has given us the ability to work with modules and obtain the equivalent results for representations. Since modules were treated in curricular courses, it has made easier the study. On the other hand, we addressed some important results such as Maschke's Theorem 2.4.1 and Wedderburn Theorem 2.5.1, together with some implications. These statements are not trivial, and thereby we had to deal with non-commutative ring theory, as well as specific facts of module theory. This is the part in which I struggled the most, given that lots of concepts were new and complex.

Furthermore, essential concepts of character theory have been studied and orthogonality relations between characters lead us to prove the First Orthogonality Relation for Characters 3.2.7 which has resulted key for our main proof of Burnside's Theorem. In addition, we have seen some examples of character computation in character tables to illustrate how characters work and how to compute them.

Finally, we have worked through the proof of the main matter. Nevertheless some previous results were needed in order to get into some complicated concepts such as algebraic integers and the well-known Schur's Lemma 4.1.6.

The implementation of this work has enriched my capacities of academic mathematical research as I have been working with many references, some of them even primary references. I learnt to use the well-known data basis of AMS and EMS and also to look through a large amount of algebra books. It has not only given me the knowledge about all these matters, but it has also provided me new skills for creating a project of this dimension, working in complicated LaTeX projects and managing bibliography.

Bibliography

[1] Helmut Bender. “A group theoretic proof of Burnside’s $p^a q^b$-theorem.” In: *Mathematische Zeitschrift* 126.4 (1972), pp. 327–338.

[2] David S. Dummit and Richard M. Foote. *Abstract algebra*. 3rd ed. Hoboken: John Wiley & Sons, 2004. ISBN: 9780471433347.

[3] David M. Goldschmidt. “A group theoretic proof of the $p^a q^b$-theorem for odd primes.” In: *Mathematische Zeitschrift* 113.Math Z (1970), pp. 373–375.

[4] Thomas W. Hungerford. *Algebra*. Graduate texts in mathematics ; 73. New York: Springer-Verlag, 1974. ISBN: 0387905189.

[5] I.M. Isaacs. *Character Theory of Finite Groups*. Dover books on advanced mathematics. Dover, 1994. ISBN: 9780486680149.

[6] Hiroshi Matsuyama. “Solvability of groups of order $2^a p^b$”. In: *Osaka J. Math* 10.2 (1973), pp. 375–378.

[7] Joseph J. Rotman. *Advanced modern algebra*. 2nd ed. Upper Saddle River, N.J: Prentice Hall, 2010. ISBN: 9780821847411.

[8] Joseph J. Rotman. *An Introduction to the theory of groups*. New York, 1995.

[9] Jean Pierre Serre. *Linear representations of finite groups*. Graduate texts in mathematics ; 42. New York: Springer-Verlag, 1977.

[10] Wikipedia. *Camille Jordan - Wikipedia, The Free Encyclopedia*. Online; accessed 02/05/2022. 2021. URL: https://en.wikipedia.org/wiki/Camille_Jordan.

[11] Wikipedia. *Hans Zassenhaus - Wikipedia, The Free Encyclopedia*. Online; accessed 03/01/2022. 2012. URL: https://en.wikipedia.org/wiki/Hans_Zassenhaus.

[12] Wikipedia. *Otto Hölder - Wikipedia, The Free Encyclopedia*. Online; accessed 02/05/2022. 2022. URL: https://en.wikipedia.org/wiki/Otto_H%C3%B6lder.