Representation theory

Sophie Morel

March 26, 2019

These are the notes of a "Topics in representation theory" class I taught in Princeton University in the Fall of 2016. I tried to resist the urge to add things, but I succumbed in a few cases, most notably the proofs of the spectral theorem in chapter V and of the character formulas in chapters IV and VI. The problems are homework and final exam problems from various iterations of the class, and are given with full solutions. [1]

Here are the main sources for the notes : The abstract representation theory results of chapter I are mostly taken from Lam's book [20]. The results about representations of finite groups over a field explained in chapters II and III can be found in Serre's book [29], and the summary of the representation theory of symmetric groups in chapter IV owes almost everything to Etingof's notes [12]. The proof of the Peter-Weyl theorem in chapter V was strongly inspired by Tao's online notes [34] and [33]. Finally, chapter VI was my attempt to specialize highest weight theory to the Lie group $\mathrm{SU}(n)$ and the complex Lie algebra $\mathfrak{sl}_n(\mathbb{C})$. I am not aware of a textbook where this is done, but I used Humphreys's book [15] as reference for the Lie algebra parts (the general exposition and especially the proof of the Weyl character formula), and the introduction of Knapp's book [18] for the results about closed subgroups of $\mathrm{GL}_n(\mathbb{C})$. Also, the proof of the Baker-Campbell-Hausdorff formula in section VI.6 of chapter VI is due to Eichler (see [10]).

The books and notes just mentioned are a very good source for anybody wanting to learn more about the various topics touched upon in these notes. Let me also mention Sepanski's book [27] for the representation theory of compact Lie groups and semisimple Lie algberas, Serre's books [31] and [30] for a very different approach to many of the same topics (Lie groups, Lie algebras, and their representations), and the book [8] of Demazure-Gabriel for more about algebraic groups. Other standard (and excellent) references on algebraic groups are the books of Borel ([4]), Humphreys ([14]) and Springer ([32]). And of course, this short bibliography would not be complete without a mention of Fulton and Harris's book [13], that covers many of the same topics as these notes and contains innumerable examples and exercises.

The problems have been taken from all of these sources and many others, and I cannot claim to remember the provenance of every single one of them. [2] I will just indicate the source of a few particular problems :

- Problem VII.1.9 is theorem (3.15) of Lam's book [20].
- The problems about representations of $\mathrm{GL}_2(\mathbb{F}_q)$ (problems VII.2.6 and VII.2.8 to VII.2.15) are giving some results of section 4.1 of Bump's book [6].
- The problems about fields of definitions of representations (problems VII.2.1 and VII.2.7) are from sections 12.1 and 12.2 of Serre's book [29].
- The construction of Witt vectors in problem VII.3.4 follows closely section II.6 of Serre's book [28].

[1] There are two exceptions, both marked with (*) : problem 6.9(5) and problem 7.3.6(2). In the first case, I succumbed to laziness. In the second case, I discovered after giving the problem as an exam question that I did not know any elementary proof.

[2] If I used your work without mentioning it, I am very sorry ! Feel free to yell at me and I will correct the oversight.

- The problems about Haar measures (problems VII.5.1 to VII.5.3) are adaptations of parts of Tao's blog entry [33].
- Most of the examples in problem VII.5.4 come from the examples and exercises in sections 1.1 and 1.2 of Sepanski's book [27].
- The starting point of the problems about linear algebraic groups (problems VII.6.5, VII.6.6, VII.6.18, VII.6.17 and VII.7.1) is example (v) in chapter I of Serre's book [31].
- Problem VII.6.16 is an example of the theory of sections II.4.5- II.4.6 of Demazure and Gabriel's book [8], though to be honest I adapted it from a similar problem about Lie groups, i.e. problems 11-13 of chapter III of Knapp's book [18].
- The problem on pseudo-characters (problem VII.7.3) is extracted from Bellaïche's notes [2] and also Dotsenko's notes [9].
- Finally, the problem on Schur-Weyl duality (problem VII.7.4) is extracted from sections 4.18-4.21 of Etingof's notes [12], and the problem on the algebraic Peter-Weyl theorem (problem VII.7.1) also owes a lot to these notes.

The formal prerequisites for the class were the two undergraduate algebra classes at Princeton (covering, among other things, the basic theory of groups, rings, and modules, and some Galois theory), but some knowledge of measure theory and Hilbert spaces was also necessary in chapter V. Here are the prerequisites chapter by chapter :

- Chapter I assumes familiarity with groups, commutative rings and modules over them, and also with tensor products (there is a review problem on that last point, see problem VII.1.1)
- Chapter II assumes that the reader is familiar with chapter I and its prerequisites.
- Chapter III assumes familiarity with chapters I and II.
- Chapter IV can be read after chapter I and sections 1-3 of chapter II.
- Chapter V is formally independent of the first four chapters, but it does assume that the reader is familiar with the basic representation theory of finite groups in characteristic 0 (section 3 of chapter I and sections 1-3 of chapter II). It also requires knowledge of measure theory (up to the Riesz representation theorem) and of Hermitian inner product spaces and Hilbert spaces.
- Chapter VI is also mostly independent of the other chapters, but it depends on chapter V via corollary VI.8.4. It also assumes some familiarity with modules over noncommutative rings and tensor products over fields.

I would like to thank all the students who took the class (Alexandre De Faveri, Timothy Ratigan, Alex Song, Roger Van Peski, Joshua Wang, Xiaoyu Xu, Murilo Zanarella and Roy Zhao) and also my graduate assistant Fabian Gundlach for being an excellent audience, for asking stimulating questions and for pointing out many mistakes in the lectures and in the problem sets.

Conventions : Unless otherwise specified, when we write $\sum_{i\in I} x_i$ (in some abelian group), we will always be assuming that all but a finite number of the x_i are 0, so that the sum is a finite sum.

Also, $\mathbb{N}$ is the set of nonnegative integers, and we admit the axiom of choice.

$M_{nm}(R)$ is the set of $n \times m$ matrices with coefficients in R, $M_n(R) = M_{nn}(R)$, $\mathrm{GL}_n(R) = M_n(R)^\times$.

${}^tA$ is the transpose of a matrix.

Contents

I Abstract representation theory **11**

I.1 Semisimple rings . . . 11
I.1.1 Definition and examples . . . 11
I.1.2 Zorn's lemma . . . 14
I.1.3 Semisimple modules and rings . . . 15
I.1.4 Schur's lemma . . . 20
I.1.5 Jordan-Hölder theorem . . . 21
I.1.6 Artinian and Noetherian modules . . . 23
I.1.7 Isotypic decomposition . . . 24
I.1.8 Simple rings . . . 26
I.1.9 Double centralizer property . . . 29
I.1.10 Structure of semisimple rings (Artin-Wedderburn theorem) . . . 31
I.2 Jacobson radical . . . 33
I.3 Applications to the representation theory of finite groups . . . 36
I.4 The representation ring . . . 42
I.5 Induction and restriction . . . 45
I.5.1 Definitions . . . 45
I.5.2 Induction and exact sequences . . . 47
I.5.3 Coinduction and exact sequences . . . 49
I.5.4 Frobenius reciprocity . . . 50
I.5.5 Comparing induction and coinduction . . . 52
I.5.6 The projection formula . . . 53
I.5.7 The case of finite groups . . . 54

II Characteristic 0 theory **55**

II.1 Characters . . . 55
II.1.1 Definition . . . 55
II.1.2 Orthogonality of characters . . . 57
II.1.3 Characters and representation ring . . . 60
II.1.4 The case $k = \mathbb{C}$ . . . 61
II.2 Representations of a product of groups . . . 62
II.3 Characters and induced representation . . . 63
II.3.1 Character of an induced representation . . . 63
II.3.2 Frobenius reciprocity with characters . . . 64
II.3.3 Mackey's formula . . . 65

II.3.4 Mackey's irreducibility criterion 66
II.4 Artin's theorem 67
II.5 Brauer's theorem 69
II.6 First application of Brauer's theorem : field of definition of a representation of G 75

III Comparison between characteristic 0 theory and characteristic p theory 77
III.1 Indecomposable modules 77
III.1.1 Definitions 77
III.1.2 Noncommutative local rings 78
III.1.3 Fitting's lemma 80
III.1.4 Krull-Schmidt-Remak theorem 80
III.1.5 Projective indecomposable modules 81
III.1.6 Lifting of idempotents 83
III.2 Application to representation rings in positive characteristic 84
III.3 Representations over discrete valuation rings 85
III.4 The cde triangle 87
III.5 Representations over a field of characteristic $p \not| |G|$ 88
III.6 Brauer's theorem in positive characteristic 89
III.7 Surjectivity of d 90
III.8 Injectivity of c 91
III.9 Image of the map e 92

IV Irreducible representations of the symmetric group $\mathfrak{S}_n$ over $\mathbb{C}$ 95
IV.1 Partitions 95
IV.2 Young tableaux and Young projectors 96
IV.3 Partitions and irreducible representations 99
IV.4 Characters of the irreducible representations V_λ 102

V Representations of compact groups 109
V.1 Topological groups, Haar measures, representations 109
V.2 Finite-dimensional representations 111
V.3 Unitary representations 112
V.3.1 Definition and first properties 112
V.3.2 The operators $T^0_{v,w}$ 114
V.3.3 Schur orthogonality 115
V.4 The space $L^2(G)$ 117
V.4.1 Definition and actions of G 117
V.4.2 The convolution product 119
V.4.3 The regular representations 122
V.5 The Peter-Weyl theorem 123
V.6 The spectral theorem 127

VI Representations of Lie algebras : the case of $\mathfrak{sl}_n(\mathbb{C})$ **131**
VI.1 Definitions . . . 131
VI.2 Universal enveloping algebra . . . 134
VI.2.1 The tensor algebra of a k-module . . . 134
VI.2.2 The universal enveloping algebra of a Lie algebra . . . 135
VI.3 The matrix exponential . . . 136
VI.4 The Lie algebra of a closed subgroup of $\mathrm{GL}_n(\mathbb{C})$ . . . 140
VI.5 From groups representations to Lie algebra representations . . . 144
VI.6 The Baker-Campbell-Hausdorff formula . . . 148
VI.7 Representations of $\mathfrak{sl}_2(\mathbb{C})$ . . . 151
VI.8 Lifting representations of $\mathfrak{su}(n)$ . . . 152
VI.9 Some irreducible representations of $\mathrm{SU}(n)$ . . . 155
VI.9.1 The exterior algebra . . . 155
VI.9.2 Exterior power representations . . . 157
VI.10 Characters . . . 158
VI.11 Weights . . . 160
VI.12 More about roots and weights . . . 162
VI.12.1 Weights of infinite-dimensional representations . . . 162
VI.12.2 Roots . . . 163
VI.13 The Weyl character formula . . . 163
VI.14 Proof of the Weyl character formula . . . 164
VI.14.1 Highest weights . . . 164
VI.14.2 The Poincaré-Birkhoff-Witt theorem and the Casimir element . . . 164
VI.14.3 Verma modules . . . 166
VI.14.4 Characters of Verma modules . . . 167
VI.14.5 Jordan-Hölder series of Verma modules . . . 169
VI.14.6 End of the proof of the Weyl character formula . . . 170

VII Exercises **173**
VII.1 Chapter I exercises . . . 173
VII.1.1 Review of tensor products . . . 173
VII.1.2 Some properties of projective modules . . . 176
VII.1.3 Division rings . . . 177
VII.1.4 Ideals of rings of matrices . . . 178
VII.1.5 Commutative semisimple rings . . . 180
VII.1.6 The $\mathbb{R}$-algebra of quaternions . . . 181
VII.1.7 Simple modules over some commutative rings . . . 182
VII.1.8 Group algebra of the quaternion group . . . 183
VII.1.9 A simple ring that is not semisimple . . . 185
VII.1.10 Central simple algebras and the Brauer group . . . 188
VII.1.11 Irreducible representations of p-groups in characteristic p . . . 196
VII.1.12 Another description of induction . . . 197
VII.1.13 Representation ring of $\mathbb{Z}/p^r\mathbb{Z}$ in characteristic p . . . 199

VII.1.14 Basic properties of induction . . . 200
VII.2 Chapter II exercises . . . 201
VII.2.1 Representation rings and field extensions . . . 201
VII.2.2 Some character tables . . . 203
VII.2.3 Calculating representation rings . . . 204
VII.2.4 Representations of products . . . 205
VII.2.5 Character of small symmetric and exterior powers . . . 206
VII.2.6 Using Mackey's irreducibility criterion (representations of $\mathrm{GL}_2(\mathbb{F}_q)$, part 1) . . . 208
VII.2.7 Rationality problems . . . 210
VII.2.8 Hecke algebra . . . 213
VII.2.9 Multiplicity-free modules . . . 215
VII.2.10 Hecke algebra and multiplicities . . . 216
VII.2.11 Characters of a finite field . . . 217
VII.2.12 Representations of $\mathrm{GL}_2(\mathbb{F}_q)$, part 2 . . . 218
VII.2.13 Representations of $\mathrm{GL}_2(\mathbb{F}_q)$, part 3 . . . 221
VII.2.14 Representations of $\mathrm{GL}_2(\mathbb{F}_q)$, part 4 . . . 224
VII.2.15 Representations of $\mathrm{GL}_2(\mathbb{F}_q)$, part 5 . . . 226
VII.2.16 Induction and characters . . . 234
VII.3 Chapter III exercises . . . 235
VII.3.1 Discrete valuation rings . . . 235
VII.3.2 Discrete valuation fields . . . 236
VII.3.3 Completion of a discrete valuation ring . . . 237
VII.3.4 Witt vectors . . . 241
VII.4 Chapter IV exercises . . . 249
VII.5 Chapter V exercises . . . 249
VII.5.1 Existence of the Haar measure on a compact group . . . 249
VII.5.2 Haar measures are unique . . . 253
VII.5.3 Unimodular groups . . . 255
VII.5.4 Some examples of topological groups . . . 257
VII.5.5 Representations of compact commutative groups . . . 273
VII.5.6 Complex representations of profinite groups . . . 274
VII.5.7 A unitary representation of G such that $G \to \mathrm{U}(V)$ is not continuous . . . 275
VII.5.8 A compact group with no faihful representation . . . 276
VII.5.9 Uniqueness of the inner product making an irreducible representation unitary . . . 276
VII.6 Chapter VI exercises . . . 278
VII.6.1 Surjectivity of the exponential map . . . 278
VII.6.2 Kernel of the adjoint representation . . . 282
VII.6.3 Lie algebras of compact groups . . . 282
VII.6.4 Some Lie algebras, and the adjoint representation . . . 283
VII.6.5 Lie algebra of a linear algebraic group . . . 283
VII.6.6 Group of automorphisms of a k-algebra . . . 286

VII.6.7 Symmetric algebra and symmetric powers of a representation 288
VII.6.8 Symmetric algebra and polynomial functions 290
VII.6.9 Some representations of $\mathfrak{sl}_2(k)$ 291
VII.6.10 Representations of $\mathfrak{sl}_2(k)$ in characteristic 0 292
VII.6.11 The Jacobson-Morozov theorem (for $\mathfrak{gl}_n(k)$) 294
VII.6.12 Clebsch-Gordan decomposition 295
VII.6.13 Dual representation 296
VII.6.14 Some representations of $\mathfrak{sl}_n(k)$ 296
VII.6.15 A generating family for the universal enveloping algebra 298
VII.6.16 Universal enveloping algebra and differential operators (and a proof of the Poincaré-Birkhoff-Witt theorem for $\mathfrak{gl}_n(k)$ if $\mathrm{char}(k) = 0$) 299
VII.6.17 Regular functions on an algebraic group 306
VII.6.18 Differentiating morphisms of algebraic groups 308
VII.6.19 Semisimple representations of the Lie algebra $\mathfrak{gl}_n(\mathbb{C})$ 311
VII.6.20 Differential of a tensor product of representations 317
VII.6.21 Casimir element 318
VII.7 Exercises involving several chapters 319
VII.7.1 The algebraic Peter-Weyl theorem (chapters V and VI) 319
VII.7.2 Polarization 323
VII.7.3 Pseudo-characters (chapters I and II) 324
VII.7.4 Schur-Weyl duality (chapters I, IV and VI) 339

I Abstract representation theory

I.1 Semisimple rings

I.1.1 Definition and examples

In these notes, a ring will be a non necessarily commutative ring with unit.

Example I.1.1.1.

- $R = \{0\}$, the only ring where $0 = 1$.
- Some commutative rings : $\mathbb{Z}$, $A[T_1, \ldots, T_n]$ (polynomials in n indeterminates over a commutative ring A) or $A[T_i, i \in I]$ (polynomials in a set of indeterminates indexed by I).
- Some noncommutative rings :
  - $M_n(R)$ ($n \times n$ matrices with coefficients in a ring R).
  - The *group algebra* $R[G]$ of a group G with coefficients in a (not necessarily commutative) ring R. As a R-module, this is the free R-module with basis G, that is, $R[G] = \bigoplus_{g \in G} R \cdot g$. The multiplication is induced by that of G, i.e. given by :

$$\left(\sum_{g \in G} \alpha_g g\right)\left(\sum_{g \in G} \beta_g g\right) = \sum_{g \in G}\left(\sum_{h_1 h_2 = g} \alpha_{h_1}\beta_{h_2}\right) g.$$

    Note that this definition also makes sense if G is a monoid.
  - The *free A-algebra* over a set X, where A is a commutative ring, is the A-algebra $A\langle X\rangle$ of noncommutative polynomials with indeterminates in X. In other words, it's the algebra of the monoid M_X whose elements are words on the elements of X and whose multiplication is concatenation. (This M_X is called the *free monoid* on X.)

Definition I.1.1.2. Let R be a ring. A left (resp. right) R-module is a commutative group M with a biadditive map $R \times M \to M$, $(a, x) \longmapsto ax$ (resp. $M \times R \to M$, $(x, a) \longmapsto xa$), such that :

- $1x = x$ (resp. $x1 = x$), $\forall x \in M$.
- $(ab)x = a(bx)$ (resp. $x(ab) = (xa)b$), $\forall a, b \in R$, $\forall x \in M$.

If we want to make it very clear that M is a left (resp. right) R-module, we write ${}_RM$ (resp. M_R) instead of M.

By convention, a R-module will be a left R-module unless otherwise specified.

Remark I.1.1.3. If the ring R is commutative, then the notions of left and right R-module coincide.

Example I.1.1.4. The ring R with left (resp. right) multiplication by itself is a left (resp. right) R-module, called the left (resp. right) *regular R-module* and sometimes denoted by ${}_RR$ (resp. R_R).

Let's define a few notions for left R-modules. We obviously have similarly defined notions for R-modules.

Definition I.1.1.5. Let M be a R-module. A *R-submodule* (or *submodule* if R is clear) of M is a subgroup N of M such that $ax \in N$ for every $a \in R$ and $x \in N$.

Example I.1.1.6. A submodule of $M = {}_RR$ is just a left ideal of R.

Definition I.1.1.7. If M is a R-module and N is a submodule of M, then the quotient group M/N has a structure of R-module given by $a(x+N) = ax + N$ for $a \in R$ and $x \in M$. This is called a *quotient R-module*.

Definition I.1.1.8. Let M be a R-module and $(M_i)_{i\in I}$ be a family of submodules of M.

We say that M is the *sum* of the M_i and write $M = \sum_{i\in I} M_i$ if, for every $x \in M$, there exist $x_i \in M$ such that $x = \sum_{i\in I} x_i$.[1]

We say that M is the *direct sum* of the M_i and write $M = \bigoplus_{i\in I} M_i$ if, for every $x \in M$, there exist *uniquely determined* $x_i \in M_i$ such that $x = \sum_{i\in I} x_i$.

Definition I.1.1.9. Let M, N be R-modules. A *R-linear map* (or *R-module morphism*) from M to N is a morphism of abelian groups $\varphi/M \to N$ such that $\varphi(ax) = a\varphi(x)$ for every $a \in R$ and $x \in M$.

Definition I.1.1.10. An *exact sequence* of R-modules is an exact sequence of abelian groups where all the abelian groups are R-modules and all the maps are R-linear.

Example I.1.1.11. Let $\varphi : M \to N$ be a R-linear map. Then $\operatorname{Ker}\varphi \subset M$ and $\operatorname{Im}\varphi \subset N$ are submodules, and we have an exact sequence of R-modules :

$$0 \to \operatorname{Ker}\varphi \to M \to N \to N/\operatorname{Im}\varphi \to 0.$$

Notation I.1.1.12. Let M, N be two R-modules. Then we write $\operatorname{Hom}_R(M,N)$ for the abelian group of R-linear maps from M to N. We also write $\operatorname{End}_R(M)$ for $\operatorname{Hom}_R(M,M)$; this is a ring, and its group of invertible elements (see definition I.1.1.16) will be denoted by $\operatorname{Aut}_R(M)$.

[1] By our general convention, all but a finite number of the x_i must be 0.

Examples of modules I.1.1.13.

- *A $\mathbb{Z}$-module is just an abelian group.*
- *R^n, seen as the set of $n \times 1$ matrices with coefficients in R, is a left $M_n(R)$-module with the operation given by matrix multiplication. If we see R^n as the set of $1 \times n$ matrices with coefficients in R, we similarly get a right R-module structure on it.*
- *Let A be a commutative ring. Then a $A[T]$-module is a A-module M with a A-linear endomorphism (the action of $T \in A[T]$).*
- *More generally, if A is a commutative ring and I is a set, then a $A[T_i, i \in I]$-module is a A-module M with a family $(u_i)_{i \in I}$ of pairwise commuting A-linear endomorphisms. (The endomorphism u_i is given by the action of T_i on M.)*
- *If A is a commutative ring and X is a set, then a $A\langle X\rangle$-module is a A-module M with a family $(u_x)_{x \in X}$ of A-linear endomorphisms. (They are not required to commute with each other anymore.)*
- *If A is a commutative ring and G is a group (or just a monoid), then a $A[G]$-module is a A-module with a morphism of groups (or monoids) $G \to \mathrm{Aut}_A(M)$. This is also called a A-linear representation of the group (or monoid) G on the A-module M.*

Definition I.1.1.14 (Ideals)**.** Remember that a left ideal of R is a left submodule of ${}_RR$, and a right ideal of R is a right submodule of R_R.

An *ideal* of R is a subset I of R that is both a left ideal and a right ideal.[2] Then the quotient abelian group R/I is also a ring.

Example I.1.1.15. Take $R = \mathbb{C}[x, \sigma]$, the ring of twisted polynomials over $\mathbb{C}$ in one indeterminate x. Here σ is the endomorphism of $\mathbb{C}$ given by the complex conjugation, and the indeterminate x does not commute with the elements of $\mathbb{C}$: if $a \in \mathbb{C}$ and $r \in \mathbb{N}$, we ask that $x^r a = \sigma^a(a)x$.

Then $I_1 = R(x^2 - 1)$ and $I_2 = R(x - i)$ are both left ideals of R. The first one, I_1, is also a right ideal (so it's an ideal) because $x^2 - 1$ is in the center of R (= commutes with every element of R). But I_2 is not an ideal, because :

$$(x - i)i = xi + 1 = -ix + 1 = -i(x + i) \notin I_2.$$

Definition I.1.1.16. An element $a \in R$ is called *left invertible* (resp. *right invertible*) if there exists $b \in R$ such that $ba = 1$ (resp. $ab = 1$). In that case, b is called a *left inverse* (resp. *right inverse*) of a.

If a is both left and right invertible, we say that it is *invertible*. In that case, if b, b' are elements of R such that $ab = b'a = 1$, then we have $b = b'$ and b is the unique element of R such that $ab = 1$ (or $ba = 1$); we say that b is the *inverse* of a and write $b = a^{-1}$.

[2] Unfortunately, this is not coherent with the convention that R-modules are left R-modules.

We write $R^\times$ for the set of invertible elements of R; it's a group, with the group law given by the multiplication of R.

If $R \neq \{0\}$ and $R^\times = R - \{0\}$, we call R a *divison ring*. Note that a commutative division ring is just a field.

Example I.1.1.17.

- The ring $\mathbb{H}$ of quaternions (see problem VII.1.6) is a division ring.
- Let $V = \bigoplus_{i\in\mathbb{N}} \mathbb{Q}e_i$ (a $\mathbb{Q}$-vector space with a basis indexed by $\mathbb{N}$) and $R = \mathrm{End}_{\mathbb{Q}}(V)$. Define $u, v, w \in R$ by :

$$u(e_i) = e_{i+1} \qquad \forall i \in \mathbb{N}$$

$$v(e_i) = \begin{cases} 0 & \text{if } i > 0 \\ e_{i-1} & \text{if } i = 0 \end{cases}$$

$$w(e_i) = \begin{cases} e_0 & \text{if } i = 0 \\ 0 & \text{if } i > 0 \end{cases}$$

  Then $vu = 1$ and $vw = 0$, so v is right-invertible but not left-invertible. (If we had $u'v = 1$ with $u' \in R$, then we would have $w = (u'v)w = u'(vw) = 0$, which is not true.)

I.1.2 Zorn's lemma

The goal of this section is to state Zorn's and show a typical use in algebra.

Theorem I.1.2.1. *The axiom of choice and Zorn's lemma are equivalent, where Zorn's lemma is the following statement :* [3]

Zorn's lemma I.1.2.2. *Let X be a* nonempty *partially ordered set. Suppose that for every $Y \subset X$ that is* totally ordered*, there exists an upper bound of Y, that is, there exists $x \in X$ such that $y \leq x$ for every $y \in Y$.*

Then X has a maximal element, that is, there exists $x \in X$ such that no other $x' \in X$ is strictly bigger than x. (In other words, every element of $X - \{x\}$ is either smaller than x or not comparable to x.)

When applying Zorn's lemma, it is very important not to forget to check that the set X is nonempty.

Example of application I.1.2.3. *Let R be a ring. Then R has maximal ideals.*

[3] See theorem 5.4 of Jech's book [16].

Proof. Let X be the set of ideals of R, ordered by inclusion. Then X is not empty because $\{0\} \in X$. If Y is a totally ordered subset of X, then $J = \bigcup_{I \in Y} I$ is an ideal of R (because Y is totally ordered),[4] and it's obviously an upper bound of Y. So X has a maximal element by Zorn's lemma.

□

I.1.3 Semisimple modules and rings

Definition I.1.3.1. Let R be a ring and M be a R-module.

1. We say that M is *simple* (or *irreducible*) if $M \neq 0$ and if the only R-submodules of M are 0 and M.
2. We say that M is *semisimple* (or *completely reducible*) if, for every submodule N of M, there exists a submodule N' of M such that $M = N \oplus N'$. (In other words, if every short exact sequence $0 \to N \to M \to M/N \to 0$ of R-modules splits.)

Remark I.1.3.2.

- A simple R-module is semisimple.
- The R-module 0 is semisimple but not simple.
- If M is semisimple, then every submodule and quotient module of M is also semisimple.
- If R is a field (or a division ring), then every R-module is semisimple.

Example I.1.3.3.
- If $n \geq 1$, then R^n is a simple $M_n(R)$-module (whether it is seen as a left or right module).
- Let K be a field and M be a $K[T]$-module such that $\dim_K(M) < +\infty$. Then M is a semisimple $K[T]$-module if and only if the endomorphism of M given by the action of T is semisimple (= diagonalizable over an algebraic closure of K).

Theorem I.1.3.4. *Let R and M be as above. The following are equivalent :*

1. *M is semisimple.*
2. *M is the direct sum of a family of simple submodules.*
3. *M is the sum of a family of simple submodules.*

Lemma I.1.3.5. *If $M \neq 0$ and M is semisimple, then it has a simple submodule.*

Proof. Let $x \in M - \{0\}$, and set $M' = Rx$. Then M' is semisimple and nonzero, so we may assume that $M = M'$.

[4] If $x, x' \in J$ and $a \in R$, then we can find $I, I' \in Y$ such that $x \in I$ and $x' \in I'$. As Y is totally ordered, we have $I \subset I'$ or $I' \subset I$. In the first case, $ax + x' \in I \subset J$, and in the second, $ax + x' \in I' \subset J'$.

We want to apply Zorn's lemma to the set X of submodules N of M such that $x \notin M$, ordered by inclusion. This set is nonempty because it contains the zero submodule. If $Y \subset X$ is totally ordered, then $N = \bigcup_{N' \in Y} N'$ is a submodule of M. (The proof is the same as in example I.1.2.3.) Also, $x \notin N$ by definition of N, so N is an element of X, and an upper bound of Y.

By Zorn's lemma, X has a maximal element N. As $x \notin N$, $N \neq M$. As M is semisimple, there exists a submodule N' of M such that $M = N \oplus N'$, and we have $N' \neq 0$ because $N \neq M$.

I claim that N' is simple. Indeed, if there were a submodule $0 \neq N'' \subsetneq N'$ of N, then we would have $N \oplus N'' \notin X$ by maximality of N in X, so $x \in N \oplus N''$, but then $N \oplus N'' = M$ (because $M = Rx$), and hence $N'' = N'$, contradiction.

□

Proof of the theorem.

(ii)⇒(iii) Direct sums are sums.

(i)⇒(iii) Assume that M is semisimple. Let M' be the sum of all the simple submodules of M, and choose a submodule M'' of M such that $M = M' \oplus M''$. If $M' \neq M$, then $M'' \neq 0$, so by the lemma M'' has a simple submodule N, but then we should have $N \subset M'$, which contradicts the fact that M' and M'' are in direct sum. So $M' = M$.

(iii)⇒(i)&(ii) Let $(M_i)_{i \in I}$ be the family of all simple submodules of M. We are assuming that $M = \sum_{i \in I} M_i$. Let N be a submodule of M. We will show that there exists $J \subset I$ such that

$$M = N \oplus \bigoplus_{j \in J} M_j.$$

This clearly implies (i), and we also get (ii) by taking $N = 0$.

To get this J, we want to apply Zorn's lemma to the set X of subsets of K of I such that the sum $N + \sum_{k \in K} M_k$ is direct, ordered by inclusion. This set X is not empty, because $\varnothing \in X$.

If $Y \subset X$ is a totally ordered subset, let $K = \bigcup_{K' \in Y} K'$ and let's show that $K \in X$ (and hence is an upper bound of Y). Let $n \in N$ and $(m_k)_{k \in K} \in \prod_{k \in K} M_k$ such that $n + \sum_{k \in K} m_k = 0$. Let $K_0 \subset K$ be a finite subset such that $m_k = 0$ for $k \in K - K_0$ (this exists by our convention at the beginning). For every $k \in K_0$, there exists $L_k \in Y$ such that $k \in L_k$. As Y is totally ordered and K_0 is finite, there exists $L \in Y$ such that $K \supset \bigcup_{k \in K_0} L_k$. Then $L \in X$, so $n + \sum_{k \in K_0} m_k = 0$ implies that $n = 0$ and $m_k = 0$ for every $k \in K_0$. By the choice of K_0, we get that n and all the m_k, $k \in K$, are 0. So $K \in X$.

By Zorn's lemma, X has a maximal element J. Let $M' = N \oplus \bigoplus_{j \in J} M_j$. We want to show that $M = M'$, so let's show that $M' \supset M_i$ for every $i \in I$. Let $i \in I$. If $M' \not\supset M_i$,

then $M' \cap M_i = 0$ (because M_i is simple). But then the sum

$$M' + M_i = (N + \sum_{j \in J} M_j) + M_i$$

is direct, so $J \cup \{i\} \in X$, and this contradicts the maximality of J.

□

Definition I.1.3.6. Let R be a ring and M be a R-module. We say that M is *finitely generated* if there exists a *finite* family $(x_i)_{i \in I}$ of M such that $M = \sum_{i \in I} Rx_i$. We say that M is *cyclic* if there exists $x \in M$ such that $M = Rx$. [5]

Theorem-Definition I.1.3.7. *Let R be a ring. The following are equivalent :*

1. *All short exact sequences of R-modules split.*
2. *All R-modules are semisimple.*
3. *All finitely generated R-modules are semisimple.*
4. *All cyclic R-modules are semisimple.*
5. *The left regular R-module ${}_RR$ is semisimple.*

If these conditions are satisfied, then we say that the ring R is semisimple.

Remark I.1.3.8. A priori, this notion should be called "left semisimple ring", and we should have a similarly defined notion of "right semisimple ring". But we will see in section I.1.10 that "left semisimple" and "right semisimple" are actually equivalent.

Proof.

(i)⇔(ii) This follows directly from the definition of semisimple modules.

(ii)⇒(iii)⇒(iv)⇒(v) is obvious.

(v)⇒(ii) Let M be a R-module. If $x \in M$, then $Rx \subset M$ is a quotient of ${}_RR$, so it's semisimple because ${}_RR$ is semisimple. As $M = \sum_{x \in M} Rx$, theorem I.1.3.4 implies that M is semisimple.

□

Example I.1.3.9.

- Division rings (and in particular fields) are semisimple, because their only left ideals are 0 and the whole ring.
- $\mathbb{Z}$ is not semisimple, because $2\mathbb{Z} \subset \mathbb{Z}$ is a not a direct factor of $\mathbb{Z}$.

[5] A cyclic left (resp. right) ideal of R is also called a principal left (resp. right) ideal.

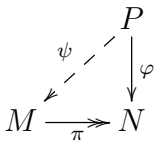
- If R is a division ring and $n \in \mathbb{N}$, then $M_n(R)$ is a semisimple ring.

  Proof. As a left $M_n(R)$-module, $M_n(R)$ is isomorphic to a direct sum of n factors R^n, and we already know that R^n is a simple $M_n(R)$-module.

  □

We will now see another characterization of semisimple rings.

Definition I.1.3.10. A R-module P is called *projective* if, for every surjective R-linear map $\pi : M \twoheadrightarrow N$ and every R-linear map $\varphi : P \to N$, there exists a R-linear map $\psi : P \to M$ such that $\pi\psi = \varphi$.



Also, a R-module F is called *free* if there exists a family $(e_i)_{i\in I}$ of elements of F such that $F = \bigoplus_{i\in I} Re_i$ and such that, for every $i \in I$, the map $R \to Re_i$, $a \longmapsto ae_i$, is an isomorphism.

Example I.1.3.11.

- Every free R-module is projective.

  Proof. If $F = \bigoplus_{i\in I} Re_i$, let $\pi : M \twoheadrightarrow N$ be a surjective R-linear map and $\varphi : P \to N$ be a R-linear map. For every $i \in I$, choose $x_i \in M$ such that $\pi(x_i) = \varphi(e_i)$. Define a R-linear map $\psi : P \to M$ by $\psi(e_i) = x_i$ for every $i \in I$. Then it's easy to check that $\pi\psi = \varphi$.

  □

- A direct summand of a projective module is projective.

  Proof. Let P be projective, and suppose that $P = P' \oplus P''$. Let's show that P' is projective. Let $\pi : M \twoheadrightarrow N$ be a surjective R-linear map and $\varphi' : P' \to N$ be a R-linear map. Let $\varphi = \varphi' + 0 : P = P' \oplus P'' \to N$. As P is projective, there exists a map $\psi : P \to M$ such that $\pi\psi = \varphi$. Write $\psi = \psi' + \psi''$, where ψ' (resp. ψ'') is a map $P' \to M$ (resp. $P'' \to M$). Then $\pi\psi' = \varphi'$.

  □

- A R-module P is projective if and only every exact sequence of R-modules $0 \to M' \to M \to P \to 0$ splits.

  Proof. If P is projective, then every such exact sequence splits by definition. Conversely, if every such exact sequence splits, then applying this to the exact sequence $0 \to K \to F \to P \to 0$ where $F = \bigoplus_{x\in P} R$, the map $F \to P$ sends $(a_x)_{x\in P}$ to

$\sum_{x\in P} a_x x$ and $K \to F$ is the kernel of this map, we see that P is a direct summand of the free R-module F, hence is projective.

□

Lemma I.1.3.12. *Let P be a R-module. The following are equivalent :*

1. *P is projective.*
2. *P is a direct summand of a free R-module.*

Proof. We saw that (ii) implies (i) in example I.1.3.11. Let's show that (i) implies (ii). Suppose that P is projective. Let $M = \bigoplus_{x\in P} R$, and define $\varphi : M \to P$ by $\varphi((a_x)_{x\in P}) = \sum_{x\in P} a_x x$. Then φ is R-linear and surjective, so there exists a R-linear map $\psi : P \to M$ such that $\varphi\psi = \mathrm{id}_P$. So we have $M \simeq P \oplus \mathrm{Ker}(\varphi)$, and we have found a free R-module M such that P is a direct summand of M.

□

Example I.1.3.13.
- By the structure theorem for finitely generated $\mathbb{Z}$-modules, any finitely generated projective $\mathbb{Z}$-module is free. In fact, *any* projective $\mathbb{Z}$-module is free.[6]
- In the previous example, we could replace $\mathbb{Z}$ by any principal ideal domain.
- If R is a Dedekind ring (for example a principal ideal domain or the ring of integers in a number field), any fractional ideal of R is a projective R-module.
- Over $R := \mathbb{Z} \times \mathbb{Z}$, $P := \mathbb{Z} \times \{0\}$ is projective (because $P \oplus (\{0\} \times \mathbb{Z}) = R$) but not free.

Theorem I.1.3.14. *Let R be a ring. The following are equivalent :*

1. *R is a semisimple ring.*
2. *All R-modules are projective.*
3. *All finitely generated R-modules are projective.*
4. *All cyclic R-modules are projective.*

Proof. (i)⇔(ii) follows from the definition, and (ii) ⇒(iii)⇒(iv) is obvious.

(iv)⇒(i) Suppose that all cyclic R-modules are projective. Let's show that ${}_R R$ is a semisimple R-module. Let $I \subset R$ be a left ideal, then R/I is a cyclic R-module, hence projective, and so the short exact sequence $0 \to I \to R \to R/I \to 0$ splits, that is, $R \simeq I \oplus R/I$.

□

[6] See theorem 3 and its corollary in Kaplansky's paper [17].

Remark I.1.3.15. We have a dual notion of *injective R-module* (a R-module I is injective if for every R-linear injective map $\varphi : I \to M$, there exists a R-linear map $\psi : M \to I$ such that $\psi\varphi = \mathrm{id}_I$). The analogue of theorem I.1.3.14 is still true but harder to prove. [7]

Remark I.1.3.16.
- A left ideal I of R is a simple R-module if and only if it is minimal among nonzero left ideals of R. So if R is a semisimple ring, it has minimal nonzero left ideals.

- Assume that I is a left ideal of R, and that the short exact sequence $0 \to I \to {}_RR \to R/I \to 0$ of R-modules splits, that is, that there exists a left ideal J of R such that $R = I \oplus J$. Write $1 = e + e'$, with $e \in I$ and $e' \in J$. Then $e^2 = e$, $e'^2 = e'$, and $ee' = e'e = 0$. (We say that e and e' are *orthogonal idempotents* of R.) Also, $I = Re$ and $J = Re'$.

  If moreover I and J are ideals, then $IJ = 0$, e and e' are central in R, I and J and rings with respective units e and e', and $R = I \times J$ as rings.

  Proof. We have $e = e(e + e') = e^2 + ee' = e^2$ with $e, e^2 \in I$ and $ee' \in J$, so $e = e^2$ and $ee' = 0$. Similarly, $e'^2 = e'$ and $e'e = 0$. Obviously, $Re \subset I$; conversely, if $x \in I$, then $x = x(e + e') = xe + xe'$ with $xe \in I$ and $xe' \in J$, so $x = xe \in Re$ and $xe' = 0$. This shows that $I = Re$. The proof that $J = Re'$ is similar.

  Assume that I and J are ideals. The $IJ \subset I \cap J = 0$. Let $a \in R$. Then

  $$a = a(e + e') = ae + ae' = (e + e')a = ea + e'a,$$

  with $ae, ea \in I$ and $ae', e'a \in J$. As $R = I \oplus J$, we have $ae = ea$ and $ae' = e'a$. If moreover $a \in I$ (resp. $a \in J$), then $ae = ea = a$ and $ae' = e'a = 0$ (resp. $ae = ea = 0$ and $ae' = e'a = 1$). To finish the proof, let's show that the map $u : I \times J \to R$, $(a, b) \longmapsto a+b$, is an isomorphism of rings. We already know that it is an isomorphism of abelian groups by hypothesis. Let $a, a' \in I$ and $b, b' \in J$. Then

  $$u((a, b))u((a', b')) = (a+b)(a'+b') = aa'+ab'+ba'+bb' = aa'+bb' = u((a, b)(a', b')).$$

  □

I.1.4 Schur's lemma

Theorem I.1.4.1 (Schur's lemma)**.** *Let R be a ring, M and N be R-modules, and $u : M \to N$ be a R-linear map.*

1. *If M is simple, then $u = 0$ or u is injective.*
2. *If N is simple, then $u = 0$ or u is surjective.*

[7] See theorem (2.9) of Lam's book [20] and the remark following it.

3. *If M and N are simple, then $u = 0$ or u is an isomorphism.*

In particular, if M is a simple R-module, then $\mathrm{End}_R(M)$ is a division ring.

Proof. This follows from the fact that $\mathrm{Ker}\, u$ is a submodule of M and $\mathrm{Im}\, u$ is a submodule of N.

□

I.1.5 Jordan-Hölder theorem

Let R be a ring and M be a R-module.

Definition I.1.5.1.

- A *Jordan-Hölder series* (or *composition series*) for M is a sequence of submodules $0 = M_k \subset \cdots \subset M_1 \subset M_0 = M$ of M such that M_i/M_{i+1} is a simple R-module for every $i \in \{0, \ldots, k-1\}$. We say that the integer k is the *length* of the series.
- If a composition series for M exists, we say that M is has *finite length*. Then the *length* $\lg(M)$ of M is the minimum of the lengths of all its Jordan-Hölder series. If M is not of finite length, we set $\lg(M) = +\infty$.
- Two Jordan-Hölder series $(M_i)_{0\leq i\leq k}$ and $(M'_i)_{0\leq i\leq l}$ are called *equivalent* if $k = l$ and there exists a permutation $\sigma \in \mathfrak{S}_k$ such that $M_i/M_{i+1} \simeq M'_{\sigma(i)}/M'_{\sigma(i)+1}$ for every $i \in \{0, \ldots, k-1\}$.

Theorem I.1.5.2. *If M has finite length, then all its Jordan-Hölder series are equivalent. In particular, all the Jordan-Hölder series of M have the same length, and the length of M is the length of any of its Jordan-Hölder series.*

Lemma I.1.5.3. *Suppose as above that M has finite length, and let N be a submodule of M. Then N and M/N have finite length, and in fact $\lg(N) \leq \lg(M)$ and $\lg(M/N) \leq \lg(M)$..*

Proof. Let $M_k \subset \cdots \subset M_0 = M$ be a Jordan-Hölder series for M such that $k = \lg(M)$. For every $i \in \{0, \ldots, k-1\}$, $(N \cap M_i)/(N \cap M_{i+1})$ is a submodule of the simple module M_i/M_{i+1}, so it is either 0 or M_i/M_{i+1}. This means that, after deleting some steps to get rid of zero quotients, the sequence $N \cap M_k \subset \cdots \subset N \cap M_0 = N$ is a Jordan-Hölder series for N, and so N has finite length $\leq k = \lg(M)$.

The proof for M/N is similar : the image in M/N of the Jordan-Hölder series $M_k \subset \cdots \subset M_0$ is a Jordan-Hölder series for M/N after we delete some indices.

□

Proof of the theorem. We do an induction on $\lg(M)$. If $\lg(M) = 0$, then $M = 0$. If $\lg(M) = 1$, then M is simple and the result is obvious. So suppose that $\lg(M) \geq 2$ and that the result is known for all R-modules of strictly smaller length.

Let $(S) = (0 = M_k \subset \cdots \subset M_1 \subset M_0 = M)$ and $(S') = (0 = M'_l \subset \cdots \subset M'_1 \subset M'_0 = M)$ be two Jordan-Hölder series for M, and assume that $k = \lg(M)$. (It is clearly enough to treat this case.) If $M_1 = M'_1$, then we can apply the induction hypothesis to M_1 (which has length $\leq k - 1$ by the lemma or the existence of (S)), and this finishes the proof.

So assume that $M_1 \neq M'_1$. Then $M_1 + M'_1 = M$. Indeed, the image of M'_1 by the obvious map $M \to M/M_1$ is a submodule, so it is 0 or M/M_1 (because M/M_1 is simple). If it is 0, then $M'_1 \subset M_1$; then the submodule M_1/M'_1 of the simple module M/M'_1 has to be 0 (if it were M/M'_1, then we would have $M = M_1$), hence $M_1 = M'_1$, contradicting out assumption. So the image of M'_1 in M/M_1 is equal to M/M_1, which means that $M = M_1 + M'_1$.

As $M = M_1 + M'_1$, the obvious maps $M'_1/(M'_1 \cap M_1) \to M/M_1$ and $M_1/(M_1 \cap M'_1) \to M/M'_1$ are isomorphisms, and in particular both $M_1/(M_1 \cap M'_1)$ and $M'_1/(M_1 \cap M'_1)$ are simple. Take a Jordan-Hölder series $0 = N_s \subset \cdots \subset N_1 \subset N_0 = M_1 \cap M'_1$ (this exists by the lemma). Then applying the induction hypothesis to M_1 (which has length $\leq k - 1$ because it has a Jordan-Hölder series of length $k - 1$), we see that its two Jordan-Hölder series $M_k \subset \cdots \subset M_1$ and $N_s \subset \cdots \subset N_0 \subset M_1$ are equivalent. In particular, $s = k - 2$, so M'_1 has a Jordan-Hölder series of length $k - 1$, i.e. $N_s \subset \cdots \subset N_0 \subset M'_1$, which implies that M'_1 also has length $\leq k - 1$. We can then apply the induction hypothesis to M'_1 to see that its two Jordan-Hölder series $M'_l \subset \cdots \subset M'_1$ and $N_s \subset \cdots \subset N_0 \subset M'_1$ are equivalent. Finally, we have shown that (S) is equivalent to $N_s \subset \cdots \subset N_0 = M_1 \cap M'_1 \subset M_1 \subset M$, and that (S') is equivalent to $N_s \subset \cdots \subset N_0 = M_1 \cap M'_1 \subset M'_1 \subset M$. As $M'_1/(M'_1 \cap M_1) \simeq M/M_1$ and $M_1/(M_1 \cap M'_1) \simeq M/M'_1$, this show that (S) and (S') are equivalent.

□

This theorem justifies the following definition :

Definition I.1.5.4. If M has finite length qnd $(M_i)_{0 \leq i \leq k}$ is a Jordan-Hölder series for M, then the simple R-modules M_i/M_{i+1}, counted with multiplicities, are called the *Jordan-Hölder constituents* of M.

Corollary I.1.5.5. *If M has finite length and N is a submodule of M, then any Jordan-Hölder factor of N (resp. M/N) is also a Jordan-Hölder factor of M, and we have* $\lg(M) = \lg(N) + \lg(M/N)$.

Proof. We already know that N and M/N have finite length by lemma I.1.5.3. Let $N_k \subset \cdots \subset N_0 = N$ and $M'_l \subset \cdots \subset M'_0 = M/N$ be Jordan-Hölder series for N and M/N, and let M_i be the inverse image of M'_i in M, for every $i \in \{0, \ldots, l\}$. Then $N_k \subset N_0 = N = M_l \subset \cdots \subset M_0 = M$ is a Jordan-Hölder series for M.

□

I.1.6 Artinian and Noetherian modules

Definition I.1.6.1. Let R be a ring.

1. We say that a R-module M is *Artinian* (resp. *Noetherian*) if every strictly decreasing (resp. strictly increasing) sequence of submodules of M is finite, that is, for any sequence $M_0 \supset M_1 \supset \dots$ (resp. $M_0 \subset M_1 \subset \dots$) of submodules of M, there exists $N \in \mathbb{N}$ such that $M_i = M_{i+1}$ for every $i \geq N$.
2. We say that R is *left Artinian* (resp. *left Noetherian*) if the left R-module ${}_RR$ is Artinian (resp. Noetherian), that is, for any sequence $I_0 \supset I_1 \supset \dots$ (resp. $I_0 \subset I_1 \subset \dots$) of left ideals of R, there exists $N \in \mathbb{N}$ such that $I_i = I_{i+1}$ for every $i \geq N$.

Proposition I.1.6.2. *Let M be a R-module. Then M has finite length if and only if it is both Artinian and Noetherian.*

Lemma I.1.6.3. *Let M be a R-module.*

1. *If $M \neq 0$ and M is Artinian (resp. Noetherian), then it admits minimal (resp. maximal) nonzero submodules.*
2. *If M is Artinian (resp. Noetherian), so is any submodule and any quotient of M is Noetherian.*

Proof. Point (ii) is obvious. In (i), we treat the Artinian case (the Noetherian case is similar). Suppose that M has no minimal nonzero submodule and that $M \neq 0$. We construct by induction on i an infinite strictly decreasing sequence $(M_i)_{i\in\mathbb{N}}$ of nonzero submodules of M, which will prove that M is not Artinian. Take $M_0 = M$. Now let $i \geq 0$, suppose that $M_0 \supsetneq \dots \supsetneq M_i$ are constructed, and let's construct M_{i+1}. As M_i cannot be a minimal nonzero submodule of M, there exists a submodule M_{i+1} of M such that $0 \subsetneq M_{i+1} \subsetneq M_i$, and we are done.

□

Proof of the proposition. Suppose that M is Artinian and Noetherian. Let's prove that M has finite length. We construct by induction on i a sequence $(M_i)_{i\in\mathbb{N}}$ of submodules of M such that, for every $i \in \mathbb{N}$, $M_i \subset M_{i+1}$ and M_{i+1}/M_i is zero or simple.

Take $M_0 = 0$. Now suppose that $i \geq 0$ and that $M_0, \dots, M_i$ are constructed. If $M_i = M$, take $M_{i+1} = M_i$. Otherwise, then by the fact that M is Artinian and by the lemma, M/M_i has a minimal nonzero submodule, and we take for M_{i+1} its inverse image in M. By minimality of M_{i+1}/M_i, this R-module is simple.

We also know that M is Noetherian, so the sequence $(M_i)_{i\in\mathbb{N}}$ must stabilize to M after a finite number of steps. As all the quotients M_{i+1}/M_i are wero or simple, we can extract from $(M_i)_{i\in\mathbb{N}}$ a Jordan-Hölder sequence for M, and so M has finite length.

Now suppose that M has finite length, and let's prove that M is Artinian and Noetherian. If M is not Noetherian, then it has an infinite sequence of submodules $M_0 \subsetneq M_1 \subsetneq \dots$. But then

$\lg(M_{i+1}) \leq \lg(M_i) + 1$ for every $i \in \mathbb{N}$, so $\lg(M_i) \geq i$, so $\lg(M)$ cannot be finite. Similarly, we see that if M is not Artinian, then it cannot have finite length.

□

Remark I.1.6.4.

- If R is commutative, we recover the usual notions of Artinian and Noetherian ring.
- As in the commutative case, any left Artinian ring is automatically left Noetherian. [8] This is not true for modules.
- If k is a field and R is a k-algebra that is finite-dimensional as a k-vector space, then R is left Artinian and left Noetherian. (Because every left ideal of R is a k-vector subspace.)
- If R is a semisimple ring, then R is left Artinian and left Noetherian.

  Proof. By theorem I.1.3.4, $R = \bigoplus_{i \in A} I_i$, where the I_i are left ideals of R that are simple as R-modules. If we show that A is finite, then ${}_R R$ will have finite length, hence be an Artinian and Noetherian R-module by proposition I.1.6.2, and we will be done. To show that A is finite, write $1 = \sum_{i \in A} x_i$, with $x_i \in I_i$ for every $i \in A$ and $x_i = 0$ for all but a finite number of i's. Let $B \subset A$ be a finite subset such that $x_i = 0$ for $i \notin B$. As $R \cdot 1 = R$, we have $R = \sum_{i \in B} Rx_i$, so $R = \bigoplus_{i \in B} I_i$, so $B = A$, so A is finite.

  □

Example I.1.6.5.

- $\mathbb{Q}[T_1, \ldots, T_n]$ is Noetherian but not Artinian (consider the sequence of ideals $(T_1) \supset (T_1^2) \supset (T_1^3) \supset \ldots$).
- $\mathbb{Q}[T_i, i \in \mathbb{N}]$ is neither Artinian nor Noetherian.
- $\mathbb{Q}[T]/(T^2)$ is borth Artinian and Noetherian.

I.1.7 Isotypic decomposition

Definition I.1.7.1. Let R be a ring. We write $S(R)$ for the set of isomorphism classes of simple R-modules.

Theorem I.1.7.2. *Let R be a ring and M be a semisimple R-module. For every $S \in S(R)$, let M_S be the sum of all the submodules of M that are isomorphic to S. Then :*

1. *We have $M = \bigoplus_{S \in S(R)} M_S$.*
2. *There exist sets I_S such that $M_S \simeq \bigoplus_{i \in I_S} S$ for every $S \in S(R)$.*

[8] See theorem (4.15) of Lam's book [20].

3. *Let N be a submodule of M. If we write $N \simeq \bigoplus_{S\in S(R)} N_S$ and $N_S \simeq \bigoplus_{S\in J_S} S$ as in (i) and (ii), then, for every $S \in S(R)$, $N_S = M_S \cap M$ for every $S \in S(R)$ and we can find an injection $J_S \hookrightarrow I_S$.*

In particular, if the set I_S is finite for some $S \in S(R)$, then $|I_S|$ depends only on M.

The nonzero M_S are called the *isotypic components* of M, and $M = \bigoplus_{S\in S(R)} M_S$ is called the *isotypic decomposition* of M. If I_S is a finite set for some $S \in S(R)$, we call $|I_S|$ the *multiplicity* of S in M.

The following lemma will be used repeatedly in the proof of the theorem.

Lemma I.1.7.3. *Let M be a R-module, and suppose that there exists a simple R-module S and s set I and an isomorphism $\varphi : M \xrightarrow{\sim} \bigoplus_{i\in I} S$. Then every simple R-submodule of M is isomorphic to S.*

Proof. Let N be a simple R-submodule of M, and suppose that N is not isomorphic to S. Then for every $i \in I$, the composition of the projection on the ith summand $\bigoplus_{i\in I} S \to S$ and of φ is a R-linear map $N \to S$, which has to be zero by Schur's lemma (theorem I.1.4.1). But then φ is zero on N, which implies that $N = 0$, contradicting the fact that N is simple.

□

Proof of the theorem.

1. First, note that $\sum_{S\in S(R)} M_S$ is the sum of all the simple submodules of M. As M is semisimple, $M = \sum_{S\in S(R)} M_S$. Now we want to show that the sum is direct. Let $S \in S(R)$, let $X = S(R) - \{S\}$. We must show that $N := M_S \cap \left(\sum_{S'\in X} M_{S'}\right) = 0$. As N is submodule of M, it is semisimple, so, if it's nonzero, then it has a simple submodule. But if N' is a simple submodule of N, then it's a simple submodule of M_S, hence isomorphic to S, and also a simple submodule of $\sum_{S'\in X} M_{S'}$, hence isomorphic to an element of X. This is not possible. So $N = 0$.
2. As M_S is a submodule of the semisimple R-module M, it is semisimple. By theorem I.1.3.4, M_S is the direct sum of a family of simple submodules. But every simple submodules of M_S is isomorphic to S.
3. Let $S \in S(R)$. Then N_S is a sum of simple R-modules isomorphic to S, so $N_S \subset M_S$. On the other hand, $N \cap M_S$ is a direct sum of simple submodules of N, and all these simple modules have to be isomorphic to S, so $N \cap M_S \subset N_S$.

   So we may assume that $M = M_S$ and $N = N_S$ for some $S \in S(R)$, and we are reduced to the following statement : If S is a simple R-module and I, J are two sets such that we have an injective R-linear map $u : N := \bigoplus_{j\in J} S \hookrightarrow M := \bigoplus_{i\in I} S$, then there exists an injection $J \hookrightarrow I$.

   We will only use this statement when I and J are finite, and in that case it follows from the Jordan-Hölder theorem (theorem I.1.5.2), but let's see how to prove it general. For any

R-modules M_1, M_2, let $\mathrm{Hom}_{fs}(M_1, M_2) \subset \mathrm{Hom}_R(M_1, M_2)$ be the subgroup of R-linear maps that are 0 outside a finite length submodule of M_1. Then we have

$$\mathrm{Hom}_{fl}(N, S) = \bigoplus_{j \in J} \mathrm{End}_R(S) \subset \mathrm{Hom}_R(N, S) = \prod_{j \in J} \mathrm{End}_R(S),$$

and similarly

$$\mathrm{Hom}_{fl}(M, S) = \bigoplus_{i \in I} \mathrm{End}_R(S) \subset \mathrm{Hom}_R(M, S) = \prod_{i \in I} \mathrm{End}_R(S).$$

Note that $\mathbb{D} := \mathrm{End}_R(S)$ is a division ring by Schur's lemma (theorem I.1.4.1), that $\mathrm{Hom}_R(M, S)$ and $\mathrm{Hom}_R(N, S)$ are naturally $\mathbb{D}$-modules (if $f \in \mathrm{Hom}_R(M, S)$ or $\mathrm{Hom}_R(N, S)$ and $r \in D$, set $r \cdot f = r \circ f$) and that $\mathrm{Hom}_{fl}(M, S)$ and $\mathrm{Hom}_{fl}(N, S)$ are $\mathbb{D}$-submodules. Moreover, we have a map $\varphi : \mathrm{Hom}_{fl}(M, S) \to \mathrm{Hom}_{fl}(N, S)$, $f \longmapsto f \circ u$, and it is clearly $\mathbb{D}$-linear. If we can show that φ is surjective, we will done by linear algebra. (More precisely, the incomplete basis theorem.) So let $g \in \mathrm{Hom}_{fl}(N, S)$. As M is a semisimple R-module, there exists a R-submodule N' of M such that $M = u(N) \oplus N'$. Define $f : M \to S$ by $f(u(x) + y) = g(x)$ if $x \in N$ and $y \in N'$. This makes sense because u is injective, is clearly R-linear, and we have $\varphi(f) = g$.

□

I.1.8 Simple rings

Definition I.1.8.1. A ring R is called *simple* if $R \neq 0$ and if the only ideals of R are 0 and R.

Remark I.1.8.2. Note that the definition of a simple ring involves ideals, and that of a semisimple ring involves *left* ideals. In particular, a simple ring has no a priori reason to be semisimple, and indeed there exist simple rings that are not semisimple.[9] So the terminology is a bit unfortunate.

Theorem I.1.8.3. *Let $\mathbb{D}$ be a division ring, $n \geq 1$ be an integer, and $R = M_n(\mathbb{D})$. Let $V = \mathbb{D}^n$. We will write ${}_RV$ when we view V as a left R-module by considering $\mathbb{D}^n$ as a space of $n \times 1$ matrices, and $V_{\mathbb{D}}$ when we view V as the right $\mathbb{D}$-module $\mathbb{D}^n_{\mathbb{D}}$.*

Then :

1. *The ring R is simple, semisimple, left Artinian and left Noetherian.*
2. *R has a unique (up to isomorphism) simple left module, which is ${}_RV$. As left R-modules, ${}_RR$ and ${}_RV^n$ are isomorphic.*
3. $\mathrm{End}_{\mathbb{D}}(V_{\mathbb{D}}) = R$.
4. $\mathrm{End}_R({}_RV) = \mathbb{D}$.

[9] See problem VII.1.9.

Proposition I.1.8.4. *If R is a ring and $n \geq 1$ is an integer, then any ideal I of $M_n(R)$ if of the form $I = M_n(J)$, where J is a uniquely determined ideal of R.*

Corollary I.1.8.5. *If R is a simple ring, then so is $M_n(R)$ for every $n \geq 1$.*

Proof of the proposition. If J is an ideal of R, then $M_n(J)$ is clearly an ideal of $M_n(R)$. Also, if J, J' are two ideals of R such that $M_n(J) = M_n(J')$, then it is obvious that $J = J'$. So we just need to prove that any ideal of $M_n(R)$ is of the form $M_n(J)$.

Let I be an ideal of $M_n(R)$, and let J be the set of $(1,1)$-entries of elements of I. We'll show that J is an ideal and that $I = M_n(J)$.

First, let $x, y \in J$ and let $a \in R$. Choose matrices $X, Y \in I$ such that the $(1,1)$-entries of X and Y are x and y respectively. Then aX, Xa and $X + Y$ are in I, and their respective $(1,1)$-entries are ax, xa and $x + y$, so $ax, xa, x + y \in J$. So J is an ideal of R.

Now let's denote by E_{ij}, for $1 \leq i, j \leq n$, the elementary matrices in $M_n(R)$. (So E_{ij} has all its entries equal to 0, except for the entry (i, j) which is equal to 1.) If $X = (x_{ij}) \in M_n(R)$, then $E_{ij}XE_{kl} = x_{jk}E_{jk}$.

Let's show that $I \subset M_n(J)$. If $X \in I$, then for all $j, k \in \{1, \ldots, n\}$, $E_{1j}XE_{k1} = x_{jk}E_{11} \in I$, so $x_{jk} \in J$, and so $X \in M_n(J)$.

Let's show that $M_n(J) \subset I$. Let $X = (x_{ij}) \in M_n(J)$. Then $X = \sum_{1 \leq i,j \leq n} x_{ij}E_{ij}$, so it suffices to show that all the $x_{ij}E_{ij}$ are in I. Fix $i, j \in \{1, \ldots, n\}$. Choose $Y \in I$ such that the $(1,1)$-entry of Y is x_{ij}. Then $E_{i1}YE_{1j} = x_{ij}E_{ij} \in I$.

□

Proof of the theorem. Let's prove (i). First, by the proposition, R is simple because $\mathbb{D}$ is. As a left $\mathbb{D}$-vector space, R is finite-dimensional. Since left ideals of R are $\mathbb{D}$-vector subspaces of R, R is left Artinian and left Noetherian.[10]

Let's prove that ${}_RV$ is a simple R-module. Take a nonzero R-submodule W of ${}_RV$. We use the same notation E_{ij} as in the proof of the proposition (for the elementary matrices in R). Let $w \in W - \{0\}$, and write $w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$. Choose $i_0 \in \{1, \ldots, n\}$ such that $w_{i_0} \neq 0$. Then

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = (w_{i_0}^{-1}E_{1i_0})w \in W.$$

[10] This would also follow from the fact that R is semisimple.

Now if $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ is any element of ${}_RV$, then

$$v = \begin{pmatrix} v_1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ v_n & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in W,$$

and so $W = V$.

As the left R-module ${}_RR$ is clearly isomorphic to ${}_RV^n$, the ring R is semisimple.

Let's prove (ii). It only remains to show that any simple R-module is isomorphic to ${}_RV$. So let M be a simple R-module. Choose $x \in M - \{0\}$. Then the map $u : {}_RR \to M$, $a \longmapsto ax$ is surjective (because its image is a nonzero submodule of M), so M is isomorphic to a quotient of the R-module ${}_RR$. As ${}_RR \simeq {}_RV^n$ with ${}_RV$, this implies that $M \simeq {}_RV$.

Let's prove (iii). Consider the map $\varphi : R \to \mathrm{End}_{\mathbb{D}}(V_{\mathbb{D}})$, $a \longmapsto (x \longmapsto xa))$. This map is well-defined, because for every $a \in R$, for every $x \in V$ and $\lambda \in \mathbb{D}$,

$$\varphi(a)(x\lambda) = a(x\lambda) = (ax)\lambda = (\varphi(a)(x))\lambda,$$

so $\varphi(a)$ is $\mathbb{D}$-linear. The map φ is obviously a map of rings, and we want to show that it is an isomorphism.

Let $a = (a_{ij}) \in R = M_n(\mathbb{D})$ such that $\varphi(a) = 0$. Then for every $j \in \{1, \dots, n\}$, if $e_j \in V$ is the element with jth entry equal to 1 and all the other entries equal to 0, we have

$$0 = \varphi(a)(e_j) = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}.$$

So $a = 0$. This proves that φ is injective.

Let $u \in \mathrm{End}_{\mathbb{D}}(V_{\mathbb{D}})$. For every $j \in \{1, \dots, n\}$, if $e_j \in V$ is defined as above, write

$$u(e_j) = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}.$$

Let $a = (a_{ij}) \in M_n(\mathbb{D})$. Then $\varphi(a)(e_j) = u(e_j)$ for every $j \in \{1, \dots, n\}$. As $(e_1, \dots, e_n)$ is obviously a basis of $V_{\mathbb{D}}$ over $\mathbb{D}$, this imples that $\varphi(a) = u$. So we have proved that φ is surjective.

Let's prove (iv). Let $E = \mathrm{End}_R({}_RV)$. We write the action of the ring E on V on the right, that is, we write $x(v) = vx$ for $v \in V$ and $x \in E$. Let $\psi : \mathbb{D} \to E$, $\lambda \longmapsto (v \longmapsto v\lambda)$. As

for φ, this map is well-defined, i.e. $\delta(\lambda)$ is R-linear for every $\lambda \in \mathbb{D}$, and it is a morphism of rings. We want to show that it is an isomorphism of rings. First, ψ is injective because $n \geq 1$ and $\mathbb{D}$ is a division algebra. Now let $x \in E$. Define $\lambda \in \mathbb{D}$ by $e_1 x = \lambda e_1 + \sum_{j=2}^{n} \mu_j e_j$, with $\mu_j \in \mathbb{D}$ and where $e_1, \ldots, e_n \in V$ are as in the proof of (iii). Let $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in V$, and let $a = \begin{pmatrix} v_1 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ v_n & 0 & \ldots & 0 \end{pmatrix} \in M_n(\mathbb{D}) = R$. Then $v = ae_1$, so

$$vx = (ae_1)x = a(e_1 x) = a \begin{pmatrix} \lambda \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix} = \begin{pmatrix} v_1\lambda \\ \vdots \\ v_n\lambda \end{pmatrix} = v\lambda.$$

So $x = \psi(\lambda)$, and ψ is surjective.

□

Corollary I.1.8.6. *If $\mathbb{D}, \mathbb{D}'$ are two division rings and $n, n' \geq 1$ are two integers such that $M_n(\mathbb{D}) \simeq M_{n'}(\mathbb{D}')$ as rings, then $\mathbb{D} \simeq \mathbb{D}'$ and $n = n'$.*

Proof. Let $R = M_n(\mathbb{D}) \simeq M_n(\mathbb{D}')$, and let M be the unique simple R-module (given by (ii) of the theorem). By (iii) of the theorem, $\mathbb{D}' \simeq \mathrm{End}_R(M) \simeq \mathbb{D}$. Hence

$$n = \dim_{\mathbb{D}}(M) = \dim_{\mathbb{D}'}(M) = n'.$$

□

I.1.9 Double centralizer property

We have seen in theorem I.1.8.3 that every $M_n(\mathbb{D})$ with $\mathbb{D}$ a division ring is simple. We'll now see a kind of converse of this. (Not an actual converse, as there are simple rings not of the form $M_n(\mathbb{D})$.)

Definition I.1.9.1. If R is a ring, we denote by R^{op} its opposite ring : it's isomorphic to R as an additive group, and its multiplication is given by

$$ab(\text{ in } R^{\mathrm{op}}) = ba(\text{ in } R).$$

Theorem I.1.9.2 (Double centralizer property)**.** *Let R be a simple ring and I be a nonzero left ideal of R. Let $D = \mathrm{End}_R(I)$, and make I a right D^{op}-module by setting $xu = u(x)$ if $x \in I$ and $u \in D$.*

Then the map $f : R \to \mathrm{End}_{D^{\mathrm{op}}}(I)$, $a \longmapsto (x \longmapsto ax)$, is an isomorphism of rings.

Proof. It's obvious that f is well-defined and is a morphism of rings.

Let's show that f is injective. As $I \neq 0$, $\mathrm{Ker}\, f \neq R$ (for example, $1 \notin \mathrm{Ker}\, f$). As $\mathrm{Ker}\, f$ is an ideal of R and R is simple, $\mathrm{Ker}\, f = 0$.

Let's show that f is surjective. Let $E = \mathrm{End}_{D^{\mathrm{op}}}(I)$. Make I a left E-module by setting $hx = h(x)$, for every $x \in I$ and $h \in E$. Then, for every $x \in I$ and $h \in E$, we have $hf(x) = f(hx)$. Indeed, if $a \in I$, then $r_a : I \to I$, $y \longmapsto ya$, is in D, so

$$h(xa) = h(r_a(x)) = h(xr_a) = h(x)r_a = h(x)a,$$

and so

$$(hf(x))(a) = h(xa) = h(x)a = f(h(x))(a).$$

This imples that $Ef(I) \subset f(I)$. But we know I is a nonzero left ideal of R, so IR is a nonzero ideal of R, hence $IR = R$ as R is simple. Apply the morphism of rings f gives $f(I)f(R) = f(R)$. Finally,

$$E = Ef(R) = Ef(I)f(R) \subset f(I)f(R) = f(R)$$

(the first equality holds because $1 \in f(R)$), and so f is surjective.

□

Corollary I.1.9.3. *If R is a simple ring with a minimal nonzero left ideal, then there exists a unique division ring $\mathbb{D}$ and a unique integer $n \geq 1$ such that $R \simeq M_n(\mathbb{D})$.*

In particular, a simple ring is left Artinian if and only if it is of the form $M_n(\mathbb{D})$ with $\mathbb{D}$ a division ring and $n \geq 1$.

Proof. We already know that $\mathbb{D}$ and n are unique if they exist (by corollary I.1.8.6).

Let $I \subset R$ be a minimal nonzero left ideal. Then I is a simple R-module, so $\mathbb{D} := \mathrm{End}_R(I)$ is a division ring by Schur's lemma (theorem I.1.4.1). Make I a right $\mathbb{D}^{\mathrm{op}}$-module as as in theorem above. (Note that $\mathbb{D}^{\mathrm{op}}$ is also a division ring.) By that theorem, $R \simeq \mathrm{End}_{\mathbb{D}^{\mathrm{op}}}(I)$. So we only need to show that I is finite-dimensional as a right $\mathbb{D}^{\mathrm{op}}$-vector space. Let $E = \mathrm{End}_{\mathbb{D}^{\mathrm{op}}}(I)$, and let

$$E_f = \{u \in E | \mathrm{rk}(u) < +\infty\},$$

where $\mathrm{rk}(u)$ is as usual the dimension (over $\mathbb{D}^{\mathrm{op}}$ of the image of u). It's easy to see that $E_f \neq 0$, and that E_f is an ideal of E. As $E \simeq R$ is a simple ring, $E_f = E$. Hence the identity of I is in E_f, and I is a finite-dimensional right $\mathbb{D}^{\mathrm{op}}$-vector space.

The last sentence follows from theorem I.1.8.3 and the fact that a left Artinian ring admits minimal nonzero left ideals (by lemma I.1.6.3).

□

I.1.10 Structure of semisimple rings (Artin-Wedderburn theorem)

The goal of this section is to show that every semisimple is a finite direct product of rings of the form $M_n(\mathbb{D})$, with $\mathbb{D}$ a division ring. This will imply in particular that the notions of "left semisimple rings" and "right semisimple rings" coincide.

Proposition I.1.10.1. *Let $R_1, \dots, R_n$ be rings, and let $R = R_1 \times \cdots \times R_n$. Then every R-module is of the form $M = M_1 \times \cdots \times M_n$, with M_i a R_i-module for $1 \leq i \leq n$.*

Proof. In R, write $1 = e_1 + \cdots + e_n$, with $e_i \in R_i$. Of course, as an element of $R_1 \times \cdots \times R_n$, e_i is the n-uple with ith entry equal to $1 \in R_i$ and all the other entries equal to 0. Note that all the e_i are central in R, that $e_i^2 = e_i$ for every i and that $e_i e_j = e_j e_i = 0$ for $i \not\mathrel{/} j$.

Let M be a R-module, set $M_i = e_i M$. Then R acts on M_i through the obvious projection $R \to R_i$, so we just need to show that $M \simeq M_1 \times \cdots \times M_n$. Consider the map $u : M_1 \times \dots M_n \to M$, $(x_1, \dots, x_n) \longmapsto x_1 + \cdots + x_n$; this is a R-linear map by the previous remark about the action of R on the M_i. If $x \in M$, then $x = xe_1 + \cdots + xe_n$ with $xe_i \in M_i$ for every i, so u is surjective. Moreover, if $x_1 + \cdots + x_n = 0$ with $x_i \in M_i$ for every i, then for every $j \in \{1, \dots, n\}$, $0 = e_j(x_1 + \cdots + x_n) = e_j x_i = x_i$. Hence u is injective.

□

Corollary I.1.10.2. *Suppose that $R = R_1 \times \cdots \times R_n$ as in the proposition.*

1. *R is semisimple if and only if all the R_i are semisimple.*
2. *Every simple R-module is of the form $0 \times \cdots \times 0 \times M_i \times 0 \times \cdots \times 0$, with $i \in \{1, \dots, n\}$ and M_i a simple R_i-module.*

Corollary I.1.10.3. *Let $\mathbb{D}_1, \dots, \mathbb{D}_r$ be division rings, and $n_1, \dots, n_r \geq 1$ be integers. Then $R := M_{n_1}(\mathbb{D}_1) \times \cdots \times M_{n_r}(\mathbb{D}_r)$ is a semisimple ring, and its simple modules (up to isomorphism) are $\mathbb{D}_1^{n_1}, \dots \mathbb{D}_r^{n_r}$.*

Conversely, we want to show that every semisimple ring is of this form.

Notation I.1.10.4. Let R be a ring and I be a left ideal of R. In the rest of this section, we denote by $\mathscr{I}_I \subset R$ the sum of all the left ideals I' of R that are isomorphic to I as R-modules. This is a left ideal of R.

Theorem I.1.10.5 (Artin-Wedderburn theorem)**.** *Let R be a semisimple ring. Let $(I_i)_{i \in A}$ be a set of representatives of the isomorphism classes of minimal nonzero left ideals of R, and let $R_i = \mathscr{I}_{I_i} \subset R$.*

Then A is finite, so we choose an identification $A = \{1, \dots, r\}$. Moreover, all the R_i are rings (with unit), $R \simeq R_1 \times \cdots \times R_r$ as rings, and for every $i \in \{1, \dots, r\}$, there exists a unique division ring $\mathbb{D}_i$ and a unique integer $n_i \geq 1$ such that $R_i \simeq M_{n_i}(\mathbb{D}_i)$.

Lemma I.1.10.6. *Let R be a ring and I be a minimal nonzero left ideal of R. Then $\mathscr{I}_I$ is an* ideal *of R. (That is, it is also a right ideal of R.)*

Moreover, if I and J are two nonisomorphic minimal nonzero left ideals of R, then $\mathscr{I}_I \mathscr{I}_J = 0$.

Proof. The main point is that minimal nonzero left ideals of R are simple R-modules.

Let's prove that $\mathscr{I}_I$ is a right ideal. Let I' be a left ideal of R such that $I' \simeq I$ as R-modules, and let $a \in R$. We want to show that $I'a \subset \mathscr{I}_I$. We have a surjective R-linear map $I' \to I'a$, $x \longmapsto xa$, and I' is a simple R-module, so $I'a = 0$ or $I'a \simeq I'$. In the first case, $I'a \subset \mathscr{I}_I$ is obvious; in the second case, $I'a$ is another left ideal of R that is isomorphic to I, so we also have $I'a \subset \mathscr{I}_I$.

Now let J be another minimal nonzero left ideal of R, and suppose that $\mathscr{I}_I \mathscr{I}_J \neq 0$. Then there exist left ideals I', J' of R and an element a of J' such that $I' \simeq I$, $J' \simeq J$ and $I'a \neq 0$. As J' is a simple R-module and $I'a \subset J'$ is a nonzero submodule of J', we have $I'a = J'$. As I' is a simple R-module, the surjective map $I' \to I'a$, $x \longmapsto xa$, is an isomorphism. So we get R-module isomorphisms $I \simeq I' \simeq I'a = J' \simeq J$. This proves the second part of the lemma.

□

Proof of the theorem. As R is a semisimple ring, ${}_RR$ is a direct sum of simple submodules (= minimal nonzero left ideals of R) by theorem I.1.3.4. So ${}_RR = \bigoplus_{i \in A} R_i$ by theorem I.1.7.2. By remark I.1.6.4 (that says that semisimple rings are left Artinian and left Noetherian), this implies that A is finite.

By the lemma, every $\mathscr{I}_i$ is an ideal of R. As $R = \bigoplus_{i=1}^{r} R_i$, remark I.1.3.16 implies that all the R_i are rings (with unit) and that $R \simeq R_1 \times \cdots \times R_r$ as rings.

We now prove that all the R_i are simple rings. Fix $i \in \{1, \dots, r\}$. Let $J \neq 0$ be an ideal of R_i. We want to show that $J = R_i$. As J is also an ideal of R, it contains a minimal nonzero left ideal I of R (by remark I.1.3.16). By definition of $I_1, \dots, I_r$, there exists $j \in \{1, \dots, j\}$ such that $I \simeq I_j$ as R-modules; as $I \subset J \subset R_i$, we must have $j = i$. So $R_i = \mathscr{I}_{I_i} = \mathscr{I}_I$. Hence it suffices to show that, if I' is a left ideal of R such that $I' \simeq I$, then $I' \subset J$. Fix such a I', and let $\varphi : I \xrightarrow{\sim} I'$ be an isomorphism of R-modules. As R is a semisimple ring, there exists a left ideal I'' of R such that $R = I \oplus I''$. We write $1 = e + e''$ with $e \in I$ and $e'' \in I''$. We have seen in remark I.1.3.16 that $e^2 = e$ and $I = Re$. So

$$I' = \varphi(I) = \varphi(Re) = \varphi(Re^2) = \varphi((Re)e) = \varphi(Ie) = I\varphi(e) \subset J$$

(as J is also a right ideal of R).

As R is semisimple, it's left Artinian (by remark I.1.6.4), so all the R_i are left Artinian, so by corollary I.1.9.3 there exist uniquemy determined division rings $\mathbb{D}_i$ and integers $n_i \geq 1$ such that $R_i \simeq M_{n_i}(\mathbb{D}_i)$ for every i.

□

Remark I.1.10.7. Let R be a semisimple ring, and use the notation of the Artin-Wedderburn theorem.

- Let K be a field, and suppose that R is a semisimple K-algebra that is finite-dimensional as a K-vector space. Then its simple factors R_i are also K-algebras, and so are the division rings $\mathbb{D}_i$; of course, $\dim_K(\mathbb{D}_i) < +\infty$. (This follows for example from the fact that $\mathbb{D}_i \simeq \mathrm{End}_{R_i}(V_i) = \mathrm{End}_R(V_i)$, where V_i is the simple R-module corresponding to the factor R_i, see theorem I.1.8.3.)

  In particular, if K is algebraically closed, then all the $\mathbb{D}_i$ are equal to K by problem VII.1.3. so $R \simeq M_{n_1}(K) \times \cdots \times M_{n_r}(K)$.

- If R is commutative, then $n_1 = \cdots = n_r = 1$ and all the $\mathbb{D}_i$ are commutative division rings (i.e. fields), so R is a finite product of fields. This recovers the result of problem VII.1.5, but without the Noetherian hypothesis on R.

I.2 Jacobson radical

We will just give some basic definitions and facts about the Jacobson radical of a ring (as much as we need for our representation theoretic purposes).

The basic idea is that the Jacobson radical of a ring R should be the minimal ideal I of R such that the ring R/I is semisimple. Actually, this is true if R is left Artinian, but the general situation is more complicated.

Definition I.2.1. Let R be a ring. The *Jacobson radical* of R is the intersection of all the maximal left ideals of R. We will denote it by $\mathrm{rad}(R)$.

Remark I.2.2. At this point, it looks like this should be called the left Jacobson radical of R, but we will see in corollary I.2.9 that $\mathrm{rad}(R)$ is also the intersection of all the maximal right ideals of R.

Definition I.2.3. Let R be a ring and M be a R-module. If $x \in M$, the *annihilator of x in R* is

$$\mathrm{Ann}_R(x) = \{a \in R | ax = 0\}.$$

This is obviously a left ideal of R. Also, the *annihilator of M in R* is

$$\mathrm{Ann}_R(M) = \bigcap_{x \in M} \mathrm{Ann}_R(x) = \{a \in R | \forall x \in M,\ ax = 0\}.$$

This is also obviously a left ideal of R.

Proposition I.2.4. *Let R be a ring and M be a R-module. Then $\mathrm{Ann}_R(M)$ is actually an ideal of R.*

Proof. We just need to show that it's a right ideal. Let $a \in \mathrm{Ann}_R(M)$ and $b \in R$. Then for every $x \in M$, $(ab)x = a(bx) = 0$ because a is also in the annihilator of $bx \in M$. Hence $ab \in \mathrm{Ann}_R(M)$.

□

Proposition I.2.5. *Let R be a ring and $x \in R$. The following are equivalent :*

1. $x \in \mathrm{rad}(R)$.
2. *For every $y \in R$, $1 - yx$ is left invertible. (See definition I.1.1.16.)*
3. *For every simple R-module M, $x \in \mathrm{Ann}_R(M)$.*

Proof.

(i)⇒(ii) : Suppose that $x \in \mathrm{rad}(R)$. Let $y \in R$. If $1 - yx$ is not left invertible, then $R(1 - yx) \subsetneq R$, so there exists a maximal left ideal $\mathfrak{m}$ of R such that $1 - yx \in \mathfrak{m}$. But $x \in \mathrm{rad}(R) \subset \mathfrak{m}$, so $yx \in \mathfrak{m}$, so $1 \in \mathfrak{m}$, which is not possible.

(ii)⇒(iii) Let M be a simple R-module, and let $m \in M$. If $xm \neq 0$, then $Rxm = M$ (because Rxm is a nonzero submodule of M), so there exists $y \in R$ such that $yxm = m$, i.e. $(1-yx)m = m$. As $1-yx$ is left invertible, this implies that $m = 0$, which contradicts the assumption that $xm \neq 0$.

(iii)⇒(i) Let $\mathfrak{m} \subset R$ be a maximal left ideal. Then $R/\mathfrak{m}$ is a simple R-module, so $x \in \mathrm{Ann}_R(R/\mathfrak{m})$, i.e. $x(R/\mathfrak{m}) = 0$, i.e. $x \in \mathfrak{m}$.

□

Corollary I.2.6. *The Jacobson radical of R is the intersection of the annihilators of all the simple R-modules. In particular, it is an ideal of R.*

Corollary I.2.7. *The rings R and $R/\mathrm{rad}(R)$ have the same simple modules.*

Corollary I.2.8. *Let $x \in R$. The following are equivalent :*

1. $x \in \mathrm{rad}(R)$.
2. *For all $y, z \in R$, $1 - yxz$ is invertible.*

Proof. We already know that (ii) implies (i) by the proposition. Let's prove that (i) implies (ii). Let $x \in \mathrm{rad}(R)$, and let $y, z \in R$. As $\mathrm{rad}(R)$ is an ideal of R by a previous corollary, $xz \in \mathrm{rad}(R)$, so $1 - yxz$ is left invertible by the proposition, so there exists $u \in R$ such that $u(1 - yxz) = 1 = u - uyxz$.

Moreover, $yxz \in \mathrm{rad}(R)$, so $u = 1 + u(yxz)$ is left invertible by the proposition. As u is left and right invertible, it is invertible, and hence $1 - yxz$ is the (unique) inverse of u and is also invertible.

□

As the characterization of $\mathrm{rad}(R)$ in the corollary above is unchanged if we reverse the order of the multiplication in R, we get the :

Corollary I.2.9. *We have* $\mathrm{rad}(R) = \mathrm{rad}(R^{\mathrm{op}})$*, where* R^{op} *is as in definition I.1.9.1. In other words, the ideal* $\mathrm{rad}(R)$ *is also the intersection of all the maximal right ideals of* R*.*

Remark I.2.10. If $I \subset \mathrm{rad}(R)$ is an ideal of R, then $\mathrm{rad}(R/I) = \mathrm{rad}(R)/I$. In particular, $\mathrm{rad}(R/\mathrm{rad}(R)) = 0$.

Theorem I.2.11. *Assume that* R *is* left Artinian*. Then the following are equivalent :*

1. *The ring* R *is semisimple.*
2. $\mathrm{rad}(R) = 0$*.*

Proof.

(i)⇒(ii) If R is semisimple, then there exists a left ideal I of R such that ${}_RR = I \oplus \mathrm{rad}(R)$. If $\mathrm{rad}(R) \neq 0$, then $I \neq R$, so there exists a maximal left ideal $\mathfrak{m}$ of R such that $I \subset \mathfrak{m}$. But then $I \subset \mathrm{rad}(R)$, so $\mathrm{rad}(R) = R$, which is only possible if $R = \{0\}$, and this contradicts $\mathrm{rad}(R) \neq 0$.

(ii)⇒(i) Let $(\mathfrak{m}_i)_{i\in A}$ be the family of all maximal left ideals of R. We have $\mathrm{rad}(R) = \bigcap_{i\in A} \mathfrak{m}_i$. As R is left Artinian, there exists a finite subset B of A such that $\mathrm{rad}(R) = \bigcap_{i\in B} \mathfrak{m}_i$. (If this were not true, we could find a sequence $B_0 \supset B_1 \supset \dots$ of finite subsets of A such that $\bigcap_{i\in B_0} \mathfrak{m}_i \supsetneq \bigcap_{i\in B_1} \mathfrak{m}_i \supsetneq \dots$, which would contradict the fact that R is left Artinian.)

Now if $\mathrm{rad}(R) = 0$, then the obvious R-module map ${}_RR \to \bigoplus_{i\in B} R/\mathfrak{m}_i$ is injective. As each $R/\mathfrak{m}_i$ is a simple R-module, their direct sum over $i \in B$ is a semisimple R-module, and so is its submodule ${}_RR$.

□

Corollary I.2.12. *If* R *is left Artinian, then* $R/\mathrm{rad}(R)$ *is semisimple, and it has the same simple modules as* R*.*

Remark I.2.13.

- In general, a ring R such that $\mathrm{rad}(R) = 0$ is called *Jacobson semisimple*. Any semisimple ring is Jacobson semisimple, but the converse is false. For example, if G is any group, then the groups algebras $\mathbb{C}[G]$ and $\mathbb{R}[G]$ are Jacobson (see theorem (6.4) of Lam's book [20]), but they are not semisimple if G is infinite by remark I.3.3.

- A left (resp. right) ideal I of R is called *nilpotent* if there exists $n \geq 1$ such that $I^n = 0$.

  If R is left Artinian, then $\mathrm{rad}(R)$ is the largest nilpotent left (resp. right) ideal of R, see theorem (4.12) of Lam's book [20]. This is not true in general. For example, if $R = \mathbb{Z}_p$ (the ring of p-adic integers), then $\mathrm{rad}(R) = p\mathbb{Z}_p$ is not nilpotent.

Example I.2.14.

- $\mathrm{rad}(\mathbb{Z}) = \bigcap_{p \text{ prime}} p\mathbb{Z} = 0$, even though $\mathbb{Z}$ is not semisimple. (Note that $\mathbb{Z}$ is not Artinian.)
- $\mathrm{rad}(\mathbb{Z}_p) = p\mathbb{Z}_p$ (where $\mathbb{Z}_p$ is the ring of p-adic integers).

  More generally, if A is a commutative local ring, then $rad(A)$ is its unique maximal ideal.
- Let $\mathbb{D}$ be a division ring and R be the ring of upper triangular $n \times n$ matrices with coefficients in $\mathbb{D}$. Then $\mathrm{rad}(R)$ is the set of strictly upper triangular matrices (i.e. upper triangular matrices with zeroes on the diagonal). Indeed, let's call this set J. Then J is an ideal of R, and $1 - x$ is invertible for every $x \in J$, so $J \subset \mathrm{rad}(R)$. Moreover, $R/J \simeq \mathbb{D}^n$ is semisimple, so $J \supset \mathrm{rad}(R)$.

I.3 Applications to the representation theory of finite groups

Let R be a ring and G be a group. Remember (exemple I.1.1.1 that the group algebra $R[G]$ of G with coefficients in R is defined to be $\bigoplus_{g \in G} Rg$, with the multiplication given by $(ag)(bh) = (ab)(gh)$ if $a, b \in R$ and $g, h \in G$.

Vocabulary I.3.1. A $R[G]$-module is a R-module M with a R-linear action of G, i.e. a morphism of monoids $\rho : G \to \mathrm{End}_R(M)$. This is also called a *(R-linear) representation* of G on the R-module M, and denoted by (M, ρ), or just M if the action is obvious.

The representation (M, ρ) is called *irreducible* (resp. *completely reducible* or *semisimple*) if the $R[G]$-module M is simple (resp. semisimple), and it is called *faithful* if ρ is injective.

A $R[G]$-linear map is also called a *(R-linear) G-equivariant map* (or just a morphism of representations). A sub-$R[G]$-module is also called a *subrepresentation*. If R is clear from the context, we write Hom_G, End_G and Aut_G instead of $\mathrm{Hom}_{R[G]}$, $\mathrm{End}_{R[G]}$ and $\mathrm{Aut}_{R[G]}$.

The *regular representation* of G is the representation corresponding to the left regular $R[G]$-module.

Let $\varepsilon : R[G] \to R$, $\sum_{g \in G} \alpha_g g \longmapsto \sum_{g \in G} \alpha_g$. This is a surjective R-linear map of rings, called the *augmentation map*. Its kernel is the *augmentation ideal* of $R[G]$.

A representation of G on the regular R-module ${}_RR$ (that is, a morphism of monoids $G \to \mathrm{End}_R({}_RR) = R$) is sometimes called a *character* of G. This terminology is mostly

used when R is a field, and we will try to avoid it in these notes, because the word "character" also has another meaning in representation theory. (As we will see in chapter II.)

Theorem I.3.2. *Let R be a ring and G be a* finite *group. Then $R[G]$ is a semisimple ring if and only R is a semisimple ring and $|G|$ is invertible in R.*

If R is a field, this theorem is called *Maschke's theorem*.

Proof.

$\Leftarrow$: Let M be a $R[G]$-module, and let N be a $R[G]$-submodule of M. As R is semisimple, there exists a R-submodule N' of M such that $M = N \oplus N'$, so there exists a surjective R-linear map $f : M \to N$ such that $f_{|N} = \mathrm{id}_N$. Define $F : M \to N$ by :

$$F(v) = |G|^{-1} \sum_{g \in G} g^{-1} f(gv)$$

(here we use the fact that $|G|$ is invertible in R).

We will prove that F is $R[G]$-linear and that $F_{|N} = \mathrm{id}_N$. First, F is obviously R-linear. If $g \in G$ and $v \in M$, then

$$F(gv) = |G|^{-1} \sum_{h \in G} h^{-1} f(hgv) = |G|^{-1} g \sum_{h \in G} (hg)^{-1} f(hgv) = gF(v).$$

So F is indeed $R[G]$-linear. Next, let $v \in N$. Then for every $g \in G$, gv is also in N, so $f(gv) = gv$. Hence

$$F(v) = |G|^{-1} \sum_{g \in G} g^{-1} f(gv) = |G|^{-1} \sum_{g \in G} g^{-1} gv = v.$$

If we can show that $M = N \oplus \mathrm{Ker}(F)$, we will be done, because $\mathrm{Ker}(F)$ is a $R[G]$-sundmodule of M thanks to the $R[G]$-linearity of F. For every $v \in M$, we have $F(v) \in N$, hence $F(F(v)) = F(v)$, hence $F(v - F(v)) = 0$ and $v = (v - F(v) + F(v) \in \mathrm{Ker}(F) + N$. Moreover, if $v \in N \cap \mathrm{Ker}(F)$, then $v = F(v) = 0$. This finishes the proof that $M = N \oplus \mathrm{Ker}(F)$.

$\Rightarrow$: Assume that $R[G]$ is semisimple. Then we have the augmentation map $\varepsilon : R[G] \to R$ (see I.3.1). It makes R into a $R[G]$-module, which is automatically semisimple by assumption. As $R[G]$ acts on R through its quotient $R[G]/\operatorname{Ker}\varepsilon = R$, the R-module ${}_R R$ is semisimple, and so R is a semisimple ring.

As in the Artin-Wedderburn theorem (theorem I.1.10.5), write $R = M_{n_1}(\mathbb{D}_1) \times \cdots \times M_{n_r}(\mathbb{D}_r)$, with $\mathbb{D}_1, \ldots, \mathbb{D}_r$ division rings and $n_1, \ldots, n_r \geq 1$ integers. For every i, let $R_i = M_{n_i}(\mathbb{D}_i)$; we have a surjective morphism of rings $R[G] \to R_i$, and hence R_i is also semisimple (this is the same proof as in the previous

paragraph : ${}_{R_i}R_i$ is a semisimple R_i-module). Also, $|G|$ is invertible in R if and only if it is invertible in every R_i. For a fixed $i \in \{1, \dots, r\}$, $|G|$ is invertible in R_i if and only if it is invertible in $\mathbb{D}_i$, if and only if it is nonzero in $\mathbb{D}_i$, if and only it is nonzero in R_i.

So we may assume that $R = M_n(\mathbb{D})$ with $n \geq 1$ and $\mathbb{D}$ a division ring, and we are now trying to prove that $|G| \neq 0$ in R. Suppose that $|G| = 0$ in R, and let $x = \sum_{g \in G} g \in R[G]$. Then x is central in $R[G]$. Indeed, for every $y = \sum_{g \in G} \alpha_g g$, we have

$$yx = \sum_{g \in G} \sum_{g_1 g_2 = g} \alpha_{g_1} g = \left(\sum_{g \in G} \alpha_g \right) \left(\sum_{g \in G} g \right)$$

and

$$xy = \sum_{g \in G} \sum_{g_1 g_2 = g} \alpha_{g_2} g = \left(\sum_{g \in G} \alpha_g \right) \left(\sum_{g \in G} g \right).$$

Note that this does not use the fact that $|G| = 0$, but is true in any group algebra as long as G is finite (otherwise, x doesn't make sense).

Also,

$$x^2 = \left(\sum_{g \in G} 1 \right) \left(\sum_{g \in G} g \right) = |G| x = 0.$$

Let $I = R[G]x$. As $R[G]$ is semisimple, there exists a left ideal J of $R[G]$ such that $R[G] = I \oplus J$. By remark I.1.3.16, there exists $e \in I$ such that $e = e^2 \neq 0$. But we have $e = yx$ with $y \in R[G]$, so $e = e^2 = (yx)(yx) = y^2 x^2 = 0$, contradiction.

□

Remark I.3.3. If G is infinite and $R \neq 0$, then $R[G]$ is never semisimple.

Proof. Let I be the augmentation ideal of $R[G]$ (see I.3.1). Suppose that R is semisimple. Then there exists a left ideal J of R such that ${}_R R = I \oplus J$, and $J \neq 0$ because $I \ /R[G]$. Let $0 \neq b = \sum_{g \in G} \beta_g g$. For every $h \in G$, $(1 - h)b \in I \cap J$, so $(1 - h)b = 0$, i.e. $b = hb$. Hence $\beta_{hg} = \beta_g$ for every $g, h \in G$, which means that all the β_g are equal. As $b \neq 0$, at least one of the β_g is nonzero, so all the β_g are nonzero, and this is only possible if G is finite.

□

Let k be a field and G be a finite group. Using the Artin-Wedderburn theorem (theorem I.1.10.5, see also remark I.1.10.7) and theorem I.3.2, we get :

Theorem I.3.4. *1. There are uniquely determined k-division algebras $\mathbb{D}_1, \dots, \mathbb{D}_r$ and integers $n_1, \dots, n_r \geq 1$ such that $\dim_k(\mathbb{D}_i)$ is finite for every i and*

$$k[G]/\operatorname{rad}(k[G]) \simeq M_{n_1}(\mathbb{D}_1) \times \cdots \times M_{n_r}(\mathbb{D}_r).$$

A complete set of representatives of the isomorphisms classes of irreducible representations of G is given by $V_1 := \mathbb{D}_1^{n_1}, \dots, V_r := \mathbb{D}_r^{n_r}$, and we have a G-equivariant isomorphism

$$k[G]/\operatorname{rad}(k[G]) \simeq V_1^{n_1} \oplus \cdots \oplus V_r^{n_r}.$$

2. If k is algebraically closed, then $\mathbb{D}_i = k$ for every i, and so

$$\sum_V (\dim_k V)^2 = \sum_{i=1}^r \dim_k(V_i)^2 = \sum_{i=1}^r n_i^2 = \dim_k(k[G]/\operatorname{rad}(k[G])) \leq \dim_k(k[G]) = |G|,$$

where the first sum is taken over the isomorphism classes of irreducible representations of G. This inequality is an equality if and only if the characteristic of k does not divide $|G|$.

Definition I.3.5. If R is a ring and G is a group, we denote by $S_R(G)$ the set of isomorphism classes of irreducible representations of G on R-modules.

Remark I.3.6. By the theorem above, $S_k(G)$ is finite if k is a field and G is a finite group.

Example I.3.7.

- Fix R and G as above. The trivial representation of G (over R) is the representation of G on ${}_RR$ given by the augmentation map $R[G] \to R$ (see I.3.1), i.e. by the trivial action of G on R. We denote it by $\mathbb{1}$.
- If G is the symmetric group $\mathfrak{S}_n$ and R is any ring, we denote by sgn the representation of G on R given by the sign morphism $G \to \{\pm 1\}$ composed with the obvious map $\{\pm 1\} \to R$.

Example I.3.8.

(1) Let $G = \mathfrak{S}_2$ and k be a field.

If $\operatorname{char}(k) \neq 2$, then $k[G] \simeq k \times k$, where the first factor corresponds to the trivial representation $\mathbb{1}$ and the second to the sign representation sgn.

If $\operatorname{char}(k) = 2$, then $\mathbb{1} = \operatorname{sgn}$ is the unique simple $k[G]$-module, $\operatorname{rad}(k[G])$ is equal to the augmentation ideal of $k[G]$ (see I.3.1), and we have $k[G]/\operatorname{rad}(k[G]) = k$.

(2) Let $G = \mathfrak{S}_3$ and k be a field such that $\operatorname{char}(k) \not| 6$. We know that $k[G]$ is a semisimple.

The only 1-dimensional representations of $\mathfrak{S}_3$ (corresponding to the morphisms of groups $\mathfrak{S}_3 \to k^\times$) are $\mathbb{1}$ and sgn, and they are nonisomorphic because $\operatorname{char}(k) \neq 2$.

Make $\mathfrak{S}_3$ act on k^3 by

$$\sigma \cdot (x_1, x_2, x_3) = (x_{\sigma^{-1}(1)}, x_{\sigma^{-1}(2)}, x_{\sigma^{-1}(3)}).$$

Let
$$V_2 = \{(x_1, x_2, x_3) \in k^3 | x_1 + x_2 + x_3 = 0\}.$$

Then $k^3 = V_2 \oplus \mathbb{1}$, and V_2 is an irreducible subrepresentation of k^3. Indeed, if V_2 were not simple, it would be a sum of two 1-dimensional subrepresentations, so we just need to see that neither $\mathbb{1}$ not sgn are isomorphic subrepresentations of V_2. Let $v = (x, y, -x-y) \in V_2$, and suppose that $\sigma \cdot v = v$ for every $\sigma \in \mathfrak{S}_3$. Then $x = y = -x - y$, so $3x = 3y = 0$, so $x = y = 0$ as $\mathrm{char}(k) \neq 3$, and hence $v = 0$. So V_2 has no subrepresentation isomorphic to $\mathbb{1}$. Now let $v = (x, y, -x-y) \in V_2$, and suppose that $\sigma \cdot v = \mathrm{sgn}(\sigma)v$ for every $\sigma \in \mathfrak{S}_3$. Then $x = -y = x+y$, so $x = y = 0$, so $v = 0$. So V_2 has no subrepresentation isomorphic to sgn.

We found three irreducible representations of $\mathfrak{S}_2$ of dimensions 1, 1 and 2. As $1^2 + 1^1 + 2^2 = 6 = |G|$, there are no other irreducible representations of G, and we have
$$k[G] \simeq \mathrm{End}_k(\mathbb{1}) \times \mathrm{End}_k(\mathrm{sgn}) \times \mathrm{End}_k(V_2) \simeq k \times k \times M_2(k)$$
as k-algebras.
Remark. If $\mathrm{char}(k) = 2$, then V_2 is still an irreducible representation of $\mathfrak{S}_3$, but we now have $\mathbb{1} = \mathrm{sgn}$. The Jacobson radical $\mathrm{rad}(k[G])$ is 1-dimensional over k, and we have $k[G]/\mathrm{rad}(k[G]) \simeq k \times M_2(k)$.

If $\mathrm{char}(k) = 3$, then $\mathbb{1} \not\simeq \mathrm{sgn}$, but V_2 is not irreducible anymore. In fact, we have an exact sequence of $k[G]$-modules $0 \to \mathbb{1} \to V_2 \to \mathrm{sgn} \to 0$. The Jacobson radical $\mathrm{rad}(k[G])$ is 4-dimensional over k, and we have $k[G]/\mathrm{rad}(k[G]) \simeq k \times k$.

(3) Let G be the quaternion group $Q = \{\pm 1, \pm i, \pm j, \pm j\} \subset \mathbb{H}$, with the mutliplication given by that of $\mathbb{H}$. (See problem VII.1.8.)

In problem VII.1.8, the following facts are proved :
$$\mathbb{R}[G] \simeq \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{H},$$
and so $S_{\mathbb{R}}(G)$ has 5 elements. More precisely, the elements of $S_{\mathbb{R}}(G)$ are :

- The trivial representation of G on $\mathbb{R}$.
- The representation of G on $\mathbb{R}$ given by the map $\begin{cases} G & \to & \mathbb{R} \\ \pm 1, k & \longmapsto & 1 \\ i, j & \longmapsto & -1 \end{cases}$
- The representation of G on $\mathbb{R}$ given by the map $\begin{cases} G & \to & \mathbb{R} \\ \pm 1, i & \longmapsto & 1 \\ j, k & \longmapsto & -1 \end{cases}$
- The representation of G on $\mathbb{R}$ given by the map $\begin{cases} G & \to & \mathbb{R} \\ \pm 1, j & \longmapsto & 1 \\ i, k & \longmapsto & -1 \end{cases}$

- The (4-dimensional) representation of G on $\mathbb{H}$ given by the inclusion $G \subset \mathbb{H}^\times$ and the action of $\mathbb{H}^\times$ on $\mathbb{H}$ by multiplication on the left.

Also,

$$\mathbb{C}[G] \simeq \mathbb{C} \times \mathbb{C} \times \mathbb{C} \times \mathbb{C} \times M_2(\mathbb{C}),$$

and so $S_{\mathbb{C}}(G)$ has 5 elements. The first four are 1-dimensional and are just the tensor products by $\mathbb{C}$ of the four 1-dimensional representations of G over $\mathbb{R}$. The fifth is 2-dimensional : to construct, we use the $\mathbb{C}$-algebra isomorphism $\mathbb{H} \otimes_{\mathbb{R}} \mathbb{C} \simeq M_2(\mathbb{C})$, which splits the 4-dimensional representation of G on $\mathbb{H} \otimes_{\mathbb{R}} \mathbb{C}$ (coming from the 4-dimensional simple $\mathbb{R}[G]$-module) into two isomorphic irreducible 2-dimensional representations.

Proposition I.3.9. *If G is a finite abelian group and k is an algebraically closed field, then every simple $k[G]$-module is of dimension 1 over k.*

Proof. Let V be a simple $k[G]$-module. Then any nonzero element v of V gives a surjective k-linear map $k[G] \to V$ (sending $x \in k[G]$ to xv), and so in particular V is a finite-dimensional vector space, and so is $\mathrm{End}_{k[G]}(V)$. The finite-dimensional k-algebra $\mathrm{End}_{k[G]}(V)$ is also a k-division algebra by Schur's lemma (theorem I.1.4.1), so $\mathrm{End}_{k[G]}(V) = k$ by problem VII.1.3. Moreover, as G is abelian, $k[G]$ is a commutative ring, so the action of any element of $k[G]$ on V is $k[G]$-linear. This means that the action of $k[G]$ on V is given by a map of k-algebras $k[G] \to \mathrm{End}_{k[G]}(V) = k \subset \mathrm{End}_k(V)$. So for any $v \in V$, the subspace kv is a $k[G]$-module of V. As V is irreducible, this implies that $\dim_k(V) = 1$.

□

Remark I.3.10. The proposition above is false in general for infinite groups. For example, if $k = \mathbb{C}$ and $G = \mathbb{C}(T)^\times$, then $\mathbb{C}(T)$ is a simple $\mathbb{C}[G]$-module with the obvious action of G by mutliplication, but it is not 1-dimensional over $\mathbb{C}$.

The reason for this is that, if G is infinite, then we cannot conclude from Schur's lemma that the algebra of endomorphisms of a simple $k[G]$-module if equal to k. However, if for example k is algebraically closed and uncountable, then we still have $\mathrm{End}_{k[G]}(V) = k$ for any simple $k[G]$-module V provided that either G is countable or $\dim_k V$ is countable. [11] So for example, if we suppose that G is commutative and countable, and that k is algebraically closed and uncountable, we can deduce that every simple $k[G]$-module is 1-dimensional over k.

Example I.3.11. The proposition above is also false if the field k is not algebraically closed. For example, take $G = \{\pm 1, \pm i\} \subset \mathbb{C}^\times$ and $k = \mathbb{R}$. Then $\mathbb{C}$, with the obvious action of G, is a 2-dimensional irreducible representation of G over $\mathbb{R}$.

[11] See lemma 2.11 of the book [3] of Bernstein and Zelevinski. The proof goes as follows : Let $u \in \mathrm{End}_{k[G]}(V)$, suppose that $u \notin k \cdot \mathrm{id}_V$. For every $\lambda \in k$, $u - \lambda \mathrm{id}_V$ is G-equivariant and nonzero, so it is invertible by Schur's lemma; let $v_\lambda = (u - \lambda \mathrm{id}_V)^{-1}$. Now choose $x \in V - \{0\}$. Then it is an easy exercise to show that the family $(v_\lambda(x))_{\lambda \in k}$ is linearly independent. If we assumed that $\dim_k V$ is countable, we get a contradiction. If we assumed that G is countable, then this forces $\dim_k V$ to be countable (because we have a surjective k-linear map $k[G] \to V$, $a \longmapsto ax$), so we also get a contradiction.

Example I.3.12. Let $G = \mathbb{Z}/n\mathbb{Z}$ ($n \geq 1$), and let k be an algberaically closed field.

- The case $\mathrm{char}(k) \not| n$: Fix a primitive nth root ζ_n of 1 in k. The k-algebra $k[G]$ is semisimple, so $k[G] \simeq k \times \cdots \times k$, where the n factors k correspond to the n irreducible representations of G given by the maps $G = \mathbb{Z}/n\mathbb{Z} \to k^\times$, $1 \longmapsto \zeta_n^i$, for $1 \leq i \leq n$.
- The case $\mathrm{char}(k)|n$: Let $p = \mathrm{char}(k)$, and write $n = p^r m$, with m prime to p. Then $G = \mathbb{Z}/p^r\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$. To give a 1-dimensional representation of G, we have to give the images $a, b \in k^\times$ of $(1,0), (0,1) \in \mathbb{Z}/p^r\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ by the map $G \to k^\times$ corresponding to the representation. We must have $a^{p^r} = 1$, hence $a = 1$ because $\mathrm{char}(k) = p$, and b can be any of the m solutions of the equation $x^m = 1$ in k. So we get m irreducible representations of G, and $k[G]/\mathrm{rad}(k[G]) \simeq k^m$.

I.4 The representation ring

Definition I.4.1. Let R be a ring.

1. We define $K(R)$ to be the quotient of the free abelian group on the basis elements $[M]$, for M a finite length[12] R-module, by all the relations of the form $[M] = [M'] + [M'']$, where $0 \to M' \to M \to M'' \to 0$ is an exact sequence of R-modules.
2. We define $PK(R)$ to be the quotient of the free abelian group on the basis elements $[P]$, for P a finite length projective R-module, by all the relations of the form $[P] = [P'] + [P'']$, where $0 \to P' \to P \to P'' \to 0$ is an exact sequence of R-modules.

[13]

We have an obvious map $PK(R) \to K(R)$, which is neither injective nor surjective in general.

Remark I.4.2. The group $K(R)$ (resp. $PK(R)$) is usually called the Grothendieck group of the category of finite length (resp. finite length projective) R-modules.

Remark I.4.3. Here are some easy properties of $K(R)$ and $PK(R)$:

(1) The exact sequence $0 \to 0 \to 0 \to 0 \to 0$ gives an equality $[0] = [0] + [0]$ in $K(R)$ and $PK(R)$, so we get $[0] = 0$ in these groups.

(2) If we have an isomorphism of (projective) R-modules $M \xrightarrow{\sim} M'$, then the sequence $0 \to M \xrightarrow{\sim} M' \to 0$ is exact, so $[M] = [M'] + [0] = [M']$ in $K(R)$ (and $PK(R)$).

(3) If $M = M_0 \supset M_1 \supset \cdots \supset M_n = 0$ are R-modules, then $[M] = \sum_{i=1}^n [M_{i-1}/M_i]$ in $K(R)$, by an easy induction on n.

Proposition I.4.4. *As a group, $K(R)$ is the free abelian group with basis $\{[M],\ V \in S(R)\}$ where $S(R)$ is the set of isomorphism classes of simple R-modules (as in definition I.1.7.1).*

[12] See definition I.1.5.1.

[13] Unfortunately, the notation is totally ad hoc.

We will prove a similar statement for $PK(R)$ in corollary III.2.1 of chapter III.

Proof. Let S be the free abelian group on the set $S(R)$, and denote by $(e_M)_{M\in S(R)}$ its canonical basis.

Define $\varphi : S \to K(R)$ by $\varphi(e_M) = [M]$.

Define $\psi : K(R) \to S$ as follows : If M is a R-module of finite length, then it has a Jordan-Hölder series $M = M_0 \supset M_1 \cdots \supset M_n = 0$ by definition, and we set

$$\psi([M]) = \sum_{i=1}^{n} e_{M_{i-1}/M_i}.$$

By theorem I.1.5.2, this does not depend on the choice of the Jordan-Hölder series.

Let's show that ψ is well-defined. If $0 \to M' \to M \xrightarrow{u} M'' \to 0$ is an exact sequence of finite length R-modules, we need to show that $\psi([M]) = \psi([M']) + \psi([M''])$. Choose Jordan-Hölder series $M' = M'_0 \supset \cdots \supset M'_n = 0$ and $M'' = M''_0 \supset \cdots \supset M''_m = 0$. For $0 \leq j \leq m$, let $M_j = u^{-1}(M''_j)$. Then $M = M_0 \supset \cdots \supset M_m = M'_0 \supset \cdots \supset M'_n = 0$ is a Jordan-Hölder series for M, which gives the desired equality immediately.

It is now obvious that φ and ψ are inverses of each other.

□

Proposition I.4.5. *If R is a semisimple ring, then $K(R) = PK(R)$, and, for all finite length R-modules M, M', we have $[M] = [M']$ in $K(R)$ if and only if $M \simeq M'$.*

Proof. If R is a semisimple ring, then every R-module is projective, so $K(R) = PK(R)$.

Let M and M' be two R-modules of finite length. As R is semisimple, we can write $M \simeq \bigoplus_{N\in S(R)} N^{\oplus r_N}$ and $M' \simeq \bigoplus_{N\in S(R)} N^{\oplus r'_N}$. Then $[M] = \sum_{N\in S(R)} r_N[N]$ and $[M'] = \sum_{N\in S(R)} r'_N[N]$. By proposition I.4.4, we have $[M] = [M']$ if and only if $r_N = r'_N$ for every $N \in S(R)$, which is equivalent to $M \simeq M'$.

□

We now apply this to representations of groups. Let k be a field, and let G be a group.

Definition I.4.6. We write $\mathrm{R}_k(G) = K(k[G])$, $\mathrm{P}_k(G) = PK(k[G])$ and $S_k(G) = S(k[G])$ (that last notation was already introduced in definition I.3.5).

We call $\mathrm{R}_k(G)$ is called the *representation ring* of G over the field k.

The name "representation ring" is explained by the following fact :

Proposition I.4.7. *We have a multiplication on $\mathrm{R}_k(G)$ given $[M][M'] = [M \otimes_k M']$, and this makes $\mathrm{R}_K(G)$ into a commutative ring, with unit element equal to $[\mathbb{1}]$.*

Proof. As tensoring over a field preserves equal sequences, the formula $[M][M'] = [M \otimes_k M']$ does define a bi-additive map $\mathrm{R}_k(G) \times \mathrm{R}_k(G) \to \mathrm{R}_k(G)$. Everything else is clear.

□

Proposition I.4.8. *If M is a $k[G]$ and N is a projective $k[G]$-module, then the $k[G]$-module $M \otimes_k N$ is also projective.*

Proof. As N is projective over $k[G]$, it is a direct summand of some free $k[G]$-module $k[G]^{\oplus I}$, [14] and then $M \otimes_k N$ is a direct summand of $M \otimes_k k[G]^{\oplus I} = (M \otimes_k k[G])^{\oplus I}$. So it is enough to show that $M \otimes_k k[G]$ is projective.

We will actually show that the $k[G]$-module $M \otimes_k k[G]$ is free over $k[G]$. Let $\underline{M}$ be the k-vector space M, considered as a representation of G with the trivial action. Define a k-linear map $u : \underline{M} \otimes_k k[G] \to M \otimes_k k[G]$ by sending $x \otimes g$ to $gx \otimes g$, if $x \in \underline{M}$ and $g \in G$. Then the k-linear map $v : M \otimes_k k[G] \to \underline{M} \otimes_k k[G]$ that sends $x \otimes g$ to $g^{-1}x \otimes g$ for $x \in M$ and $g \in G$ is an inverse of u, so u is an isomorphism of k-modules. Let's show that u is G-equivariant. Let $g, h \in G$ and $x \in \underline{x}$. Then :

$$h \cdot u(x \otimes g) = h \cdot (gx \otimes g) = (hgx) \otimes (hg)$$

and

$$u(h \cdot (x \otimes g)) = u(x \otimes (hg)) = (hgx) \otimes (hg)$$

are equal. Finally, we have found an isomorphism of $k[G]$-modules $M \otimes_k k[G] \simeq \underline{M} \otimes_k k[G]$. As $\underline{M} \otimes_k k[G]$ is just a (possibly infinite) direct sum of copies of $k[G]$, the $k[G]$-module $M \otimes_k k[G]$ is free.

□

Corollary I.4.9. *The tensor product over k induces a bi-additive map $\mathrm{R}_k(G) \times \mathrm{P}_k(G) \to \mathrm{P}_k(G)$, which makes $\mathrm{P}_k(G)$ into a $\mathrm{R}_K(G)$-module, and the obvious map $\mathrm{P}_k(G) \to \mathrm{R}_k(G)$ is $\mathrm{R}_k(G)$-linear.*

Remark I.4.10. If G is a finite group, then a $k[G]$-module V has finite length if and only if it is a finite-dimensional k-vector space.

Proof. If $\dim_k(V)$ is finite, then V is Artinian and Noetherian as a k-module, so it is Artinian and Noetherian as a $k[G]$-module, and hence has finite length by proposition I.1.6.2.

Conversely, suppose that V is a finite length $k[G]$-module. Then V has a Jordan-Hölder series $V = V_0 \supset \cdots \supset V_n = 0$. We have $\dim_k(V) = \sum_{i=1}^{n} \dim_k(V_{i-1}/V_i)$, and each V_{i-1}/V_i is a simple $k[G]$-module, so we only need to prove that every simple $k[G]$-module is a finite-dimensional k-vector space. But we already saw this : a simple $k[G]$-module is a quotient of the right regular module $k[G]$, and $\dim_k(k[G]) = |G|$ is finite.

□

[14] The free $k[G]$-module with basis I.

Remark I.4.11. If G is a finite group and K/k is a field extension, then $[V] \longmapsto [V \otimes_k K]$ induces a morphism of rings $\mathrm{R}_k(G) \to \mathrm{R}_K(G)$.

This morphism is injective if $\mathrm{char}(k) = 0$ or if K/k is a separable algebraic extension, [15] but it is not always surjective. We give counterexamples below.

Example I.4.12.

(1) Take $G = \mathbb{Z}/n\mathbb{Z}$ and k a field containing all the nth roots of unity and such that $\mathrm{char}(k) \not| n$. Then $\mathrm{R}_k(G) \simeq \mathbb{Z}^n$ as a group, and $\mathrm{R}_k(G)$ is isomorphic to the group algebra $\mathbb{Z}[\mathbb{Z}/n\mathbb{Z}]$ as a ring. (See example I.3.12.)

(2) Take $G = \{\pm 1, \pm i\} \subset \mathbb{C}^\times$. Then, as groups, $\mathrm{R}_{\mathbb{R}}(G) \simeq \mathbb{Z}^3$ (see example I.3.11) and $\mathrm{R}_{\mathbb{C}}(G) \simeq \mathbb{Z}^4$ (see proposition I.3.9). The map $\mathrm{R}_{\mathbb{R}}(G) \to \mathrm{R}_{\mathbb{C}}(G)$ is $(a, b, c) \longmapsto (a, b, c, c)$, and it is not surjective.

(3) If $G = \mathbb{Z}/p\mathbb{Z}$ and k is a field of characteristic p, then $\mathrm{R}_k(G) \simeq \mathbb{Z}$ as a ring, because the only simple $k[G]$-module is the trivial representation. (See example I.3.12.)

(4) If $G = \{\pm 1, \pm i, \pm j, \pm k\} \subset \mathbb{H}^\times$ as in example I.3.8(3), then, using the calculations of this example, we get isomorphisms of groups $\mathrm{R}_{\mathbb{R}}(G) \simeq \mathbb{Z}^5$ and $\mathrm{R}_{\mathbb{C}}[G] \simeq \mathbb{Z}^5$, and we see that the map $\mathrm{R}_{\mathbb{R}}[G] \to \mathrm{R}_{\mathbb{C}}[G]$ is given by $(a, b, c, d, e) \longmapsto (a, b, c, d, 2e)$. This is also not surjective.

I.5 Induction and restriction

Let R be a ring and $\alpha : H \to G$ be a morphism of groups. Then α a morphism of rings $\varphi : R[H] \to R[G]$. In many applications, G is finite, α is the inclusion of a subgroup of G, and R is a field, but we'll do as much as we can in the general setting.

I.5.1 Definitions

Definition I.5.1.1 (Restriction)**.** If M is a $R[G]$-module, we can see it as a $R[H]$-module by making $R[H]$ act via the morphism $\varphi : R[H] \to R[G]$. The resulting $R[H]$-module is denoted by $\mathrm{Res}^G_H M$ (or $\alpha^* M$ if we need to make α explicit) and called the restriction of M to H (along α).

If $u : M \to N$ is a morphism of $R[G]$-modules, we write $\mathrm{Res}^G_H(u) : \mathrm{Res}^G_H M \to \mathrm{Res}^G_H N$ for the same map u, now seen as a morphism of $R[H]$-modules.

Remark I.5.1.2. It is clear from the definition that Res^G_H preserves exact sequences.

[15] See problem VII.2.1.

Hence, if $\alpha(H)$ has finite index in G and k is a field, Res^G_H induces morphisms of groups $\mathrm{Res}^G_H : \mathrm{R}_k(G) \to \mathrm{R}_k(G)$ and $\mathrm{P}_k(G) \to \mathrm{P}_k(H)$ (see definition I.4.6), and the first morphism is actually a morphism of rings. (The condition on α is needed to preserve the finite length condition on the representations appearing in the definition of the representation rings.)

Definition I.5.1.3 (Induction)**.** If M is a $R[H]$-module, we set

$$\mathrm{Ind}^G_H M = R[G] \otimes_{R[H]} M,$$

where $R[H]$ acts on the left on $R[G]$ via the morphism φ. This is called the *induction* of M from H to G, and is sometimes also denoted by $\alpha_! M$.

If $u : M \to N$ is a morphism of $R[H]$-modules, we write $\mathrm{Ind}^G_H(u) : \mathrm{Ind}^G_H M \to \mathrm{Ind}^G_H N$ for the $R[G]$-linear map $\mathrm{id}_{R[G]} \otimes u$.

Definition I.5.1.4 (Coinduction)**.** If M is a $R[H]$-module, we set

$$\mathrm{CoInd}^G_H M = \mathrm{Hom}_{R[H]}(R[G], M),$$

where $R[G]$ is seen as a left $R[H]$-module via $\varphi : R[H] \to R[G]$. We make this into a left $R[G]$-module using the right regular action of $R[G]$ on itself. More concretely, if $x \in R[G]$ and $u \in \mathrm{CoInd}^G_H M$, then $x \cdot u$ is defined by $(x \cdot u)(y) = u(yx)$, for every $y \in R[H]$.

The $R[G]$-module $\mathrm{CoInd}^G_H M$ is called the *coinduction* of M from H to G, and it is sometimes denoted by $\alpha_* M$.

If $u : M \to N$ is a morphism of $R[H]$-modules, we write $\mathrm{CoInd}^G_H(u) : \mathrm{CoInd}^G_H M \to \mathrm{CoInd}^G_H N$ for the $R[G]$-linear map sending $v \in \mathrm{Hom}_{R[H]}(R[G], M)$ to $u \circ v$.

Remark I.5.1.5. Here is a more concrete description of the coinduction. Let M be a $R[H]$-module. Then restricting maps $R[G] \to M$ along the inclusion $G \subset R[G]$ induces an isomorphism of R-modules

$$\mathrm{CoInd}^G_H M \xrightarrow{\sim} \{f : G \to M | \forall h \in H, \forall g \in G,\ f(\alpha(h)g) = hf(g)\}.$$

The action of G on the right-hand side is given in the following way : If $f : G \to M$ and if $x \in G$, then $x \cdot f : G \to M$ is defined by $(x \cdot f)(g) = f(gx)$, for every $g \in G$.

Proof. Denote by ψ the R-module morphism defined above. It is injective because G generates the $R[H]$-module $R[G]$, so a $R[H]$-linear map $u : R[G] \to M$ is uniquely determined by its restriction to G.

Let's show that ψ is surjective. Let $f : G \to M$ satisfying the condition in the formula above, and define a R-linear map $u : R[G] \to M$ by

$$u(\sum_{g \in G} c_g g) = \sum_{g \in G} c_g f(g).$$

If we can show that u is $R[H]$-linear, we'll be done because we'll then have $\psi(u) = f$. But if $h \in H$ and $x = \sum_{g\in G} c_g g \in R[G]$, then

$$u(hx) = u(\sum_{g\in G} c_g \alpha(h)g) = \sum_{g\in G} c_g f(\alpha(h)g) = \sum_{g\in G} c_g h f(g) = hu(x).$$

The last sentence of the remark is clear.

□

The analogue of remark I.5.1.2 for induction and coinduction is more complicated (and not always true). We will consider this question below.

I.5.2 Induction and exact sequences

The following obvious result will be used several times :

Proposition I.5.2.1. *If $\alpha : H \to G$ is injective, then $\varphi : R[H] \to R[G]$ makes $R[G]$ into a free (left or right) $R[H]$-module. More precisely, let $(g_i)_{i\in I}$ be a complete set of representatives of $\alpha(H) \subset G$ (resp. $G/\alpha(H)$) in G; then $(g_i)_{i\in I}$ is a basis of $R[G]$ as a left (resp. right) $R[H]$-module.* □

Definition I.5.2.2. Suppose that $G = \{1\}$. Then, for every $R[H]$-module M, $\mathrm{Ind}_H^1 M$ is called the R-module of *coinvariants* of M under H and denoted by M_H.

Remark I.5.2.3. It follows directly from the definition that M_H is the quotient of M by the R-submodule generated by all $hm - m$, with $h \in H$ and $m \in M$.

Theorem I.5.2.4. *For every exact sequence $0 \to M' \to M \to M'' \to 0$ of $R[H]$-modules, the sequence $\mathrm{Ind}_H^G M' \to \mathrm{Ind}_H^G M \to \mathrm{Ind}_H^G M'' \to 0$ is exact.*

Moreover, we can also deduce that the sequence $0 \to \mathrm{Ind}_H^G M' \to \mathrm{Ind}_H^G M \to \mathrm{Ind}_H^G M'' \to 0$ is exact in the following two situations :

1. *$\alpha : H \to G$ is injective.*
2. *$\mathrm{Ker}(\alpha)$ is finite, R is semisimple and $|\mathrm{Ker}(\alpha)|$ is invertible in R.*

Lemma I.5.2.5. *Let M be a $R[H]$-module.*

1. *If $\alpha : H \to G$ is surjective, then $\mathrm{Ind}_H^G = M_{\mathrm{Ker}\,\alpha}$, with the following action of G : For every $g \in G$ and $x \in M_{\mathrm{Ker}\,\alpha}$, choose preimage $h \in H$ of g by α and $m \in M$ of x by the obvious quotient map, and then gx is the image in $M_{\mathrm{Ker}\,\alpha}$ of hm.*
2. *In general, $\mathrm{Ind}_H^G = \mathrm{Ind}_{\alpha(H)}^G M_{\mathrm{Ker}\,\alpha}$.*

Proof. 1. Let $u : \mathrm{Ind}_H^G M = R[G] \otimes_{R[H]} M \to M_{\mathrm{Ker}\,\alpha}$ be defined as follows : If $a \in R[G]$ and $m \in, M$, choose $b \in R[H]$ such that $\varphi(b) = a$ and take for $u(a \otimes m)$ the image of bm in $M_{\mathrm{Ker}\,\alpha}$. This does not depend on the choice of b, because $\mathrm{Ker}\,\varphi$ is the $R[H]$-submodule of $R[H]$ generated by all $h - 1$, for $h \in \mathrm{Ker}\,\alpha$, and it does define a map on $R[G] \otimes_{R[H]} M$, because the formula for $u(a \otimes m)$ is additive in a and m and takes the same value on $(a\varphi(x), m)$ and (a, xm) if $x \in R[H]$.

Let $v' : M \to R[G] \otimes_{R[H]} M$ be the map $m \longmapsto 1 \otimes m$. If $m \in M$ and $h \in \mathrm{Ker}\,\alpha$, then

$$v'(hm - m) = 1 \otimes (hm) - 1 \otimes m = \alpha(h) \otimes m - 1 \otimes m = 0.$$

So v' defines a R-linear map $v : M_{\mathrm{Ker}\,\alpha} \to \mathrm{Ind}_H^G M$.

It is now very easy to check that u and v are inverses of each other, and to read the formula for the action of G on $M_{\mathrm{Ker}\,\alpha}$ on the definition of u.

2. This follows from the fact that $\mathrm{Ind}_H^G M = \mathrm{Ind}_{\alpha(H)}^G \mathrm{Ind}_H^{\alpha(H)}$ (which is just the transitivity of $\otimes$) and from (i).

□

Lemma I.5.2.6. *Suppose that R is semisimple, and that G is a finite group such that $|G|$ is invertible in R. Then, for every exact sequence of $R[G]$-modules $0 \to M' \to M \to M'' \to 0$, the sequence $0 \to M'_G \to M_G \to M''_G \to 0$ is also exact.*

Proof. The hypotheses imply that the ring $R[G]$ is semisimple (see theorem I.3.2). By the first part of the theorem (whose proof does not use this lemma), we only need to show that taking coinvariants preserves injectivity in our situation. So let $M \to N$ be an injective $R[G]$-linear map. As $R[G]$ is semisimple, there exists a $R[G]$-submodule M' of N such that $N = M \oplus M'$. It's clear on the definition of coinvariants that $N_G = M_G \oplus M'_G$. In particular, the map $M_G \to N_G$ is injective.

□

Proof of the theorem. The first part follows from the general properties (more precisely, the right exactness) of the tensor product.

Let's prove the second part. In situation (i), the right $R[H]$-module $R[G]$ is free by proposition I.5.2.1, so taking tensor products by $R[G]$ over $R[H]$ preserves exact sequences. Suppose that we are in situation (ii). Then, by lemma I.5.2.5, $\mathrm{Ind}_H^G M = \mathrm{Ind}_{\alpha(H)}^G M_{\mathrm{Ker}\,\alpha}$ for every $R[H]$-module M, and so the statement follows from lemma I.5.2.6 and from situation (i).

□

Remark I.5.2.7. The second part of the theorem is not true in general, because taking coinvariants does not always preserve exact sequences. [16]

[16] See any book on group homology, for example Brown's book [5].

Corollary I.5.2.8. *Suppose that $R = k$ is a field. If G and H are finite and $\mathrm{char}(k)$ does not divide $|\mathrm{Ker}(\alpha)|$, then Ind_H^G induces a morphism of groups $\mathrm{R}_k(H) \to \mathrm{R}_k(G)$.*

Remark I.5.2.9. Note that this is not a morphism of rings. For example, it sends the unit $[\mathbb{1}]$ of $\mathrm{R}_k(H)$ to $[k[G] \otimes_{k[\alpha(H)]} k]$. [17]

I.5.3 Coinduction and exact sequences

Definition I.5.3.1. Suppose that $G = \{1\}$. Then, for every $R[H]$-module M, $\mathrm{CoInd}_H^1 M$ is called the R-module of *invariants* of M under H and denoted by M^H.

Remark I.5.3.2. It follows directly from the definition that

$$M^H = \{m \in M | \forall h \in H,\ hm = m\}.$$

Theorem I.5.3.3. *For every exact sequence $0 \to M' \to M \to M'' \to 0$ of $R[H]$-modules, the sequence $0 \to \mathrm{CoInd}_H^G M' \to \mathrm{CoInd}_H^G M \to \mathrm{CoInd}_H^G M''$ is exact.*

Moreover, we can also deduce that the sequence $0 \to \mathrm{CoInd}_H^G M' \to \mathrm{CoInd}_H^G M \to \mathrm{CoInd}_H^G M'' \to 0$ is exact in the following two situations :

1. *$\alpha : H \to G$ is injective.*
2. *$\mathrm{Ker}(\alpha)$ is finite, R is semisimple and $|\mathrm{Ker}(\alpha)|$ is invertible in R.*

Lemma I.5.3.4. *Let M be a $R[H]$-module.*

1. *If $\alpha : H \to G$ is surjective, then $\mathrm{CoInd}_H^G = M^{\mathrm{Ker}\,\alpha}$, with the following action of G : For every $g \in G$ and $m \in M^{\mathrm{Ker}\,\alpha}$, choose a preimage $h \in H$ of g by α, and then gm is defined to be hm.*
2. *In general, $\mathrm{CoInd}_H^G = \mathrm{CoInd}_{\alpha(H)}^G M^{\mathrm{Ker}\,\alpha}$.*

Proof. 1. Let $u : \mathrm{CoInd}_H^G M = \mathrm{Hom}_{R[H]}(R[G], M) \to M^{\mathrm{Ker}\,\alpha}$ be defined as follows : If $f : R[G] \to M$ is a $R[H]$-linear map, take $u(f) = f(1)$. This is well-defined because, for every $h \in \mathrm{Ker}\,\alpha$, $hf(1) = f(\alpha(h)1) = f(1)$.

Let $v : M^{\mathrm{Ker}\,\alpha} \to \mathrm{Hom}_{R[H]}(R[G], M)$ be the map defined as follows : If $m \in M$ and $a \in R[G]$, choose $b \in R[H]$ such that $\varphi(b) = a$, and set $v(m)(a) = bm$. This does not depend on the choice of b, because $\mathrm{Ker}\,\varphi$ is the $R[H]$-submodule of $R[H]$ generated by all $h - 1$, for $h \in \mathrm{Ker}\,\alpha$.

It is now very easy to check that u and v are inverses of each other, and the statement about the action of G on $M^{\mathrm{Ker}\,\alpha}$ is obvious.

[17] But see corollary I.5.6.2 for a property that it does have.

2. This follows from the fact that $\mathrm{CoInd}_H^G M = \mathrm{CoInd}_{\alpha(H)}^G \mathrm{CoInd}_H^{\alpha(H)}$ (i.e. that $\mathrm{Hom}_{R[H]}(R[G], M) = \mathrm{Hom}_{R[\alpha(H)]}(R[G], \mathrm{Hom}_{R[H]}(R[\alpha(H)], M))$, which is a general property of Hom) and from (i).

□

Lemma I.5.3.5. *Suppose that R is semisimple, and that G is a finite group such that $|G|$ is invertible in R. Then, for every exact sequence of $R[G]$-modules $0 \to M' \to M \to M'' \to 0$, the sequence $0 \to M'^G \to M^G \to M''^G \to 0$ is also exact.*

Proof. The hypotheses imply that the ring $R[G]$ is semisimple (see theorem I.3.2). By the first part of the theorem (whose proof does not use this lemma), we only need to show that taking invariants preserves surjectivity in our situation. So let $M \to N$ be an surjective $R[G]$-linear map. As $R[G]$ is semisimple, there exists a $R[G]$-submodule N' of M such that $M = N \oplus N'$. It's clear on the definition of invariants that $M^G = N^G \oplus N'^G$. In particular, the map $M^G \to N^G$ is surjective.

□

Proof of the theorem. The first part follows from the general properties (more precisely, the left exactness) of Hom.

Let's prove the second part. In situation (i), the left $R[H]$-module $R[G]$ is free by proposition I.5.2.1, so taking $\mathrm{Hom}_{R[H]}(R[G], \cdot)$ preserves exact sequences. Suppose that we are in situation (ii). Then, by lemma I.5.3.4, $\mathrm{CoInd}_H^G M = \mathrm{CoInd}_{\alpha(H)}^G M^{\mathrm{Ker}\,\alpha}$ for every $R[H]$-module M, and so the statement follows from lemma I.5.3.5 and from situation (i).

□

Remark I.5.3.6. The second part of the theorem is not true in general, because taking invariants does not always preserve exact sequences. [18]

Corollary I.5.3.7. *If $R = k$ is a field, G and H are finite and $\mathrm{char}(k)$ does not divide $|\mathrm{Ker}(\alpha)|$, then CoInd_H^G induces a morphism of groups $\mathrm{R}_k(H) \to \mathrm{R}_k(G)$.* [19]

I.5.4 Frobenius reciprocity

Definition I.5.4.1. Let M be a $R[H]$-module. We denote by ε_M the morphism $\mathrm{Res}_H^G \mathrm{CoInd}_H^G M \to M$ sending $u \in \mathrm{Hom}_{R[H]}(R[G], M)$ to $u(1) \in M$, and by η_M the morphism $M \to \mathrm{Res}_H^G \mathrm{Ind}_H^G M$, $m \longmapsto 1 \otimes m$.

Proposition I.5.4.2. *For every $R[H]$-module M, the maps ε_M and η_M are $R[H]$-linear.*

[18] See any book on group cohomology, for example Brown's book [5].

[19] Note that CoInd_H^G is actually canonically isomorphic to Ind_H^G in that case, by corollary I.5.5.2 and proposition I.5.5.3.

Proof. It is clear that η_M is $R[H]$-linear. Let $u \in \mathrm{Hom}_{R[H]}(R[G], M)$ and $x \in R[H]$. Then

$$\varepsilon_M(x \cdot u) = (x \cdot u)(1) = u(x) = xu(1)$$

because u is $R[H]$-linear.

□

Theorem I.5.4.3 (Frobenius reciprocity)**.** *Let M be a $R[G]$-module and N be a $R[H]$-module.*

1. *The morphism of groups*

$$\Phi : \mathrm{Hom}_{R[G]}(M, \mathrm{CoInd}_H^G N) \to \mathrm{Hom}_{R[H]}(\mathrm{Res}_H^G M, N)$$

sending $u \in \mathrm{Hom}_{R[G]}(M, \mathrm{CoInd}_H^G N)$ to $\varepsilon_N \circ \mathrm{Res}_H^G(u)$ is an isomorphism. It is R-linear if R is commutative.

2. *The morphism of groups*

$$\Phi' : \mathrm{Hom}_{R[G]}(\mathrm{Ind}_H^G N, M) \to \mathrm{Hom}_{R[H]}(N, \mathrm{Res}_H^G M)$$

sending $u \in \mathrm{Hom}_{R[G]}(\mathrm{Ind}_H^G N, M)$ to $\mathrm{Res}_H^G(u) \circ \eta_N$ is an isomorphism. It is R-linear if R is commutative.

Remark I.5.4.4. This theorem is not specific to group algebras and stays true (with the obvious modifications in the definitions and the same proof) if we replace the morphism of rings $R[H] \to R[G]$ by any morphism of rings $R_1 \to R_2$.

Also, point (ii) follows from a more general statement, called the adjunction between $\otimes$ and Hom. (See problem VII.1.1.)

Proof. The statements about the R-linearity of Φ and Ψ if R is commutative are obvious.

Write $R_1 = R[H]$ and $R_2 = R[G]$.

1. Consider the map

$$\Psi : \mathrm{Hom}_{R[H]}(\mathrm{Res}_H^G M, N) = \mathrm{Hom}_{R_1}(M, N) \to$$
$$\mathrm{Hom}_{R[G]}(M, \mathrm{CoInd}_H^G N) = \mathrm{Hom}_{R_2}(M, \mathrm{Hom}_{R_1}(R_2, N))$$

sending $u : M \to N$ to the map $\Psi(u) : m \longmapsto (a \longmapsto u(am))$. If u is R_1-linear, then $\Psi(u)$ sends M to $\mathrm{Hom}_{R_1}(R_2, N)$, and we check easily that it is R_2-linear : indeed, for every $b \in R_2$ and $m \in M$, if $a \in R_2$, then

$$\Psi(u)(bm))(a) = u(abm) = (\Psi(u)(m))(ab) = (b \cdot (\Psi(u)(m)))(a).$$

To prove the statement, we only need to show that Φ and Ψ are inverses of each other.

Let u be a R_1-linear map from M to N. Then $\Psi(u) : M \to \mathrm{Hom}_{R_1}(R_2, N)$ is the R_2-linear map $m \longmapsto (a \longmapsto u(am))$, and $\Phi\Psi(u) : M \to N$ is the R_1-linear map sending $m \in M$ to $(\Psi(u)(m))(1) = u(m)$. So $\Psi\Phi(u) = u$.

Let v be a R_2-linear map from M to $\mathrm{Hom}_{R_1}(R_2, N)$. Then $\Phi(v)$ is the R_1-linear map $m \longmapsto v(m)(1)$, and, for every $m \in M$ and $a \in R_2$,

$$(\Psi\Phi(v)(m))(a) = \Phi(v)(am) = (v(am))(1) = (a \cdot v(m))(1) = v(m)(a).$$

So $\Psi\Phi(v) = v$.

2. Consider the map

$$\Psi' : \mathrm{Hom}_{R[H]}(N, \mathrm{Res}^G_H M) = \mathrm{Hom}_{R_1}(N, M) \to$$
$$\mathrm{Hom}_{R[G]}(\mathrm{Ind}^G_H N, M) = \mathrm{Hom}_{R_2}(R_2 \otimes_{R_1} N, M)$$

sending a R_1-linear map $u : N \to M$ to the R_2-linear map $\Psi'(u) : R_2 \otimes_{R_1} N \to M$, $a \otimes n \longmapsto au(n)$. (This map $\Psi'(u)$ is well-defined because $(a, n) \longmapsto au(n)$ is additive in each variable, and because we have $\varphi(a)u(n) = u(\varphi(a)n)$ if $a \in R_1$.)

To prove the statement, we only need to show that Φ' and Ψ' are inverses of each other.

Let u be a R_1-linear map from N to M. Then for every $n \in N$,

$$\Phi'\Psi'(u)(n) = \Psi'(u)(1 \otimes n) = u(n).$$

So $\Phi'\Psi'(u) = u$.

Let v be a R_2-linear map from $R_2 \otimes_{R_1} N$ to M. Then for every $a \in R_2$ and $n \in N$,

$$\Psi'\Phi'(v)(a \otimes n) = a\Phi'(v)(n) = av(1 \otimes n) = v(a \otimes n).$$

So $\Psi'\Phi'(v) = v$.

□

I.5.5 Comparing induction and coinduction

Proposition I.5.5.1. *Suppose that $\alpha : H \to G$ is injective, and let M be a $R[H]$-module. Then we have a canonical isomorphism of $R[G]$-modules between $\mathrm{Ind}^G_H M$ and the set of $f : G \to M$ such that f is supported on a finite union of right cosets of H in G and that $\forall h \in H, \forall g \in G,\ f(\alpha(h)g) = hf(g)$.*

The action of G on the second module is given in the following way : If $f : G \to M$ and if $x \in G$, then $x \cdot f : G \to M$ is defined by $(x \cdot f)(g) = f(gx)$, for every $g \in G$.

Proof. See problem VII.1.12.

□

Corollary I.5.5.2. *Suppose that $\alpha : H \to G$ is injective. Then, for every $R[H]$-module M, we have a canonical $R[G]$-linear injective map $\mathrm{Ind}_H^G M \to \mathrm{CoInd}_H^G M$.*

If moreover the image of α has finite index in G, this map is an isomorphism.

Proof. This follows immediately from remark I.5.1.5 and proposition I.5.5.1.

□

Proposition I.5.5.3. *Suppose that $\alpha : H \to G$ is surjective. Then, for every $R[H]$-module M, we have a canonical $R[G]$-linear map*

$$\mathrm{CoInd}_H^G M = M^{\mathrm{Ker}\,\alpha} \to \mathrm{Ind}_H^G M = M_{\mathrm{Ker}\,\alpha},$$

induced by the identity of M.

If moreover R is a semisimple ring, $\mathrm{Ker}(\alpha)$ is finite and $|\mathrm{Ker}(\alpha)|$ is invertible in R, then this morphism is always an isomorphism.

Proof. We have $\mathrm{CoInd}_H^G M = M^{\mathrm{Ker}\,\alpha}$ and $\mathrm{Ind}_H^G M = M_{\mathrm{Ker}\,\alpha}$ by lemmas I.5.3.4 and I.5.2.5, which gives the map.

Suppose that R is a semisimple ring, that $K := \mathrm{Ker}(\alpha)$ is finite and that $|K|$ is invertible in R. Then $R[K]$ is a semisimple ring by theorem I.3.2, so there exists a $R[K]$-submodule N of M such that $M = M^K \oplus N$. Hence $M_K = (M^K)_K \oplus N_K = M^K \oplus N_K$, which show that the map $u : M^K \to M_K$ induced by id_M is surjective.

To show that u is injective, we construct its inverse. Consider the R-linear map $v' : M \to M^K$ sending $m \in M$ to $\frac{1}{|K|}\sum_{g\in K} gm$. This makes sense because $|K|$ is invertible in R, and it clearly lands in M^K. Also, $v'(gm - m) = 0$ for every $g \in K$ and $m \in M$, so v' induces a map $v : M_K \to M^K$, and it is easy to check that v is the inverse of u.

□

I.5.6 The projection formula

In this section, we assume that the ring R is commutative.

Proposition I.5.6.1. *If M is a $R[H]$-module and N is a $R[G]$-module, then we have a canonical $R[G]$-linear isomorphism*

$$\mathrm{Ind}_H^G(M \otimes_R (\mathrm{Res}_H^G N)) \simeq (\mathrm{Ind}_H^G M) \otimes_R N.$$

Proof. We write

$$M_1 = \mathrm{Ind}_H^G(M \otimes_R (\mathrm{Res}_H^G N)) = R[G] \otimes_{R[H]} (M \otimes_R N)$$

and

$$M_2 = (\mathrm{Ind}_H^G M) \otimes_R N = (R[G] \otimes_{R[H]} M) \otimes_R N.$$

Note that there is obvious R-linear isomorphism between M_1 and M_2, but it is not $R[G]$-linear in general.

Instead, consider the R-linear maps

$$\varphi : \left\{ \begin{array}{ccc} M_1 & \to & M_2 \\ g \otimes (v \otimes w) & \longmapsto & (g \otimes v) \otimes (gw) \end{array} \right.$$

and

$$\psi : \left\{ \begin{array}{ccc} M_2 & \to & M_1 \\ (g \otimes v) \otimes w & \longmapsto & g \otimes (v \otimes (g^{-1}w)). \end{array} \right.$$

It's easy to see that these maps are well-defined and inverses of each other, so we just need to check that φ is G-linear. Let $g, h \in G$ and $v \in M$ and $w \in N$. Then

$$\varphi(h(g\otimes(v\otimes w))) = \varphi((hg)\otimes(v\otimes w)) = ((hg)\otimes v)\otimes(hgw) = h((g\otimes v)\otimes w) = h\varphi(g\otimes(v\otimes w)).$$

□

Corollary I.5.6.2. *Suppose that R is a field k. (So that $\mathrm{R}_k(G)$ and $\mathrm{R}_k(H)$ are rings.) Then, for every $x \in \mathrm{R}_k(H)$ and $y \in \mathrm{R}_k(G)$, we have*

$$\mathrm{Ind}_H^G(x \,\mathrm{Res}_H^G(y)) = (\mathrm{Ind}_H^G x)y.$$

In other words, if we make $\mathrm{R}_k(G)$ act on $\mathrm{R}_k(H)$ via the morphism of rings Res_H^G, then $\mathrm{Ind}_H^G : \mathrm{R}_k(H) \to \mathrm{R}_k(G)$ is $\mathrm{R}_k(G)$-linear.

In particular, the image of Ind_H^G is an ideal of $\mathrm{R}_k(G)$.

I.5.7 The case of finite groups

We now suppose that the groups G and H are finite. Then we know that :

1. For every $R[H]$-module M, there is a canonical $R[G]$-module map $\mathrm{CoInd}_H^G M \to \mathrm{Ind}_H^G M$. If α is injective, or if R is a semisimple ring and $|\operatorname{Ker}\alpha|$ is invertible in R, this is an isomorphism. (By corollary I.5.5.2, proposition I.5.5.3 and the transitivity of induction and of coinduction.)
2. Res_H^G preserves exact sequences. If α is injective or if R and semisimple and $|\operatorname{Ker}\alpha|$ is invertible in R, so do Ind_H^G and CoInd_H^G. (By remark I.5.1.2 and theorems I.5.2.4 and I.5.3.3.)

II Characteristic 0 theory

In this chapter, unless otherwise specified, k will be a field of characteristic 0 and all the groups are finite. So for any group G, the k-algebra $k[G]$ is semisimple (see theorem I.3.2 of chapter I).

A representation of a group G will be a representation of G on a finite-dimensional k-vector space, i.e. a finite length $k[G]$-module, and we'll usually write Hom_G instead of $\mathrm{Hom}_{k[G]}$. The *regular representation* of G is the representation corresponding to the left regualr $k[G]$-module.

II.1 Characters

II.1.1 Definition

Let G be a group.

Definition II.1.1.1. A function $f : G \to k$ is called *central* if for all $g, h \in G$, $f(gh) = f(hg)$. We write $\mathscr{C}(G, k)$ for the k-algebra of central functions from G to k.

Definition II.1.1.2. Let (V, ρ) be a representation of G. The *character of* V is the function

$$\chi_V : \left\{ \begin{array}{ccc} G & \to & k \\ g & \longmapsto & \mathrm{Tr}(\rho(g)). \end{array} \right.$$

Proposition II.1.1.3. *Let (V, ρ_V) and (W, ρ_W) be representations of G. Then :*

1. $\chi_V(1) = \dim_k V$.
2. $\chi_V \in \mathscr{C}(G, k)$.
3. $\chi_{V \oplus W} = \chi_V + \chi_W$.
4. $\chi_{V \otimes_k W} = \chi_V \chi_W$.

Remember that the action of G on $V \otimes_k W$ is defined by $g(v \otimes w) = (gv) \otimes (gw)$.

Proof. Only point (iv) is not trivial. Choose k-bases $(e_1, \ldots, e_n)$ and $(f_1, \ldots, f_m)$ of V and W. Let $g \in G$. In the chosen bases of V and W, write $\rho_V(g)$ and $\rho_W(g)$ as matrices $(x_{ij})_{1 \leq i,j \leq n}$ and

$(y_{ij})_{1\leq i,j\leq m}$. Then $(e_i \otimes f_j)_{1\leq i\leq n, 1\leq j\leq m}$ is a basis of $V \otimes_k W$, and the matrix of $\rho_V(g) \otimes \rho_W(g)$ in this basis is $(x_{i,i'}y_{j,j'})_{1\leq i,i'\leq n, 1\leq j,j'\leq m}$. Hence

$$\chi_{V\otimes_k W}(g) = \sum_{i=1}^{n}\sum_{j=1}^{m} x_{i,i}y_{j,j} = (\sum_{i=1}^{n} x_{i,i})(\sum_{j=1}^{m} y_{j,j}) = \chi_v(g)\chi_W(g).$$

□

Corollary II.1.1.4. *The map $V \longmapsto \chi_V$ induces a morphism of rings $\mathrm{R}_k(G) \to \mathscr{C}(G,k)$.*

Remark II.1.1.5. Let K/k be an extension of fields. Then, for every representation of G over k, we have $\chi_{V\otimes_k K} = \chi_V$, so we get a commutative diagramm

$$\begin{array}{ccc} \mathrm{R}_k(G) & \longrightarrow & \mathscr{C}(G,k) \\ \downarrow & & \downarrow \\ \mathrm{R}_K(G) & \longrightarrow & \mathscr{C}(G,K) \end{array}$$

where the horizontal arrows are those of the previous corollary, the left vertical arrow is given by $[V] \longmapsto [V \otimes_k K]$ and the right vertical arrow is the obvious inclusion.

Remark II.1.1.6. If V is a representation of G, we have a k-algebra map $k[G] \to \mathrm{End}_k(G)$, so we can extend χ_V to a function $\chi_V : k[G] \to k$, and properties (iii) and (iv) of proposition II.1.1.3 still hold.

Definition II.1.1.7. Let V, W be representations of G.

- The k-vector space $\mathrm{Hom}_k(V,W)$ becomes a representation of G if we make $g \in G$ act by $(g \cdot f)(v) = gf(g^{-1}v)$, for $f \in \mathrm{Hom}_k(V,W)$ and $v \in V$.
- In particular, $V^* := \mathrm{Hom}_k(V,k)$ is a representation of G (we use the trivial action of G on k), and we have $(g \cdot f)(v) = f(g^{-1}v)$ for all $g \in G$, $f \in V^*$ and $v \in V$.

Definition II.1.1.8. (See definition I.5.3.1 of chapter I.) For every representation V of G, we set

$$V^G = \{v \in V | \forall g \in G,\ gv = v\}.$$

This is the space of *invariants* of G in V.

Remark II.1.1.9. If V and W are representations of G, then

$$\mathrm{Hom}_G(V,W) = \mathrm{Hom}_k(V,W)^G.$$

Proposition II.1.1.10. *Let V, W be representations of G. Then the map*

$$\left\{ \begin{array}{ccc} V^* \otimes_k W & \to & \mathrm{Hom}_k(V,W) \\ v \otimes w & \longmapsto & (v \longmapsto f(v)w) \end{array} \right.$$

is a G-equivariant isomorphism.

Proof. Let's call this map φ. It is well-defined because the map $V \times W \to \mathrm{Hom}_k(V,W)$, $(v,w) \longmapsto (v \longmapsto f(v)w)$, is k-bilinear.

First we show that φ is G-equivariant. Let $f \in V^*$, $w \in W$, $g \in G$ and $v \in V$. Then

$$(\varphi(g(f \otimes w)))(v) = (\varphi((gf) \otimes (gw)))(v) = (gf)(v)gw = f(g^{-1}v)gw,$$

and

$$(g\varphi(f \otimes w))(v) = g(\varphi(f \otimes w)(g^{-1}v)) = gf(g^{-1}v)w.$$

Let's show that φ is bijective. Let $(e_i)_{i\in I}$ be a basis of W as a k-vector space. Then we have a k-linear isomorphism $u : \bigoplus_{i\in I} \mathrm{Hom}_k(V,k) \xrightarrow{\sim} \mathrm{Hom}_k(V,W)$ sending $(f_i)_{i\in I}$ to $v \longmapsto \sum_{i\in I} f_i(v)e_i$, and a k-linear isomorphism $v : \bigoplus_{i\in I} V^* \xrightarrow{\sim} V^* \otimes_k W$ sending $(f_i)_{i\in I}$ to $\sum_{i\in I} f_i \otimes e_i$. Now we just have to notice that $\varphi = u \circ v^{-1}$.

□

Proposition II.1.1.11. *Let V, W be representations of G. Then, for every $g \in G$:*

1. $\chi_{V^*}(g) = \chi_V(g^{-1})$;
2. $\chi_{\mathrm{Hom}_k(V,W)} = \chi_V(g^{-1})\chi_W(g)$.

Proof. Point (ii) follows from (i) and from propositions II.1.1.3 and II.1.1.10. Let's prove (i). Write ρ for the action morphism $G \to \mathrm{End}_k(V)$. Choose a basis $\mathscr{B}$ of V as a k-vector space, and let M be the matrix of $\rho(g)^{-1}$ in $\mathscr{B}$. Then g acts by the matrix ${}^tM$ on V^*, so the result follows from the fact that $\mathrm{Tr}(M) = \mathrm{Tr}({}^tM)$.

□

II.1.2 Orthogonality of characters

Let's first reformulate Schur's lemma in our situation.

Theorem II.1.2.1 (Schur's lemma)**.** *Let V, W be irreducible representations of G. Then $\mathrm{Hom}_k(V,W) = 0$ unless $V \simeq W$, and $\mathrm{End}_k(V)$ is a finite-dimensional k-division algebra. If moreover k is algebraically closed, then $\mathrm{End}_k(V) = k$.*

Also, remember (definition I.3.5 of chapter I) that $S_k(G)$ is a set of representatives of the isomorphism classes of irreducible representations of G over k.

Proof. Everything but the last sentence follows from theorem I.1.4.1 of chapter I and the fact that $\dim_k(V)$ is finite. The last statement follows from problem VII.1.3.

□

Theorem II.1.2.2. *Let V, W be representations of G. Then*

$$\frac{1}{|G|}\sum_{g\in G}\chi_V(g^{-1})\chi_W(g) = \frac{1}{|G|}\sum_{g\in G}\chi_{V^*}(g)\chi_W(g) = \dim_k(\mathrm{Hom}_G(V,W)).$$

We will use the following two lemmas, which are particular cases of the theorem.

Lemma II.1.2.3. *Suppose that k is algebraically closed, and let (V,ρ) be an irreducible representation of V. Then*

$$\sum_{g\in G}\chi_V(g) = \begin{cases} 0 & \text{if } V \not\simeq \mathbb{1} \\ |G| & \text{if } V \simeq \mathbb{1}. \end{cases}$$

Proof. Extend $\rho : G \to \mathrm{End}_k(V)$ to $\rho : k[G] \to \mathrm{End}_k(V)$, and let $c = \sum_{g\in G} g \in k[G]$. Then c is central in $k[G]$ (in fact, $hc = ch = c$ for every $h \in G$), so $\rho(c) \in \mathrm{End}_k(G)$ is a G-equivariant endomorphism of V. By Schur's lemma, there exists $\lambda \in k$ such that $\rho(c) = \lambda \mathrm{id}_V$, and so we have

$$\sum_{g\in G}\chi_V(c) = \mathrm{Tr}(\rho(c)) = \lambda \dim V.$$

Moreover, for every $h \in G$,

$$\lambda\rho(h) = \rho(c)\rho(h) = \rho(ch) = \rho(c) = \lambda\mathrm{id}_V.$$

So, if $\lambda \neq 0$, then $V \simeq \mathbb{1}$, and then of course

$$\sum_{g\in G}\chi_V(g) = \sum_{g\in G} 1 = |G|.$$

□

Lemma II.1.2.4. *Let V be a representation of G. Then*

$$\sum_{g\in G}\chi_V(g) = |G|\dim_k(V^G).$$

Proof. By remark II.1.1.5 and problem VII.2.1(1), we may assume that k is algebraically closed. Write $V = \bigoplus_{W\in S_k(G)} W^{\oplus n_W}$. Then

$$\sum_{g\in G}\chi_V(g) = \sum_{W\in S_k(G)} n_W \sum_{g\in G}\chi_W(g).$$

By lemma II.1.2.3, this is equal to $n_{\mathbb{1}}|G|$.

On the other hand, if $W \in S_k(G)$ and $W \simeq \mathbb{1}$, then $W^G = 0$ (because W^G is a subrepresentation of W). So $V^G = \mathbb{1}^{n_{\mathbb{1}}}$, and $\dim_k(V^G) = n_{\mathbb{1}}$.

□

Proof of the theorem. By propositions II.1.1.3 and II.1.1.11,

$$\frac{1}{|G|}\sum_{g\in G}\chi_V(g^{-1})\chi_W(g) = \frac{1}{|G|}\sum_{g\in G}\chi_{V^*}(g)\chi_W(g) = \frac{1}{|G|}\sum_{g\in G}\chi_{\mathrm{Hom}_k(V,W)}(g).$$

By lemma II.1.2.4 and remark II.1.1.9, this is equal to

$$\dim_k(\mathrm{Hom}_k(V,W)^G) = \dim_k(\mathrm{Hom}_G(V,W)).$$

□

Corollary II.1.2.5. *If V and W are irreducible and nonisomorphic, then*

$$\sum_{g\in G}\chi_{V^*}(g)\chi_W(g) = 0.$$

Proof. This follows from the theorem and from Schur's lemma.

□

Corollary II.1.2.6. *Suppose that k is algebraically closed, and let V be a representation of G. Then V is irreducible if and only if $\sum_{g\in G}\chi_{V^*}(g)\chi_V(g) = |G|$.*

Proof. The theorem says that $\sum_{g\in G}\chi_{V^*}(g)\chi_v(g) = |G|\dim_k(\mathrm{End}_G(V))$.

If V is irreducible, then $\mathrm{End}_k(V) = k$ by Schur's lemma.

Conversely, suppose that $\sum_{g\in G}\chi_{V^*}(g)\chi_v(g) = |G|$. Write $v = \bigoplus_{i\in I} V_i^{\oplus n_i}$, where the V_i are irreducible and pairwise nonisomorphic. Then by corollary II.1.2.5 and what we just saw above,

$$\sum_{g\in G}\chi_{V^*}(g)\chi_V(g) = \sum_{i,j\in I} n_i n_j \sum_{g\in G}\chi_{V_i^*}(g)\chi_{V_j}(g) = |G|\sum_{i\in I} n_i^2.$$

So only one of the n_i can be nonzero, and moreover it has to be equal to 1. Hence V is irreducible.

□

Remark II.1.2.7. If we don't assume that k is algebraically closed, then the same proof shows that any representation V of G satisfying $\sum_{g\in G}\chi_{V^*}(g)\chi_v(g) = |G|$ has to be irreducible.

Corollary II.1.2.8. *The family $(\chi_V)_{V\in S_k(G)}$ is linearly independent in $\mathscr{C}(G,k)$.*

Proof. Suppose that $\sum_{W\in S_k(G)}\alpha_W\chi_W = 0$, with $\alpha_W \in k$. Then, for every $V \in S_k(G)$, we have

$$0 = \sum_{g\in G}\Big(\sum_{W\in S_k(G)}\alpha_W\chi_W(g)\Big)\chi_{V^*}(g) = \sum_{W\in S_k(G)}\alpha_W\sum_{g\in G}\chi_W(g)\chi_{V^*}(g) = \alpha_V|G|,$$

hence $\alpha_V = 0$.

□

Corollary II.1.2.9. *For any representations V, V' of G, we have $V \simeq V'$ if and only if $\chi_V = \chi_{V'}$.*

In particular, the map $R_k(G) \to \mathscr{C}(G,k)$ is injective.

Proof. Write $V = \bigoplus_{W \in S_k(G)} W^{\oplus n_W}$ and $V' = \bigoplus_{W \in S_k(G)} W^{\oplus n'_W}$. Then $\chi_V = \sum_{W \in S_k(G)} n_W \chi_W$ and $\chi_{V'} = \sum_{W \in S_k(G)} n'_W \chi_W$.

By corollary II.1.2.8, $\chi_V = \chi_{V'}$ if and only if $n_W = n'_W$ for every $W \in S_k(G)$, which is also equivalent to $V \simeq V'$.

□

II.1.3 Characters and representation ring

Theorem II.1.3.1. *Suppose that the field k is algberaically closed. Then the family $(\chi_W)_{W \in S_k(G)}$ is a basis of $\mathscr{C}(G,k)$.*

Lemma II.1.3.2. *If $f \in \mathscr{C}(G,k)$ is such that*

$$\sum_{g \in G} f(g)\chi_{W^*}(g) = 0$$

for every $W \in S_k(G)$, then $f = 0$.

Proof. If $\rho : G \to \mathrm{End}_k(V)$ is a representation of G, we set

$$\rho(f) = \sum_{g \in G} f(g)\rho(g) \in \mathrm{End}_k(V).$$

Then, for every $g \in G$,

$$\rho(g)\rho(f) = \sum_{h \in G} f(g)\rho(gh) = \sum_{h \in G} f(g)\rho(ghg^{-1}g) = \sum_{h \in G} f(ghg^{-1})\rho(ghg^{-1})\rho(g) = \rho(f)\rho(g),$$

because f is a central function. So, if V is irreducible, then $\rho(f) \in \mathrm{End}_G(V) = k$ (by Schur's lemma), hence we can write $\rho(f) = \lambda \mathrm{id}_V$ with $\lambda \in k$, and we have

$$\lambda \dim(V) = \mathrm{Tr}(\rho(f)) = \sum_{g \in G} f(g)\chi_V(g) = 0,$$

which gives λ and finally $\rho(f) = 0$.

By semisimplicity of $k[G]$, we get that $\rho(f) = 0$ for any representation ρ of G. Applying this to the regular representation ρ_{reg} gives

$$0 = \rho_{reg}(f)1 = \sum_{g \in G} f(g)g$$

in $k[G]$, i.e., $f(g) = 0$ for every $g \in G$.

□

Proof of the theorem. We already know that this family is linearly independent, so we just need to show that it generates $\mathscr{C}(G, k)$.

Let $f \in \mathscr{C}(G, k)$. For every $W \in S_k(G)$, let

$$\alpha_W = \frac{1}{|G|} \sum_{g \in G} f(g) \chi_{W^*(g)}.$$

If we knew $(\chi_W)_{W \in S_k(G)}$ was a basis $\mathscr{C}(G, k)$, this α_W would be the coefficient of f corresponding to χ_W, by corollaries II.1.2.5 and II.1.2.6. So we set

$$f' = f - \sum_{W \in S_k(G)} \alpha_W \chi_W,$$

and we try to prove that $f' = 0$. For every $W \in S_k(G)$, using corollary II.1.2.5 gives

$$\sum_{g \in G} f'(g) \chi_{W^*}(g) = \sum_{g \in G} f(g) \chi_{W^*}(g) - \alpha_W \sum_{g \in G} \chi_W(g) \chi_{W^*}(g) = 0,$$

so the result follows from the lemma.

□

Corollary II.1.3.3. *Suppose that k is algberaically closed. Then we have $|S_k(G)| = \dim_k \mathscr{C}(G, k)$, and this is also equal to the number of conjugacy classes in G.* □

Corollary II.1.3.4. *Suppose that k is algebraically clsoed. Then the map $\mathrm{R}_k(G) \to \mathscr{C}(G, k)$, $[V] \longmapsto \chi_V$, of corollary II.1.1.4 induces an isomorphism of k-algebras $\mathrm{R}_k(G) \otimes_{\mathbb{Z}} k \xrightarrow{\sim} \mathscr{C}(G, k)$.* □

Remark II.1.3.5. Many of the results of the previous three sections are true for any field k (whose characteristic does not divide $|G|$). The most important results that require k to be algebraically closed are II.1.2.6 and theorem II.1.3.1 (more precisely, the fact that the characters of irreducible representations generate $\mathscr{C}(G, k)$) and its corollaries II.1.3.3 and II.1.3.4.

II.1.4 The case $k = \mathbb{C}$

Proposition II.1.4.1. *Let k be any field, and let (V, ρ) be a representation of a group G. Then for every $g \in G$, all the eigenvalues of $\rho(g)$ are $|G|$th roots of 1 in $\overline{k}$.*

Proof. Let $g \in G$. We have $\rho(g^{|G|}) = \rho(g)^{|G|} = \mathrm{id}_V$, hence the characteristic polynomial of $\rho(g)$ divides $T^{|G|} - 1$.

□

Corollary II.1.4.2. *Let (V, ρ) be a representation of V over $\mathbb{C}$. Then, for every $g \in G$, $\chi_{V^*}(g) = \overline{\chi_V(g)}$.*

Proof. Let $g \in G$, and let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $\rho(g)$ (with multiplicities). Then

$$\chi_{V^*}(g) = \chi_V(g^{-1}) = \lambda_1^{-1} + \cdots + \lambda_n^{-1} = \overline{\lambda}_1 + \cdots + \overline{\lambda}_n.$$

□

Corollary II.1.4.3. *Define a Hermitian inner product on the finite-dimensional $\mathbb{C}$-vector space $\mathscr{C}(G, \mathbb{C})$ by*

$$f_1 \cdot f_2 = \frac{1}{|G|} \sum_{g \in G} f_1(g)\overline{f_2(g)}.$$

Then $(\chi_W)_{W \in S_{\mathbb{C}}(G)}$ is an orthonomal basis of $\mathscr{C}(G, \mathbb{C})$.

II.2 Representations of a product of groups

Let G_1 and G_2 be two groups. If V_1 (resp. V_2) is a representation of G_1 (resp. G_2), then $V_1 \otimes_k V_2$ becomes a representation of $G_1 \times G_2$ with the action $(g_1, g_2)(v_1 \otimes v_2) = (g_1 v_1) \otimes (g_2 v_2)$.

It is easy to show (see proposition II.1.1.3) that

$$\chi_{V_1 \otimes_k V_2}(g_1, g_2) = \chi_{V_1}(g_1)\chi_{V_2}(g_2).$$

Theorem II.2.1. *Suppose that k is algebraically closed, and let V_1 (resp. V_2) be a representation of G_1 (resp. G_2).*

1. *The representation $V_1 \otimes_k V_2$ of $G_1 \times G_2$ is irreducible if and only if both V_1 and V_2 are irreducible.*
2. *Every irreducible representation of $G_1 \times G_2$ is of the form $V_1 \otimes_k V_2$.*

This theorem is proved in problem VII.2.4. This same problem also contains a counterexample to point (i) if k is not algebraically closed.

II.3 Characters and induced representation

Remember that in the present setting, the induction and the coinduction coincide (see section I.5.7 of chapter I).

We fix a finite group G, and we assume again that k is any field whose characteristic does not divide $|G|$.

Notation. If $f_1, f_2 \in \mathscr{C}(G,k)$, write $\langle f_1, f_2\rangle_G$ (or just $\langle f_1, f_2\rangle$ if G is clear from the context) for

$$\frac{1}{|G|}\sum_{g\in G} f_1(g)f_2(g^{-1}) = \sum_{g\in G} f_1(g^{-1})f_2(g).$$

Note that $\langle f_1, f_2\rangle = \langle f_2, f_1\rangle$, and that this differs slightly from the inner product of section II.1.4 if $k = \mathbb{C}$.

With this notation, theorem II.1.2.2 becomes : For any representations V and W of G, we have

$$\langle \chi_V, \chi_V\rangle = \dim_k(\mathrm{Hom}_G(V,W)).$$

II.3.1 Character of an induced representation

Let H be a subgroup of G.

Definition II.3.1.1. If $f \in \mathscr{C}(H,k)$, define $\mathrm{Ind}_H^G f : G \to k$ by

$$\mathrm{Ind}_H^G f(g) = \frac{1}{|H|}\sum_{s\in G | s^{-1}gs\in H} f(s^{-1}gs).$$

Theorem II.3.1.2. *1. For every $f \in \mathscr{C}(H,k)$, $\mathrm{Ind}_H^G f \in \mathscr{C}(G,k)$.*

2. If V is a representation of H, then $\mathrm{Ind}_H^G \chi_V = \chi_{\mathrm{Ind}_H^G V}$. [1]

Proof. It's easy enough to prove (i) directly, and we can also deduce it from (ii). (Indeed, we may assume that k is algebraically closed, and then any $f \in \mathscr{C}(V,k)$ is a linear combination of characters of representations.)

Let's prove (ii). Let (V,ρ_V) be a representation of H, and write $(W,\rho_W) = \mathrm{Ind}_H^G(V,\rho_V)$. Let $g_1,\dots,g_r$ be a system of representatives of G/H. Then $k[G] = \bigoplus_{i=1}^r g_i k[H]$, so

$$W = k[G]\otimes_{k[H]} V = \bigoplus_{i=1}^r W_i,$$

[1] See problem VII.2.16 for a generalization to a morphism $H \to G$ that is not necessarily injective.

with

$$W_i = g_i k[H] \otimes_{k[H]} V.$$

Let $g \in G$. Then we have $\sigma \in \mathfrak{S}_r$ and $h_1, \ldots, h_r \in H$ such that $gg_i = g_{\sigma(i)}h_i$, for every $i \in \{1, \ldots, r\}$. Then $\rho_W(g)(W_i) = W_{\sigma(i)}$ for every i. Choose a basis $(e_1, \ldots, e_n)$ of V. Then for every $i \in \{1, \ldots, r\}$, $(g_i \otimes e_1, \ldots, g_i \otimes e_n)$ is a basis of W_i, and we have

$$\rho_W(g)(g_i \otimes e_s) = g_{\sigma(i)} \otimes \rho_V(h_i)e_s.$$

So

$$\mathrm{Tr}(\rho_W(g) \sum_{i|\sigma(i)=i} \mathrm{Tr}(\rho_V(h_i)) = \sum_{i|g_i^{-1}gg_i \in H} f(g_i^{-1}gg_i).$$

But we have $G = \coprod_{i=1}^r g_i H$ and, if $s \in g_i H$, then :

(a) $s^{-1}gs \in H$ if and only if $g_i^{-1}gg_i \in H$;

(b) $f(s^{-1}gs) = f(g_i^{-1}gg_i)$.

This finishes the proof.

□

Remark II.3.1.3. If $f \in \mathscr{C}(G, k)$, we write $\mathrm{Res}_H^G f$ for $f_{|H}$. Then $\mathrm{Res}_H^G \chi_V = \chi_{\mathrm{Res}_H^G V}$ for every representation V of G.

II.3.2 Frobenius reciprocity with characters

We still assume that H is a subgroup of G.

Theorem II.3.2.1. *If $f_1 \in \mathscr{C}(H, k)$ and $f_2 \in \mathscr{C}(G, k)$, then*

$$\langle f_1, \mathrm{Res}_H^G f_2 \rangle_H = \langle \mathrm{Ind}_H^G f_1, f_2 \rangle_G.$$

Compare with theorem I.5.4.3 of chapter I.

Proof. We can deduce this theorem from theorem I.5.4.3 of chapter I : We may assume that k is algebraically closed, and then $\mathscr{C}(G, k)$ and $\mathscr{C}(H, k)$ are generated by characters of representations, so we may assume that $f_1\chi_V$ and $f_2 = \chi_W$, with V (resp. W) a representation of H (resp. G). Then, by theorem II.1.2.2, the left hand side is equal to $\dim_k(\mathrm{Hom}_H(V, \mathrm{Res}_H^G W))$ and the right hand side to $\dim_k(\mathrm{Hom}_G(\mathrm{Ind}_H^G V, W))$.

But it is also very easy to prove the theorem directly. Indeed, we have

$$\langle f_1, \mathrm{Res}_H^G f_2 \rangle_H = \frac{1}{|H|} \sum_{h \in H} f_1(h) f_2(h).$$

On the other hand,

$$\begin{aligned}\langle \mathrm{Ind}_H^G f_1, f_2\rangle_G &= \frac{1}{|G|}\sum_{g\in G}\frac{1}{|H|}\sum_{s\in G|s^{-1}gs\in H} f_1(s^{-1}gs)f_2(g)\\ &= \frac{1}{|G|}\sum_{g\in G}\frac{1}{|H|}\sum_{s\in G|s^{-1}gs\in H} f_1(s^{-1}gs)f_2(s^{-1}gs)\\ &= \frac{1}{|G|}\frac{1}{|H|}\sum_{h\in H}\sum_{s\in G} f_1(h)f_2(h) = \frac{1}{|H|}\sum_{h\in H} f_1(h)f_2(h).\end{aligned}$$

□

II.3.3 Mackey's formula

Let R be a ring, let G be a finite group, and let H, K be subgroups of G. In this section, we will consider representations of these groups over R.

Let (V,ρ) be a representation of H. The question we want to answer is : What does $\mathrm{Res}_K^G \mathrm{Ind}_H^G V$ look like ?

Let $g_1,\dots,g_r$ be a system of representatives of the double classes in $K\backslash G/H$. In other words, we have

$$G = \coprod_{i=1}^r Kg_iH.$$

For every $i\in\{1,\dots,r\}$, let $H_i = g_iHg_i^{-1}\cap K$, and let (V_i,ρ_i) be the representation of H_i on V given by $\rho_i(h) = \rho(g_i^{-1}hg_i)$.

Theorem II.3.3.1 (Mackey's formula)**.** *We have an isomorphism of $R[K]$-modules*

$$\mathrm{Res}_K^G \mathrm{Ind}_H^G V \simeq \bigoplus_{i=1}^r \mathrm{Ind}_{H_i}^K V_i.$$

Proof. For every $j\in\{1,\dots,r\}$, let $(x_i)_{i\in I_j}$ be a system of representatives of K/H_j. Then $\{x_ig_j, 1\le j\le r, i\in I_j\}$ is a system of representatives of G/H. Indeed,

$$G = \coprod_{j=1}^r Kg_jH = \coprod_{j=1}^r\coprod_{i\in I_j} x_iH_jg_jH = \coprod_{j=1}^r\coprod_{i\in I_j} x_ig_j(g_j^{-1}H_jg_j)H = \coprod_{j=1}^r\coprod_{i\in I_j} x_ig_jH$$

(because $g_j^{-1}H_jg_j\subset H$ for every j).

So $W := \mathrm{Ind}_H^G V = \bigoplus_{j=1}^r W_j$, where $W_j = \bigoplus_{i\in I_j} x_ig_jk[H]\otimes_{k[H]} V$. Note that $W_j\subset W$ is stable by K.

Fix $j \in \{1, \dots, r\}$. Let's show that the $R[K]$-modules W_j and $\mathrm{Ind}_{H_j}^K V_j$ are isomorphic. (This will finish the proof.) Consider the R-linear map $\varphi : \mathrm{Ind}_{H_j}^K V_j \to W_j$ sending $\sum_{i \in I_j} x_i \otimes v_i$ to $\sum_{i \in I_j} (x_i g_j) \otimes v_i$. Consider the R-linear map $\psi : W_j \to \mathrm{Ind}_{H_j}^K V_j$ sending $\sum_{i \in K I_j} (x_i g_j) \otimes v_i$ to $\sum_{i \in I_j} x_i \otimes v_i$. It is clear that φ and ψ are inverses of each other. So we just need to show that φ is K-linear. Let $y \in K$. Then we have $\sigma \in \mathfrak{S}_{I_j}$ and $h_i \in H_j$, $i \in I_j$, such that $y x_i = x_{\sigma(i)} h_i$ for every $i \in I_j$. Then

$$\begin{aligned}
\varphi(y(\sum_{i \in I_j} x_i \otimes v_i)) &= \varphi(\sum_{i \in I_j} (y x_i) \otimes v_i) = \varphi(\sum_{i \in I_j} (x_{\sigma(i)} h_i) \otimes v_i) = \varphi(\sum_{i \in I_j} x_{\sigma(i)} \otimes (\rho_j(h_i) v_i)) \\
&= \varphi(\sum_{i \in I_j} x_{\sigma(i)} \otimes (\rho(g_j^{-1} h_i g_j) v_i)) = \sum_{i \in I_j} (x_{\sigma(i)} g_j) \otimes (\rho(g_j^{-1} h_i g_j) v_i) \\
&= \sum_{i \in I_j} (x_{\sigma(i)} h_i g_j) \otimes v_i = \sum_{i \in I_j} (y x_i g_j) \otimes v_i = y \varphi(\sum_{i \in I_j} x_i \otimes v_i)
\end{aligned}$$

(we use the fact that $g_j^{-1} h_i g_j \in H$ to move it from the right to the left of a tensor product).

□

II.3.4 Mackey's irreducibility criterion

Fix a finite group G. We come back to our field of coefficients k, and we suppose that $\mathrm{char}(k)$ does not divide G and that k is algebraically closed.

Notation. If V and W are representations of G, we write

$$\langle V, W \rangle_G = \langle \chi_V, \chi_W \rangle_G = \dim_k \mathrm{Hom}_G(V, W).$$

Let H be a subgroup of G and (V, ρ) be a repersentation of H. Even if V is irreducible, it is not always true that $\mathrm{Ind}_H^G V$ is irreducible. (For example, $\mathrm{Ind}_{\{1\}}^G \mathbb{1} = k[G]$ is only irreducible if $G = \{1\}$.) So we want a criterion to decide when $\mathrm{Ind}_H^G V$ is irreducible.

For every $g \in G$, write $H_g = g H g^{-1} \cap H$, and define two representations ρ^g and $\mathrm{Res}_g(\rho)$ of H_g on V by

- $\rho_g(h) = \rho(g^{-1} h g)$;
- $\mathrm{Res}_g(\rho)(h) = \rho(h)$.

Theorem II.3.4.1. *The following are equivalent :*

1. $W := \mathrm{Ind}_H^G V$ *is irreducible.*
2. V *is irreducible, and for every* $g \in G - H$, $\mathrm{Hom}_{H_g}(\rho^g, \mathrm{Res}_g(\rho)) = 0$ *(i.e.* ρ^g *and* $\mathrm{Res}_g(\rho)$ *have no common irreducible subrepresentation).*

Proof. We use the fact that W is irreducible if and only if $\langle W, W\rangle_G = 1$. (This is corollary II.1.2.6.) Let's calculate $\langle W, W\rangle_G$. The Frobenius reciprocity formula (theorem II.3.2.1) gives

$$\langle W, W\rangle_G = \langle V, \mathrm{Res}_H^G W\rangle_H.$$

Mackey's formula (theorem II.3.3.1) gives

$$\mathrm{Res}_H^G W \simeq \bigoplus_{g\in H\backslash G/H} \mathrm{Ind}_{H_g}^H(\rho^g),$$

hence

$$\langle W, W\rangle_G = \sum_{g\in H\backslash G/H} \langle V, \mathrm{Ind}_{H_g}^H(\rho^g)\rangle_H = \sum_{g\in H\backslash G/H} \langle \mathrm{Ind}_{H_g}^H(\rho^g), V\rangle_H.$$

Using the Frobenius reciprocity formula again show that this is equal to

$$\sum_{g\in H\backslash G/H} \langle \rho^g, \mathrm{Res}_g(\rho)\rangle_H.$$

Note that all the terms in this sum are ≥ 0. Also, if $g \in H$, then $H_g = H$ and $\rho_g \simeq \mathrm{Res}_g(\rho) = \rho$, so $\langle \rho^g, \mathrm{Res}_g(V)\rangle_H = \langle V, V\rangle_H$.

Finally, we see that W is irreducible if and only if $\langle W, W\rangle_G = 1$, if and only if $\langle V, V\rangle_H = 1$ and $\langle \rho^g, \mathrm{Res}_g(\rho)\rangle_H = 0$ for every $g \in G - H$, if and only V is irreducible and $\mathrm{Hom}_{H_g}(\rho^g, \mathrm{Res}_g(\rho)) = 0$ for every $g \in G - H$.

□

II.4 Artin's theorem

Let k and G be as before. We write $\mathrm{R}(G) = \mathrm{R}_k(G)$.

We still use the notation $\langle V, W\rangle_G$ of section II.3.4. As this number is an integer and only depends on χ_V and χ_W, it induces a symmetric $\mathbb{Z}$-bilinear map

$$\langle ., .\rangle_G : \mathrm{R}(G) \times \mathrm{R}(G) \to \mathbb{Z}.$$

Proposition II.4.1. *Write* $\mathrm{R}(G)_{\mathbb{Q}} = \mathrm{R}(G) \otimes_{\mathbb{Z}} \mathbb{Q}$.

1. *The $\mathbb{Z}$-bilinear map $\langle ., .\rangle_G : \mathrm{R}(G) \times \mathrm{R}(G) \to \mathbb{Z}$ induces a $\mathbb{Q}$-linear isomorphism*

$$\mathrm{R}(G)_{\mathbb{Q}} \xrightarrow{\sim} R(G)^*_{\mathbb{Q}} := \mathrm{Hom}_{\mathbb{Q}}(\mathrm{R}(G)_{\mathbb{Q}}, \mathbb{Q})$$

   sending x to $y \longmapsto \langle x, y\rangle_G$.

2. *Let H be a subgroup of G. If we use the isomorphism of (i) to identify $\mathrm{R}(G)_{\mathbb{Q}}$ and $\mathrm{R}(H)_{\mathbb{Q}}$ with their duals, then the transpose of Ind_H^G is Res_H^G.*

Proof. 1. Denote by $(e_V)_{V\in S_k(G)}$ the basis $([V])_{V\in S_k(G)}$ of $\mathrm{R}(G)_{\mathbb{Q}}$, and by (e_V^*) the dual basis. By theorem II.1.2.2 and corollary II.1.2.5, the map of (i) sends each e_V to a nonzero multiple of e_V^*.

Note that, if k is algebraically closed, then this map sends each e_V to e_V^* itself by corollary II.1.2.6, and it actually induces an isomorphism $\mathrm{R}(G) \xrightarrow{\sim} \mathrm{Hom}_{\mathbb{Z}}(\mathrm{R}(G), \mathbb{Z})$.

2. This is just a reformulation of the Frobenius reciprocity formula (theorem II.3.2.1.)

□

Theorem II.4.2 (Artin)**.** *In this theorem, we don't assume that k is algebraically closed (but we do suppose that $\mathrm{char}(k) = 0$).*

Let X be a set of subgroups of G. Consider the map

$$\mathrm{Ind}_X = \bigoplus_{H\in X} \mathrm{Ind}_H^G : \bigoplus_{H\in X} \mathrm{R}(H) \to \mathrm{R}(G).$$

Then the following are equivalent :

1. $G = \bigcup_{H\in X} \bigcup_{g\in G} gHg^{-1}$;
2. $\mathrm{Ind}_X \otimes_{\mathbb{Z}} \mathbb{Q}$ *is surjective, that is, for every $x \in \mathrm{R}(G)$, there exists an integer $d \geq 1$ and elements $x_H \in \mathrm{R}(H)$, $H \in X$, such that*

$$dx = \sum_{H\in X} \mathrm{Ind}_H^G x_H.$$

Proof.

(ii)⇒(i) We use the hypothesis that $\mathrm{char}(k) = 0$ to identify $\mathrm{R}(G)_{\mathbb{Q}}$ and $\mathrm{R}(H)_{\mathbb{Q}}$ to subrings of $\mathscr{C}(G, k)$ and $\mathscr{C}(H, k)$ (via characters, see corollary II.1.1.4).

Let $S = \bigcup_{H\in X} \bigcup_{g\in G} gHg^{-1}$. Then if $H \in X$ and $x_H \in \mathrm{R}(H)_{\mathbb{Q}}$, $\mathrm{Ind}_H^G x_H$ is zero on $G - S$. By (ii), this implies that $x = 0$ on $G - S$ for every $x \in \mathrm{R}(G)_{\mathbb{Q}}$. By theorem II.1.3.1, this implies that, for every $f \in \mathscr{C}(G, k)$, $f_{G-S} = 0$. Hence $G - S = \varnothing$.

(i)⇒(ii) To show that $\mathrm{Ind}_X \otimes_{\mathbb{Z}} \mathbb{Q}$ is surjective, we just need to show that it transpose (a $\mathbb{Q}$-linear map between the dual spaces) is injective. By (ii) of the proposition above, the transpose of $\mathrm{Ind}_X \otimes_{\mathbb{Z}} \mathbb{Q}$ is

$$\bigoplus_{H\in X} \mathrm{Res}_H^G : \mathrm{R}(G)_{\mathbb{Q}} \to \bigoplus_{H\in X} \mathrm{R}(H)_{\mathbb{Q}}.$$

To show that this map is injective, it suffices to show that the sum of the restriction maps

$$\mathscr{C}(G, k) \to \bigoplus_{H\in X} \mathscr{C}(H, k)$$

is injective. (As in the first part of the proof, for every subgroup H of G, $\mathrm{R}(H)_{\mathbb{Q}}$ injects naturally in $\mathscr{C}(H,k)$ because $\mathrm{char}(k)=0$.) But this follows directly from condition (i).

□

Corollary II.4.3. *Take for X the set of cyclic subgroups of G. Then $\mathrm{Ind}_X \otimes_{\mathbb{Z}} \mathbb{Q}$ is surjective.*

If k is algebraically closed, we can reformulate this result as follows : For every representation V of G, there exist cyclic subgroups $C_1,\dots,C_r$ of G, 1-dimensional representations V_i of C_i and rational numbers $\alpha_1,\dots,\alpha_r$ such that

$$[V]=\sum_{i=1}^{r}\alpha_i \,\mathrm{Ind}_{C_i}^{G}[V_i].$$

Proof. The family X satisfies the condition of Artin's theorem, because every $g \in G$ is an element of the cyclic subgroup that it generates.

The reformulation when k is algebraically closed follows from the fact that every irreducible representation of a commutative group is 1-dimensional in that case (by proposition I.3.9 of chapter I).

□

II.5 Brauer's theorem

In this section, we assume that G is a finite group and that k is an algebraically closed field of characteristic 0. [2] We write $\mathrm{R}(G)=\mathrm{R}_k(G)$, and we use characters to identify $\mathrm{R}(G)$ to a subring of $\mathscr{C}(G,k)$ (by corollary II.1.1.4).

Definition II.5.1. Let p be a prime number. A finite group H is called *p-elementary* if $H=C\times P$, with C a cyclic group of order prime to p and P a p-group.

For every prime number p, let $X(p)$ be the set of p-elementary subgroups of G. Let X be the union of all the $X(p)$ for p prime.

Theorem II.5.2 (Brauer's theorem)**.** *Let*

$$\mathrm{Ind}_X=\bigoplus_{H\in X}\mathrm{Ind}_H^G : \bigoplus_{H\in X}\mathrm{R}(H)\to\mathrm{R}(G).$$

Then Ind_X is surjective.

[2] There is a generalization of Brauer's theorem for characteristic 0 fields that are not algebraically closed; see section 12.6 of Serre's book [29].

Theorem II.5.3. *Let p be a prime number, write $|G| = p^r m$ with p prime to m, and let V_p be the image of*

$$\mathrm{Ind}_p := \bigoplus_{H \in X(p)} \mathrm{Ind}_H^G : \bigoplus_{H \in X(p)} \mathrm{R}(H) \to \mathrm{R}(G).$$

Then $m \in V_p$. In particular, $R(G)/V_p$ is a finite group of order prime to p.

Proof that theorem II.5.3 implies theorem II.5.2. We have $\mathrm{Im}(\mathrm{Ind}_X) = \sum_{p \text{ prime}} V_p$, so $\mathrm{R}(G)/\mathrm{Im}(\mathrm{Ind}_X)$ is a finite group of order prime to every prime number, i.e., the trivial group.

□

Proof of theorem II.5.3.

(1) Let $n = |G|$ and $\mathscr{O} = \mathbb{Z}[1, \zeta_n, \ldots, \zeta_n^{n-1}] \subset k$, where ζ_n is a primitive nth root of 1 in k. Then :

- For every $x \in \mathrm{R}(G)$, x (seen as a function on G) takes values in $\mathscr{O}$. Indeed, for every representation (V, ρ) of G and every $g \in G$, $\rho(g)$ has eigenvalues in $\mathscr{O}$ by proposition II.1.4.1, and so $\chi_V(g) \in \mathscr{O}$.
- We have $\mathscr{O} \cap \mathbb{Q} = \mathbb{Z}$. Indeed, let $\alpha \in \mathscr{O} \cap \mathbb{Q}$, and let $f \in \mathbb{Q}[T]$ be its minimal polynomial over $\mathbb{Q}$. Then $f \in \mathbb{Z}[T]$ as α is integral over $\mathbb{Z}$ (as an element of $\mathscr{O}$), and $\deg(f) = 1$ as $\alpha \in \mathbb{Q}$.
- By the previous point, the map $\mathscr{O}/\mathbb{Z} \to (\mathscr{O} \otimes_{\mathbb{Z}} \mathbb{Q})/\mathbb{Q}$ is injective, so $\mathscr{O}/\mathbb{Z}$ is torsion-free, so it is free as a $\mathbb{Z}$-module, so $\mathscr{O}$ itself has a $\mathbb{Z}$-basis of the form $(1, \alpha_1, \ldots, \alpha_c)$.
- The image of $\mathscr{O} \otimes \mathrm{Ind}_p : \bigoplus_{H \in X(p)} \mathscr{O} \otimes_{\mathbb{Z}} \mathrm{R}(H) \to \mathscr{O} \otimes_{\mathbb{Z}} \mathrm{R}(G)$ is $\mathscr{O} \otimes_{\mathbb{Z}} V_p$, and we have $(\mathscr{O} \otimes_{\mathbb{Z}} V_p) \cap \mathrm{R}(G) = V_p$. Indeed, we have

$$\mathscr{O} \otimes_{\mathbb{Z}} V_p = V_p \oplus \bigoplus_{i=1}^{c} \alpha_i V_p \subset \mathscr{O} \otimes_{\mathbb{Z}} \mathrm{R}(G) = \mathrm{R}(G) \oplus \bigoplus_{i=1}^{c} \alpha_i \mathrm{R}(G),$$

  and, if $y = x + \sum_{i=1} \alpha_i x_i \in \mathscr{O} \otimes_{\mathbb{Z}} V_p$ (with $x, x_i \in V_p$), then $x \in \mathrm{R}(G)$ if and only if all the x_i are 0.

By the last point above, it suffice to prove that $m \in \mathscr{O} \otimes_{\mathbb{Z}} V_p$.

(2) <u>The character θ_C</u>

If C is a cyclic group of G of order c, then let

$$\theta_C : \left\{ \begin{array}{ccc} C & \to & \mathbb{Z} \\ x & \longmapsto & \left\{ \begin{array}{ll} c & \text{if } C = \langle x \rangle \\ 0 & \text{otherwise.} \end{array} \right. \end{array} \right.$$

<u>Claim</u> : $|G| = \sum_{C \subset G \text{ cyclic}} \mathrm{Ind}_C^G(\theta_C)$.

Proof. Let $\theta'_C = \operatorname{Ind}_C^G \theta_C$. Let $x \in G$. Then

$$\theta'_C(x) = \frac{1}{c} \sum_{s \in G | sxs^{-1} \in C} \theta_C(sxs^{-1}) = \sum_{s \in G | \langle sxs^{-1} \rangle = C} 1.$$

For for every $s \in G$, sxs^{-1} generates exactly one cyclic subgroup of G. Hence

$$\sum_{C \subset G \text{ cyclic}} \theta'_C(x) = \sum_{s \in G} 1 = |G|.$$

□

Claim : For every cyclic subgroup C of G, $\theta_C \in \mathrm{R}(C)$.

Proof. By induction on $c := |C|$. The result is obvious if $c = 1$. If $c > 1$, then the previous claim gives

$$c = \sum_{B \subset C \text{ cyclic}} \operatorname{Ind}_B^C \theta_B = \theta_C + \sum_{B \subsetneq C \text{ cyclic}} \operatorname{Ind}_B^C \theta_B.$$

We have $c \in \mathrm{R}(C)$, and all the $\theta_B \in \mathrm{R}(B)$ for every $B \subsetneq C$ by the induction hypothesis, so this gives $\theta_C \in \mathrm{R}(C)$.

□

(3) Claim : Let $f \in \mathscr{C}(G, \mathbb{Z})$ such that $f(G) \subset n\mathbb{Z}$. (Remember that $n = |G|$.) Then we can write

$$f = \sum_{C \subset G \text{ cyclic}} \alpha_c \operatorname{Ind}_C^G x_C,$$

with $\alpha_C \in \mathscr{O}$ and $x_C \in \mathrm{R}(C)$.

In particular, $f \in \mathscr{O} \otimes_{\mathbb{Z}} \mathrm{R}(G)$.

Proof. Write $f = nf'$, with $f' \in \mathscr{C}(G, \mathbb{Z})$. We have $n = \sum_{C \subset G \text{ cyclic}} \operatorname{Ind}_C^G \theta_C$, hence

$$f = \sum_{C \subset G \text{ cyclic}} \operatorname{Ind}_C^G(\theta_C) f' = \sum_{C \subset G \text{ cyclic}} \operatorname{Ind}_C^G(\theta_C \operatorname{Res}_C^G f')$$

by corollary I.5.6.2 of chapter I. Write $f_C = \theta_C \operatorname{Res}_C^G f'$.

Let's show that $f_C \in \mathscr{O} \otimes_{\mathbb{Z}} \mathrm{R}(C)$ for every $C \subset G$ cyclic. (This will finish the proof.) Note that $f_C \in \mathscr{C}(C, \mathbb{Z})$, and $f_C(C) \subset |C|\mathbb{Z}$. So for every $\chi \in \mathrm{R}(C)$, then

$$\langle f_C, \chi \rangle_C = \frac{1}{|C|} \sum_{x \in C} f_C(x) \chi(x) \in \mathscr{O},$$

and hence

$$f_C = \sum_{W \in S_k(C)} \langle f_C, \chi_W \rangle_C \chi_W \in \mathscr{O} \otimes_{\mathbb{Z}} \mathrm{R}(C).$$

□

(4) An element $x \in G$ is called *p-unipotent* (resp. *p-regular* if its order is a power of p (resp. prime to p).

Claim : For every $x \in G$, there exists a unique pair (x_r, x_u) of elements of G satisfying the following conditions :

(a) x_r is p-regular and x_u is p-unipotent;

(b) $x = x_r x_u = x_u x_r$.

Moreover, x_r and x_u are powers of x.

Proof. First, if x_r and x_u satisfy (a) and (b), then they have to be powers of x. Indeed, let a (resp. b) be the order of x_r (resp. x_u). Then $(a, b) = 1$, so there exists an integer $N \geq 1$ such that a divides N and $N = 1 \mod b$, and then $x^N = x_r^N x_u^N = x_u$, and $x^{1-N} = x x_u^{-1} = x_r$.

Let's show the uniqueness statement. So suppose that we have two paits (x_r, x_u) and (x'_r, x'_u) satisfying (a) and (b). By (a), we can find an integer $N \geq 1$ such that $x_u^{p^N} = (x'_u)^{p^N} = 1$ and $x_r^{p^N} = x_r$, $(x'_r)^{p^N} = x'_r$. Then using (b), we get $x^{p^N} = x_r = x'_r$, and this also gives $x_u = x'_u$.

Let's show the existence statement. By the first part of the proof, we may assume that G is generated by x, hence that G is cyclic. So we may assume that $G = \mathbb{Z}/n\mathbb{Z}$ and $x = 1$. Write $n = p^r m$ with p not dividing m. By the Chinese remainder theorem, $G \simeq \mathbb{Z}/p^r\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$, and we can take $x_r = (0, 1)$ and $x_u = (1, 0)$.

□

Claim : Let $\chi \in \mathscr{O} \otimes_{\mathbb{Z}} \mathrm{R}(G)$ be such that $\chi(G) \subset \mathbb{Z}$, let $x \in G$, and write $x = x_r x_u$ as above. Then $\chi(x) = \chi(x_r) \mod p$.

Proof. We may assume that $G = \langle x \rangle$. Let $\chi_1, \ldots, \chi_n$ be the characters of the irreducible representations of G over k, which are all 1-dimensional by proposition I.3.9 of chapter I. We write $\chi = \sum_{i=1}^n a_i \chi_i$, with $a_i \in \mathscr{O}$. Let $q = p^r$ be the order of x_u. Then $x^q = x_r^q$, so, for every $i \in \{1, \ldots, n\}$, $\chi_i(x)^q = \chi_i(x_r)^q$ (χ_i is compatible with multiplication because it is the character of a 1-dimensional representation). So

$$\chi(x)^q = (\sum_{i=1}^n a_i \chi_i(x))^q = \sum_{i=1}^n a_i^q \chi_i(x)^q = \sum_{i=1}^r a_i^q \chi_i(x_r)^q = \chi(x_r)^q \mod p\mathscr{O}.$$

As $\chi(x), \chi(x_r) \in \mathbb{Z}$ and $p\mathscr{O} \cap \mathbb{Z} = p\mathbb{Z}$, we get $\chi(x)^q = \chi(x_r)^q \mod p$. Finally, as we know that $a^p = a \mod p$ for every $a \in \mathbb{Z}$, this implies that $\chi(x) = \chi(x_r) \mod p$.

□

(6) If $x \in G$, we write

$$Z_G(x) = \{g \in G | gx = xg\}.$$

(This is the *centralizer of x in G*.)

Claim : Let $x \in G$ be p-regular. Let $H = C \times P$ be p-elementary, with $C = \langle x \rangle$ and $P \subset Z_G(c)$ a Sylow p-subgroup (i.e. such that p does not divide $|Z_G(x)/P|$). Then there exists $\psi \in \mathscr{O} \otimes_{\mathbb{Z}} \mathrm{R}(H)$ such that $\psi(H) \subset \mathbb{Z}$ and that, if $\psi' = \mathrm{Ind}_H^G \psi$, then :

(i) $\psi'(x) \neq 0 \mod p$;

(ii) $\psi'(s) = 0$, for every $s \in G$ a p-regular element that is not conjugate to x.

Proof. Let $c = |C|$ and $p^r = |P|$. Let

$$\psi_C : \left\{ \begin{array}{ccc} C & \to & \mathbb{Z} \\ y & \longmapsto & \left\{ \begin{array}{ll} c & \text{if } y = x \\ 0 & \text{otherwise.} \end{array} \right. \end{array} \right.$$

As $\psi_C(C) \subset c\mathbb{Z}$, $\psi_C \in \mathscr{O} \otimes_{\mathbb{Z}} \mathrm{R}(C)$ by (3). Let

$$\psi : \left\{ \begin{array}{ccc} H = C \times P & \to & \mathbb{Z} \\ (x, y) & \longmapsto & \psi_C(x) \end{array} \right.$$

Then $\psi \in \mathscr{O} \otimes_{\mathbb{Z}} \mathrm{R}(H)$. (Indeed, if $\psi_C = \sum_{V \in S_k(C)} a_V \chi_V$ with $a_V \in \mathscr{O}$, then $\psi = \sum_{V \in S_k(C)} a_V \chi_{V \otimes \mathbb{1}_H}$.)

Let $s \in G$ be p-regular. Then

$$\psi'(s) = \frac{1}{cp^r} \sum_{y \in G | ysy^{-1} \in H} \psi(ysy^{-1}).$$

Let $y \in G$. If $ysy^{-1} \in H$, then $ysy^{-1} \in C$ (because ysy^{-1} is p-regular), so $\psi(ysy^{-1}) \neq 0$ if and only $ysy^{-1} = x$. Hence $\psi'(s) = 0$ if s is not conjugate to x. Also,

$$\psi'(x) = \frac{1}{cp^r} \sum_{y \in G | yxy^{-1} = x} \psi(x) = \frac{1}{p^r} \sum_{y \in Z_g(x)} 1 = \frac{1}{p^r} |Z_G(x)| \neq 0 \mod p.$$

□

(7) Claim : There exists $\psi \in \mathscr{O} \otimes_{\mathbb{Z}} V_p$ such that $\psi(G) \subset \mathbb{Z}$ and $\psi(x) \neq 0 \mod p$ for every $x \in G$.

Proof. Let $(x_i)_{i \in I}$ be a system of representatives of the p-regular conjugacy classes in G. For every $i \in I$, we can find by (6) a $\psi_i \in \mathscr{O} \otimes_{\mathbb{Z}} V_p$ such that $\psi_i(G) \subset \mathbb{Z}$, $\psi_i(x_i) \neq 0 \mod p$ and $\psi_i(x_j) = 0$ for every $j \neq i$. Let $\psi = \sum_{i \in I} \psi_i$. Then $\psi \in \mathscr{O} \otimes_{\mathbb{Z}} V_p$ and $\psi(G) \subset \mathbb{Z}$. If $x \in G$, write $x = x_r x_u$ as in (4). Then there exists a unique $i \in I$ such that x_r is conjugate to x_i, and we have (again by (4))

$$\psi(x) = \psi(x_r) = \psi(x_i) = \psi_i(x_i) \neq 0 \mod p.$$

□

(8) As before, write $n := |G| = p^r m$ with p prime to m. Let's prove that $m \in V_p$, which is the statement of the theorem. For this, choose a $\psi \in \mathscr{O} \otimes_{\mathbb{Z}} V_p$ as in (7). Let $N = \varphi(p^r) = |(\mathbb{Z}/p^r\mathbb{Z})^\times|$. Then for every $\ell \in \mathbb{Z}$ prime to p, $\ell^N = 1 \mod p^r$. So for every $x \in G$, $\psi(x)^N = 1 \mod p^r$. So $m(\psi^N - 1) \in \mathscr{C}(G, \mathbb{Z})$ takes its values in $n\mathbb{Z}$, and by (3) this implies that $m(\psi^N - 1) \in \mathscr{O} \otimes_{\mathbb{Z}} V_p$. As $\mathscr{O} \otimes_{\mathbb{Z}} V_p$ is an ideal of $\mathscr{O} \otimes_{\mathbb{Z}} \mathrm{R}(G)$ by corollary I.5.6.2 of chapter I, $m\psi^N \in \mathscr{O} \otimes_{\mathbb{Z}} V_p$. Finally, we get $m = m\psi^N - m(\psi^N - 1) \in \mathscr{O} \otimes_{\mathbb{Z}} V_p$.

□

Definition II.5.4. A representation V of G is called *monomial* if there exists a subgroup H of G and a 1-dimensional representation W of H such that $V = \mathrm{Ind}_H^G V$.

The following corollary is often called "Brauer's theorem" too.

Corollary II.5.5. *For every representation V of G, there exist monomial representations $V_1, \ldots, V_r$ of G and integers $n_1, \ldots, n_r \in \mathbb{Z}$ such that, in $\mathrm{R}(G)$, we have an equality*

$$[V] = \sum_{i=1}^{r} n_i [V_i].$$

Thanks to theorem II.5.2, this corollary is immediate once we have the following proposition.

Proposition II.5.6. *Let p be a prime number and H be a p-elementary group. Then every irreducible representation of H is monomial.*

Lemma II.5.7. *Let P be a p-group, and suppose that P is not abelian. Denote by $Z(P)$ the center of P. Then there exists an abelian normal subgroup A of P such that $Z(P) \subsetneq A$.*

Proof. The quotient $P/Z(P)$ is a nontrivial p-group, so its center is nontrivial. Choose $A' \subset Z(P/Z(P))$ cyclic of order p, and let A be its inverse image in P. Clearly $Z(P) \subsetneq A$, and A is normal in P because it's the inverse image of a normal subgroup of $P/Z(P)$. Also, A is abelian because it is generated by $Z(P)$ and by a lift of a generator of A'.

□

Proof of the proposition. Write $H = C \times P$, with C cyclic of order prime to p and P a p-group. By theorem II.2.1, irreducible representations of H are all of the form $V_1 \otimes_k V_2$, where V_1 (resp. V_2) is an irreducible representation of C (resp. P). By proposition I.3.9 of chapter I, V_1 is 1-dimensional, so we just need to show that V_2 is monomial (If $V_2 = \mathrm{Ind}_{P'}^P W_2$, then $V_1 \otimes_k V_2 = \mathrm{Ind}_{C \times P'}^H (V_1 \otimes_k V_2$.)

So we may assume that $H = P$ is a p-group. We prove the result by induction on $|P|$. If P is abelian (for example if $|P| = p$), then every irreducible representation of P is 1-dimensional by proposition I.3.9 of chapter I. So assume that P is not abelian. Let (V, ρ) be an irreducible

representation of P. If $\operatorname{Ker}\rho \neq \{1\}$, then applying the induction hypothesis to $P/\operatorname{Ker}\rho$,[3] we see that ρ is monomial. So we may assume that ρ is faithful.

By the lemma, there exists a normal abelian subgroup A of P such that $Z(P) \subsetneq A$, where $Z(P)$ is the center of P. Let $V = V_1 \oplus \cdots \oplus V_n$ be the isotypic decomposition of $\operatorname{Res}^P_A V$. (See section I.1.7 of chapter I.) Because A is abelian, proposition I.3.9 of chapter I implies that A acts on each V_i through a morphism of groups $\rho_i : A \to k^\times$. We can't have $V_1 = V$, because otherwise $\rho(A)$ would be contained in the center of $\rho(P)$, so A would be contained in the center of P (as ρ is faithful), which contradicts the choice of A.

Let $g \in P$ and $i \in \{1, \ldots, n\}$. Then, if $v \in V_i$ and $y \in A$,

$$\rho(y)\rho(g)v = \rho(g)\rho(g^{-1}yg)v = \rho(g)\rho_i(g^{-1}yg)v = \rho_i(g^{-1}yg)\rho(g)v$$

(because A is normal in G and $\rho_i(g^{-1}yg) \in k$). So $\rho(g)$ sends V_i bijectively to the isotypic component of $\operatorname{Res}^P_A V$ corresponding to the map $H \to k^\times$, $y \longmapsto \rho_i(g^{-1}yg)$. In other words, the action of P on V permutes the V_i, so we get an action of P on the set $\{V_1, \ldots, V_n\}$. As V is irreducible, this action is transitive. Hence all the V_i are isomorphic as k-vector space, and so

$$\dim_k V_1 = \cdots = \dim_k V_n = \frac{1}{n}\dim_k V.$$

Let

$$H = \{g \in G | \rho(g)V_1 = V_1\},$$

then H is a subgroup of G and $|G/H| = n > 1$, that is, $H \not= G$. As H stabilizes V_1, we get a representation of H on V_1, which we will denote by V_H. Define $\varphi : \operatorname{Ind}^G_H V_H \to V$ by $\varphi(g \otimes v) = \rho(g)v$. This map φ is well-defined by definition of H, and it is clearly G-equivariant. It is surjective because $V = \sum_{g\in G} \rho(g)V_1$. Moreover, we have

$$\dim_k \operatorname{Ind}^G_H V_H = |G/H| \dim_k V_H = n \dim_k V_1 = \dim_k V,$$

so φ is an isomorphism, and $V \simeq \operatorname{Ind}^G_H V_H$ as representations of G. Also, as V is irreducible, V_H is irreducible. Applying the induction hypothesis to H (and the transitivity of induction), we see that V is monomial.

□

II.6 First application of Brauer's theorem : field of definition of a representation of G

In this section, we assume that k is a field of characteristic 0 and denote by $\overline{k}$ an algebraic closure of k. Remember that the map $\mathrm{R}_k(G) \to \mathrm{R}_{\overline{k}}(G)$ is injective (by corollary II.1.2.9).

[3] And using problem VII.1.14.

Theorem II.6.1. *Let m be the least common multiple of the orders of all the elements of G. Then, if k contains all the mth roots of 1 in $\overline{k}$, the map $\mathrm{R}_k(G) \to \mathrm{R}_{\overline{k}}(G)$ is an isomorphism.*

Using problem VII.2.7, we get the following reformulation.

Corollary II.6.2. *Under the hypothesis of the theorem, every representation of G over $\overline{k}$ is realizable over k. (That is, is of the form $V \otimes_k \overline{k}$, where V is a representation of G over k.)*

Proof of the theorem. Let $x \in \mathrm{R}_{\overline{k}}(G)$. By Brauer's theorem (in the form of corollary II.5.5), there exist subgroups $H_1, \dots, H_r$ of G, 1-dimensional representations $(V_1, \rho_1), \dots, (V_r, \rho_r)$ of $H_1, \dots, H_r$ over $\overline{k}$ and integers $n_1, \dots, n_r \in \mathbb{Z}$ such that

$$x = \sum_{i=1}^{r} n_i \operatorname{Ind}_{H_i}^{G} [V_i].$$

Let $i \in \{1, \dots, r\}$. For every $g \in H_i$, $\rho_i(g) \in \overline{k}^{\times}$ is a mth root of 1, so it is actually in k, and so $[V_i]$ is in the image of $\mathrm{R}_k(H_i) \to \mathrm{R}_{\overline{k}}(H_i)$. Hence x is in the image of $\mathrm{R}_k(G) \to \mathrm{R}_{\overline{k}}(G)$.

□

III Comparison between characteristic 0 theory and characteristic p theory

If k is a field of characteristic p and G is a finite group of order not prime to p, then the ring $k[G]$ is not semisimple anymore, but it is still left Artinian.

So to understand what happens a bit better, we'll start with some generalities about modules on not necessarily semisimple rings.

III.1 Indecomposable modules

In this section, R is a ring. Unless otherwise specified, any R-module will be assumed to be of finite length, hence to have a Jordan-Hölder series. See sections I.1.5 and I.1.6 of chapter I for definitions of all these terms.

Remember that we defined the *Jacobson radical* $\mathrm{rad}(R)$ of R in definition I.2.1 of chapter I as the intersection of all the maximal left ideals of R. By corollaries I.2.6 and I.2.12 of chapter I, $\mathrm{rad}(R)$ is an ideal of R, the quotient $R/\mathrm{rad}(R)$ has the same simple modules as R, and it is a semisimple ring if R is left Artinian.

III.1.1 Definitions

Definition III.1.1.1. If M is a R-module, we write $\mathrm{rad}(M) = \mathrm{rad}(R)M$. This is a submodule of M.

Remark III.1.1.2. Suppose that R is left Artinian. Then a R-module M is semisimple if and only if $\mathrm{rad}(M) = 0$.

Proof. If $\mathrm{rad}(M) = 0$, then M is a $R/\mathrm{rad}(R)$-module, so it is semisimple because $R/\mathrm{rad}(R)$ is a semisimple ring.

If M is a semisimple R-module, then $M = \bigoplus_{i\in I} M_i$ with all the M_i simple , by theorem I.1.3.4 of chapter I. On each M_i, R acts through $R/\mathrm{rad}(R)$ by proposition I.2.5 of chapter I, so $\mathrm{rad}(R)M = 0$.

□

Definition III.1.1.3. A R-module M is called *indecomposable* if for every direct sum decomposition $M = M' \oplus M''$, we have $M' = 0$ or $M'' = 0$.

Remark III.1.1.4. If M is simple, then M is indecomposable. The converse is false.

For example, take $R = \mathbb{F}_p[\mathbb{Z}/p\mathbb{Z}]$ and $M = {}_RR$. Then M is indecomposable but not simple.

III.1.2 Noncommutative local rings

Definition III.1.2.1. A (possibly noncommutative) ring S is called *local* if $S \neq \{0\}$ and S has a unique maximal left ideal.

Remark III.1.2.2. If S is a commutative ring, then it is local in the sense of definition III.1.2.1 if and only if it is local in the usual sense.

Theorem III.1.2.3. *Let S be a ring. The following conditions are equivalent :*

1. *S is local.*
2. *S has a unique maximal right ideal.*
3. *$\mathrm{rad}(S)$ is a maximal left ideal of S.*
4. *$\mathrm{rad}(S)$ is a maximal right ideal of S.*
5. *$S \neq \{0\}$ and, for every $x \in S$, either x or $1 - x$ is invertible.*
6. *$S/\mathrm{rad}(S)$ is a division algebra.*
7. *$S \neq \{0\}$, and every $x \in S - \mathrm{rad}(S)$ is invertible.*

Note that if S is local, then $\mathrm{rad}(S)$ is the unique maximal left ideal and the unique maximal right ideal of S. It is also the unique maximal ideal of S, but in the noncommutative case, a ring that has a unique maximal ideal is not necessarily local. [1]

Proof. The equivalence of (i) and (iii) follows from the definition of $\mathrm{rad}(S)$, and the equivalence of (ii) and (iv) follows from corollary I.2.9 of chapter I. Also, (vii) implies (v) by corollary I.2.8 of chapter I (which implies that $1 - x$ is invertible if $x \in \mathrm{rad}(S)$), and it's clear that (vii) implies (vi).

Let's prove that (iii) implies (vii). Let $x \in S - \mathrm{rad}(S)$. We want to show that x is invertible. We have $Rx \not\subset \mathrm{rad}(S)$. As $\mathrm{rad}(S)$ is a maximal left ideal of R, this implies that $Rx = R$, so there exists $y \in R$ such that $yx = 1$. If $xy \in \mathrm{rad}(S)$, then, by corollary I.2.8 of chapter I, $1 - yxy = 0$ is invertible, which is impossible because $S \neq \{0\}$; so $xy \notin \mathrm{rad}(S)$. Reasoning as

[1] This condition is actually equivalent to the fact that $S/\mathrm{rad}(S)$ is a simple ring, so a counterexample is a simple ring that is not a division algebra. See problem VII.1.9 for an example of such a ring.

before, we see that there exists $z \in S$ such that $zxy = 1$. So y is left and right invertible, hence invertible, and $x = y^{-1}$ is also invertible.

A similar reasoning shows that (iv) implies (vii).

Let's prove that (v) implies (iii), (iv) and (vi). Let $x \in S - \mathrm{rad}(S)$. By corollary I.2.8 of chapter I, there exist $y, z \in S$ such that $1 - yxz$ is not invertible. By condition (v), this implies that yxz is invertible, hence that $x = y^{-1}z^{-1}$ is also invertible. So any left (resp. right) ideal of S that strictly contains $\mathrm{rad}(S)$ contains an invertible element, hence is equat to S. This gives (iii) and (iv). For (vi), we have just seen that any element of S that is not in $\mathrm{rad}(S)$ is invertible, so every nonzero element of $S/\mathrm{rad}(S)$ is invertible.

Let's prove that (vi) implies (iii). Let $x \in S - \mathrm{rad}(S)$. By (vi), there exists $y \in S$ such that $yx \in 1 + \mathrm{rad}(S)$ and $xy \in 1 + \mathrm{rad}(S)$. By corollary I.2.8 of chapter I, this implies that xy and yx are invertible, hence that x (and y) are invertible. So any left ideal strictly containing $\mathrm{rad}(S)$ contains an invertible element, which gives (iii) as before.

□

Let's consider the particular case of left Artinian rings.

Proposition III.1.2.4. *Let S be a left Artinian ring. Then $\mathrm{rad}(S)^N = 0$ for N big enough. In fact, if $n = \lg({}_SS)$, then $\mathrm{rad}(S)^n = 0$.*

In particular, every element of $\mathrm{rad}(S)$ is nilpotent.

Proof. Let $S = I_0 \supset I_1 \supset \cdots \supset I_n = 0$ be a Jordan-Hölder series for ${}_SS$. The I_i are left ideals of S, and I_i/I_{i+1} is a simple S-module for every $i \in \{0, \ldots, n-1\}$. As $\mathrm{rad}(S)$ annihilates every simple S-module, we have $\mathrm{rad}(S)I_i \subset I_{i+1}$ for every $i \in \{0, \ldots, n-1\}$, and so $\mathrm{rad}(S)^n = \mathrm{rad}(S)^n I_0 \subset I_n = 0$.

□

Corollary III.1.2.5. *Let S be a left Artinian ring. Then the following are equivalent :*

1. *S is local.*
2. *Every element of S is nilpotent or invertible.*

Proof.

(i)⇒(ii) Let $x \in S$, and suppose that x is not nilpotent. Then $x \notin \mathrm{rad}(S)$ by the proposition, so x is invertible by theorem III.1.2.3.

(ii)⇒(i) If $x \in S$ is nilpotent, then the sum $\sum_{n \geq 0} x^n$ is finite, hence defines an element of S, and this element is an inverse of $1 - x$. So (ii) implies that x or $1 - x$ is invertible for every $x \in S$, and theorem III.1.2.3 says that S is local.

□

III.1.3 Fitting's lemma

Proposition III.1.3.1 (Fitting's lemma)**.** *Let R be a ring and M be a R-module of finite length. Let $f \in \mathrm{End}_R(M)$. Then, for n big enough,*

$$M = \mathrm{Ker}(f^n) \oplus \mathrm{Im}(f^n).$$

Proof. As M has finite length, every non-increasing (or non-decreasing) sequence of R-submodules of M has to stabilize, by proposition I.1.6.2. Applying this to the non-decreasing sequence $(\mathrm{Ker}(f^n))_{n\geq 0}$ and the non-increasing sequence $(\mathrm{Im}(f^n))_{n\geq 0}$, we find an integer $N \geq 0$ such that, for every $n \geq N$, $\mathrm{Ker}(f^n) = \mathrm{Ker}(f^{n+1})$ and $\mathrm{Im}(f^n) = \mathrm{Im}(f^{n+1})$.

Let $n \geq N$, and let's show that $M = \mathrm{Ker}(f^n) \oplus \mathrm{Im}(f^n)$.

First, if $x \in M$, then $f^n(x) \in \mathrm{Im}(f^{2n})$, so there exists $y \in M$ such that $f^n(x) = f^{2n}(y)$, and so $x - f^n(x) \in \mathrm{Ker}(f^n)$ and $x = f^n(x) + (x - f^n(x)) \in \mathrm{Im}(f^n) + \mathrm{Ker}(f^n)$. This proves that $M = \mathrm{Im}(f^n) + \mathrm{Ker}(f^n)$.

Now take $x \in \mathrm{Ker}(f^n) \cap \mathrm{Im}(f^n)$. Write $x = f^n(y)$ with $y \in M$. Then $y \in \mathrm{Ker}(f^{2n}) = \mathrm{Ker}(f^n)$, so $x = f^n(y) = 0$. This proves that $\mathrm{Ker}(f^n) \cap \mathrm{Im}(f^n) = 0$.

□

Corollary III.1.3.2. *Let R and M be as in the proposition. Then M is indecomposable if and only if $\mathrm{End}_R(M)$ is local.*

Proof. Suppose that M is indecomposable, and let $f \in \mathrm{End}_R(M)$. Then there exists an integer $n \geq 1$ such that $M = \mathrm{Ker}(f^n) \oplus \mathrm{Im}(f^n)$. As M is indecomposable, either $\mathrm{Im}(f^n) = 0$, and then f is nilpotent, or $\mathrm{Ker}(f^n) = 0$ and $\mathrm{Im}(f^n) = M$, and then f^n is invertible, and so is f.

Suppose that M is not indecomposable, and write $M = M' \oplus M''$, with $M', M'' \neq 0$. Let $\pi : M \to M'$ be the projection with kernel M''. Then π is not invertible because $\mathrm{Ker}\,\pi = M'' \neq 0$, and $1 - \pi$ is not invertible because $\mathrm{Ker}(1 - \pi) = M' \neq 0$. So $\mathrm{End}_R(M)$ is not local.

□

III.1.4 Krull-Schmidt-Remak theorem

Theorem III.1.4.1. *Let M be a R-module (of finite length). Then we have $M = M_1 \oplus \cdots \oplus M_r$, with the M_i indecomposable. Morover, the M_i are uniquely determined up to reordering.*

Proof. Existence of the decomposition : We do an induction on $\lg(M)$. If $\lg(M) = 1$, then M is simple and the result is obvious. If $\lg(M) \geq 2$ and M is not indecomposable, write

$M = M' \oplus M''$ with $M', M'' \neq 0$. Then $\lg(M'), \lg(M'') < \lg(M)$, so we can apply the induction hypothesis to M' and M'' to get the result.

Uniqueness of the decomposition : Assume that $M = M_1 \oplus \cdots \oplus M_r = M'_1 \oplus \cdots \oplus M'_S$ with all the M_i and M'_j indecomposable. We do an induction on r. If $r = 1$, M is indecomposable and the result is obvious. Suppose that $r \geq 2$. For every $j \in \{1, \ldots, s\}$, let $u_j \in \mathrm{End}_R(M_1)$ be the composition $M_1 \hookrightarrow M \twoheadrightarrow M'_j \hookrightarrow M \twoheadrightarrow M_1$, where all the maps are obvious injections or projections. Then $\mathrm{id}_{M_1} = \sum_{j=1}^{s} u_j$, so there exists j such $u_j \notin \mathrm{rad}(\mathrm{End}_R(M_1))$. We may assume that $j = 1$.

As M_1 is indecomposable, $\mathrm{End}_R(M_1)$ is local by corollary III.1.3.2, so u_1 is invertible by theorem III.1.2.3. Write $u_1 = vw$, where w is the composition $M_1 \hookrightarrow M \twoheadrightarrow M'_1$ and v is the composition $M'_1 \hookrightarrow M \twoheadrightarrow M_1$. Then $(u_1^{-1}v)w = \mathrm{id}_{M_1}$, so w is injective and $M'_1 = w(M_1) \oplus \mathrm{Ker}(u_1^{-1}v)$. As M'_1 is indecomposable and $w(M_1) \neq 0$, $\mathrm{Ker}(u_1^{-1}v) = 0$, hence w is an isomorphism, and so is v.

Let $x \in M_1 \cap (\bigoplus_{j\geq 2} M'_j)$. Then the projection of x on M'_1 is 0, so $w(x) = 0$, so $x = 0$ since w is injective. Hence M_1 and $\bigoplus_{j\geq 2} M'_j$ are in direct sum. Moreover, if $x \in M$, then the surjectivity of w implies that there exists $x_1 \in M_1$ such that $x - x_1 \in \sum_{j\geq 2} M'_j$. So $M = M_1 + \sum_{j\geq 2} M'_j$. Finally, we get $M = M_1 \oplus \bigoplus_{j\geq 2} M'_j$ and $M_1 \simeq M'_1$. The result now follows from the induction hypothesis, applied to $M/M_1 \simeq \bigoplus_{i=2}^{r} M_j \simeq \bigoplus_{j=2}^{s} M'_j$.

□

III.1.5 Projective indecomposable modules

Remember that projective modules are defined in definition I.1.3.10 of chapter I.

Proposition III.1.5.1. *Let P be a projective R-module, M be a R-module and I be an ideal of R. Then the reduction modulo I map*

$$\mathrm{Hom}_R(P, M) \to \mathrm{Hom}_R(P/IP, M/IM) = \mathrm{Hom}_{R/I}(P/IP, M/IM)$$

is surjective.

Proof. Denote by $\pi : P \to P/IP$ and $\pi' : M \to M/IM$ the projections. Let $u \in \mathrm{Hom}_R(P/IP, M/IM)$. We have a commutative diagram

$$\begin{array}{ccc} P & \overset{u'}{\dashrightarrow} & M \\ {\scriptstyle\pi}\downarrow & \searrow^{u\pi} & \downarrow{\scriptstyle\pi'} \\ P/IP & \xrightarrow[u]{} & M/IM \end{array}$$

As π' is surjective and P is projective, there exists a map $u' : P \to M$ such that $\pi' u' = u\pi$.

□

Corollary III.1.5.2. *Assume that R is left Artinian. Let P be a projective R-module of finite length. Then P is indecomposable if and only if $P/\operatorname{rad}(P)$ is simple.*

Lemma III.1.5.3. *If M is a nonzero R-module of finite length, then $\operatorname{rad}(M) \neq M$.*

Proof. Let $M' \subsetneq M$ be a maximal proper submodule. Then M/M' is simple, so $\operatorname{rad}(R)$ acts trivially on M/M', so $\operatorname{rad}(M) = \operatorname{rad}(R)M \subset M'$.

□

Proof of the corollary. If P is not indecomposable, write $P = P_1 \oplus P_2$ with P_1, P_2 nonzero. Then $P/\operatorname{rad}(P) = P_1/\operatorname{rad}(P_1) \oplus P_2/\operatorname{rad}(P_2)$, and $P_1/\operatorname{rad}(P_1), P_2/\operatorname{rad}(P_2) \neq 0$ by the lemma, so $P/\operatorname{rad}(P)$ is not simple.

Now assume that $M := P/\operatorname{rad}(P)$ is not simple. As R is left Artinian, $R/\operatorname{rad}(R)$ is semisimple, so M is a semisimple module, and so we can write $M = M_1 \oplus M_2$ with $M_1, M_2 \neq 0$. Let $\pi_1 \in \operatorname{End}_R(M_1)$ be the composition $M \twoheadrightarrow M_1 \hookrightarrow M$, where the maps are the obvious projection and inclusion. By the proposition, there exists $\pi \in \operatorname{End}_R(P)$ such that $\pi \mod \operatorname{rad}(P) = \pi_1$. Then neither π nor $1 - \pi$ are invertible (because neither π_1 nor $1 - \pi_1$ are), so $\operatorname{End}_R(P)$ is not local, so P cannot be indecomposable by corollary III.1.3.2.

□

Notation III.1.5.4. We write $PI(R)$ for the set of isomorphism classes of finite length projective indecomposable R-modules, and $S(R)$ for the set of isomorphism classes of simple R-modules.

Proposition III.1.5.5. *Suppose that R is left Artinian and left Noetherian.*[2] *Then the map $P \longmapsto P/\operatorname{rad}(P)$ induce a bijection $PI(R) \to S(R)$.*

Proof. This map is well-defined by corollary III.1.5.2.

Let's show that it is surjective. Let M be a simple R-module. Any $x \in M - \{0\}$ gives a surjective map $R \twoheadrightarrow M$ (sending $a \in R$ to ax). As ${}_RR$ is a R-module of finite length, we can apply the Krull-Schmidt-Remark theorem to it and write $R = P_1 \oplus \cdots \oplus P_n$, where the P_i are indecomposable module that are automatically projective as direct summands of a free module. Then $R/\operatorname{rad}(R) = \bigoplus_{i=1}^n P_i/\operatorname{rad}(P_i)$ surjects to M, and every $P_i/\operatorname{rad}(P_i)$ is simple by corollary III.1.5.2, so M is isomorphic to one of $P_i/\operatorname{rad}(P_i)$ by Schur's lemma (theorem I.1.4.1 of chapter I).

Let's show that the map of the proposition is injective. Let P, P' be two projective indecomposable modules of finite length, and suppose that we have a R-module isomorphism $u : P/\operatorname{Ker}(P) \xrightarrow{\sim} P'/\operatorname{Ker}(P')$. By proposition III.1.5.1, there exists a R-module map $u' : P \to P'$ such that $u = u' \mod \operatorname{rad}(R)$. Let $N \subset P'$ be a proper maximal submodule. Then P'/N is simple, so $\operatorname{rad}(R)(P'/N) = 0$, so $\operatorname{rad}(P') = \operatorname{rad}(R)P' \subset N$. As $P'/\operatorname{rad}(P')$ is

[2] Note that "left Artinian" implies "left Noetherian", see theorem (4.15) of Lam's book [20].

simple, this implies that $N = \mathrm{rad}(P')$, and that $\mathrm{rad}(P')$ is the unique proper maximal submodule of P. As $u'(P) \not\subset \mathrm{rad}(P')$ (because u is surjective), $u'(P) = P'$, and so u' is surjective. As P' is projective, we have $P \simeq P' \oplus \mathrm{Ker}(u')$. But P is indecomposable, so $\mathrm{Ker}(u') = 0$, and u' is an isomorphism.

□

Remark III.1.5.6. If M is a simple R-module, its inverse image in $PI(R)$ is a "minimal" projective R-module such that $P \twoheadrightarrow M$. This is called a *projective envelope* or *projective cover* of M. In fact, projective envelopes exist for any finite length R-module, and they are unique up to isomorphism. For more about them, see section 24 of Lam's book [20] (projective covers are introduced in definition (24.9).)

III.1.6 Lifting of idempotents

Definition III.1.6.1. Let S be a ring. An element $e \in S$ is called *idempotent* if $e^2 = 2$.

Theorem III.1.6.2. *Let S be a ring, and let I be an ideal of S such that every element of I is nilpotent.* [3] *Let $\overline{e} \in S/I$ be idempotent. Then there exists $e \in S$ idempotent such that $\overline{e} = e \mod I$.*

Proof. Note that $\overline{f} := 1 - \overline{e}$ is also idempotent, and that we have $\overline{e}\overline{f} = \overline{f}\overline{e} = 0$. (We say that $\overline{e}$ and $\overline{f}$ are *orthogonal idempotents*.) The idea is to try to lift both $\overline{e}$ and $\overline{f}$ to orthogonal idempotents of S.

Let e be any lift of S, and let $f = 1 - e$. Then $ef = fe \in I$, and $e + f = 1 \mod I$. By the assumption on I, there exists $k \geq 1$ such that $(ef)^k = e^k f^k = 0$. Note that $e^k = \overline{e}^k = \overline{e} \mod I$. Let $e' = e^k$ and $f' = f^k$. Then $e'f' = f'e' = 0$, and $e' + f' = e^k + f^k = \overline{e} + \overline{f} = 1 \mod I$. Let $x = 1 - (e' + f')$. As $x \in I$, there exists $n \geq 1$ such that $x^n = 0$. So $u := 1 + x + x^2 + \cdots + x^{n-1}$ is an inverse of $1 - x = e' + f'$, and it commutes with e' and f' (because x does). Let $e'' = ue'$ and $f'' = uf'$. Then $e'' = \overline{e} \mod I$ (because $u = 1 \mod I$), we have $e''f'' = f''e'' = 0$, and $e'' + f'' = u(e' + f') = 1$. So $(e'')^2 = (e'')^2 + e''f'' = e''(e'' + f'') = e''$, and we have our idempotent lift of $\overline{e}$.

□

Corollary III.1.6.3. *Let S be a ring and I be an ideal of S. Suppose that the obvious map $S \to \hat{S} := \varprojlim_n S/I^n$ is an isomorphism. Then any idempotent of S/I lifts to an idempotent of S.*

Proof. Let $\overline{e} \in S/I$ be idempotent. By the theorem, we can construct by induction on $n \geq 1$ a sequence of idempotents $e_n \in S/I^n$ such that $e_1 = \overline{e}$ and that e_{n+1} lifts e_n for every n. Then $e := (e_n)_{n \geq 1}$ is an element of $\hat{S}$, and its preimage in S is an idempotent lifting $\overline{e}$.

[3] This is called a *nil ideal*, and is not the same as a nilpotent ideal.

□

III.2 Application to representation rings in positive characteristic

Let R be a ring. Remember that we defined (in definition I.4.1 of chapter I) $PK(R)$ to be the quotient of the free abelian group on the basis elements $[P]$, for P a finite length projective R-module, by all the relations of the form $[P] = [P'] + [P'']$, where $0 \to P' \to P \to P'' \to 0$ is an exact sequence of R-modules.

Corollary III.2.1. *1. $PK(R)$ is a free $\mathbb{Z}$-module with basis $([P])_{P \in PI(R)}$.*

2. If P and P'' are projective R-modules of finite length, then $P \simeq P'$ as R-modules if and only if $[P] = [P']$ in $PK(R)$.

Proof. Point (i) is proved just as for $K(R)$ (see the proof of proposition I.4.4 of chapter I).

To prove (ii), take P and P' as in the statement, and write $P = P_1 \oplus \cdots \oplus P_r$, $P' = P'_1 \oplus \cdots \oplus P'_s$, with the P_i and the P'_j indecomposable (by theorem III.1.4.1). The P_i and P'_j are also automatically projective, so they are in $PI(R)$.

By theorem III.1.4.1 again, $P \simeq P'$ if and only if there exists a bijection $\sigma : \{1, \ldots, r\} \xrightarrow{\sim} \{1, \ldots, s\}$ such that $P_i \simeq P'_{\sigma(i)}$ for every $i \in \{1, \ldots, r\}$. By point (i), this is equivalent to $[P] = [P']$.

□

We now suppose that k is a field and that G is a group.

We write $\mathrm{P}_k(G)$ for $KP(k[G])$ (as in definition I.4.6 of chapter I), and we also write $PI_k(G)$ for $PI(k[G])$.

Remark III.2.2. We have seen in corollary I.4.9 of chapter I that the tensor product over k defines a $\mathrm{R}_k(G)$-module structure on $\mathrm{P}_k(G)$, and that the obvious map $\mathrm{P}_k(G) \to R_k(G)$ is $R_k(G)$-linear.

Remark III.2.3. By proposition I.4.4 of chapter I, proposition III.1.5.5 and corollary III.2.1, $\mathrm{R}_k(G)$ and $\mathrm{P}_k(G)$ are free $\mathbb{Z}$-modules of the same rank. But we still do not know what the map $\mathrm{P}_k(G) \to \mathrm{R}_k(G)$ is like ! (Unless G is finite and $\mathrm{char}(k) \not| |G|$, then it is just the identity.)

In fact, we can prove that this map is injective and that $\mathrm{R}_k(G)/\mathrm{P}_k(G)$ is a finite p-group, where $p = \mathrm{char}(k)$, but this is far from obvious. See theorem III.8.2.

III.3 Representations over discrete valuation rings

Let Λ be a commutative ring and G be a finite group. We assume that every $\Lambda[G]$-module is of finite type over Λ (unless otherwise specified).

Definition III.3.1. We denote by $\mathrm{P}_\Lambda(G)$ the quotient of the free $\mathbb{Z}$-module on all the $[P]$, for P a projective $\Lambda[G]$-module that is of finite type as a Λ-module, by the relations $[P] = [P'] + [P'']$, for every exact sequence $0 \to P' \to P \to P'' \to 0$.

Remark III.3.2. If $\Lambda \to \Lambda'$ is a morphism of commutative rings, then $P \longmapsto P \otimes_\Lambda \Lambda'$ induces a morphism of groups $\mathrm{P}_\Lambda(G) \to \mathrm{P}_{\Lambda'}(G)$.

Indeed, if P is a projective $\Lambda[G]$-module, then it is a direct summand of some free $\Lambda[G]$-module F, and then $P \otimes_\Lambda \Lambda'$ is a direct summand of the free $\Lambda'[G]$-module $F \otimes_\Lambda \Lambda'$.

Proposition III.3.3. *Let P be a $\Lambda[G]$-module. Then the following are equivalent :*

1. *P is a projective $\Lambda[G]$-module.*
2. *P is projective as a Λ-module, and there exists $u \in \mathrm{End}_\Lambda(P)$ such that*

$$\forall x \in P,\ \sum_{g \in G} gu(g^{-1}x) = x.$$

Proof. We write P_0 for P seen as a Λ-module. Let $Q = \Lambda \otimes_\Lambda P_0$. We have a surjective $\Lambda[G]$-linear map $q : Q \twoheadrightarrow P$, $x \otimes y \longmapsto xy$.

I claim that the map $\varphi : \mathrm{End}_\Lambda(P_0) \to \mathrm{Hom}_{\Lambda[G]}(P, Q)$ sending u to $\sum_{g \in G} g \otimes ug^{-1} : x \longmapsto \sum_{g \in G} g \otimes u(g^{-1}x)$ is well-defined and an isomorphism of Λ-mdules.

Indeed, it is easy to see that $\varphi(u)$ is $\Lambda[G]$-linear for every $u \in \mathrm{End}_\Lambda(P_0)$, so φ is well-defined.

Let's show that φ is injective. Let $u \in \mathrm{End}_\Lambda(P_0)$. Note that $Q = \bigoplus_{g \in G} g \otimes P_0$ as a Λ-module. So, if $x \in P$ is such that

$$0 = \varphi(u)(x) = \sum_{g \in G} g \otimes u(g^{-1}x),$$

we have $u(g^{-1}x) = 0$ for every $g \in G$, and in particular $u(x) = 0$. Hence we have $u = 0$ if $\varphi(u) = 0$.

Let's show that φ is surjective. Let $v \in \mathrm{Hom}_{\Lambda[G]}(P, Q)$. Then we can write $v(x) = \sum_{g \in G} g \otimes u_g(x)$ for every $x \in P$, with the u_g in $\mathrm{End}_\Lambda(P_0)$. By $\Lambda[G]$-linearity of v, for every $h \in G$ and $x \in P$,

$$v(h^{-1}x) = \sum_{g \in G} g \otimes u_g(h^{-1}x) = \sum_{g \in G} (h^{-1}g) \otimes u_g(x),$$

so $u_g(x) = u_{h^{-1}g}(h^{-1}g)$ for every $g, h \in G$ and $x \in P$. In particular, $u_g(x) = u_1(g^{-1}x)$ for every $g \in G$ and $x \in P$, and so $v = \varphi(u_1)$.

Now we come back to the proof of the proposition.

(i)⇒(ii) If P is projective, it's a direct factor of some free module $\Lambda[G]^{(I)}$. As $\Lambda[G]$ is a free Λ-module, $\Lambda[G]^{(I)}$ is also free as a Λ-module, so P_0 is projective. Also, as $q : Q \to P$ is surjective and Λ-linear, there exists a Λ-linear map $s : P \to Q$ such that $qs = \mathrm{id}_P$. Write $s = \varphi(u)$, with $u \in \mathrm{End}_\Lambda(P_0)$. Then $\mathrm{id}_P = qs = \sum_{g\in G} gug^{-1}$, which gives (ii).

(ii)⇒(i) If P_0 is a projective Λ-module, then Q is a projective $\Lambda[G]$-module. Also, $s := \varphi(u) : P \to Q$ satisfies $qs = \mathrm{id}_P$, so P is a direct summand of Q, hence is also a projective $\Lambda[G]$-module.

□

Now we will specialize to the case that Λ is a discrete valuation ring. Remember that discrete valuation rings were defined in problem VII.3.1.

Theorem III.3.4. *Suppose that Λ is a discrete valuation ring with residue field k and maximal ideal $\mathfrak{m}$.*

1. *If P is a $\Lambda[G]$-module that is free of finite type[4] over Λ, then P is a projective $\Lambda[G]$-module if and only if $\overline{P} := P \otimes_\Lambda k$ is a projective $k[G]$-module.*
2. *If P and P' are projective $\Lambda[G]$-modules, then $P \simeq P'$ as $\Lambda[G]$-modules if and only if $P \otimes_\Lambda k \simeq P' \otimes_\Lambda k$ as $k[G]$-modules.*
3. *Suppose that the discrete valuation ring Λ is* complete. *(See problem VII.3.3.) If $\overline{P}$ is a projective $k[G]$-module, then there exists a unique (up to isomorphism) projective $\Lambda[G]$-module P such that $\overline{P} \simeq P \otimes_\Lambda k$.*

Proof. 1. We already know that $\overline{P}$ is a projective $k[G]$-module if P is a projective $\Lambda[G]$-module. Let's prove the converse. Suppose that $\overline{P}$ is a projective $k[G]$-module. By proposition III.3.3, there exists $\overline{u} \in \mathrm{End}_k(\overline{P})$ such that $\sum_{g\in G} g\overline{u}g^{-1} = \mathrm{id}_{\overline{P}}$. As P is a projective Λ-module, there exists by proposition III.1.5.1 a $u \in \mathrm{End}_\Lambda(P)$ lifting $\overline{u}$. Then $u' := \sum_{g\in G} gug^{-1} \in \mathrm{End}_{\Lambda[G]}(P)$ is equal to $\mathrm{id}_{\overline{P}}$ modulo $\mathfrak{m}$. So $\det(u') = 1 \mod \mathfrak{m}$, so $\det(u') \in \Lambda'$, so u' is invertible, and we have $\sum_{g\in G} g(u(u')^{-1})g^{-1} = \mathrm{id}_P$. By proposition III.3.3 again, P is a projective $\Lambda[G]$-module.

2. Let $\overline{u} : P \otimes_\Lambda k \xrightarrow{\sim} P' \otimes_\Lambda k$ be an isomorphism of $k[G]$-modules. By propositions III.1.5.1, there exists a $\Lambda[G]$-module map $u : P \to P'$ lifting $\overline{u}$. We want to show that u is invertible, and for this it suffices to show that it is an isomorphism of Λ-modules. This follows from Nakayama's lemma, but we can also do it directly : We know that P and P' are projective Λ-modules of finite type, hence they are free Λ-modules of finite type. Their ranks are

[4]Equivalently, projective of finite type.

equal, because they are equal to the dimension over k of $P \otimes_\Lambda k \simeq P' \otimes_\Lambda k$. If we choose Λ-bases of P and P', then the matrix A of u in these bases is square, and $\det(A) \neq 0 \mod \mathfrak{m}$ because $\overline{u}$ is invertible. Hence $\det(A) \in \Lambda^\times$, and u is an isomorphism of Λ-modules.

3. The uniqueness follows from (ii). For the existence, using the fact that $\overline{P}$ is projective (and of finite type as a k-module), write $\overline{F} = \overline{P} \otimes \overline{P}'$, with $\overline{F} = k[G]^{\oplus n}$, let $F = \Lambda[G]^{\oplus n}$, and let $B = \mathrm{End}_{\Lambda[G]}(F)$. Then the map $B \to \mathrm{End}_{k[G]}(\overline{F})$ is surjective by proposition III.1.5.1, so it identifies $\mathrm{End}_{k[G]}(\overline{F})$ with $B/\mathfrak{m}B$. Let $\overline{e} \in \mathrm{End}_{k[G]}(\overline{F})$ be the projection on $\overline{P}$ with kernel $\overline{P}'$. By corollary III.1.6.3 (which applies because B is a free Λ-module of finite type, so its ideal $\mathfrak{m}B$ satisfies the condition of the corollary), there exists an idempotent $e \in B$ lifting $\overline{e}$. Then we have $F = \mathrm{Im}(e) \oplus \mathrm{Ker}(e)$, and $P := \mathrm{Im}(e)$ is a projective $\Lambda[G]$-module such that $P \otimes_\Lambda k = \mathrm{Im}(\overline{e}) = \overline{P}$.

□

Corollary III.3.5. *If Λ is a discrete valuation ring with residue field k, then $P \to P \otimes_\Lambda k$ induces an injective map $\psi : \mathrm{P}_\Lambda(G) \to \mathrm{P}_k(G)$. This map is bijective if Λ is a complete discrete valuation ring.*

Proof. We have already seen that the map exists. Let's show that it is injective. Every element of $\mathrm{P}_\Lambda(G)$ can be written as $[P] - [P']$ where P, P' are two projective $\Lambda[G]$-modules. So let P, P' be projective $\Lambda[G]$-modules such that $[P \otimes_\Lambda k] = [P' \otimes_\Lambda k]$. By corollary III.2.1, this implies that $P \otimes_\Lambda k \simeq P' \otimes_\Lambda k$ as $k[G]$-modules. By (ii) of the theorem, this implies that $P \simeq P'$ as $\Lambda[G]$-modules. Finally, the last sentence follows directly from (iii) of the theorem.

□

III.4 The cde triangle

In this section, we fix a complete discrete valuation ring Λ with uniformizing element ϖ, maximal ideal $\mathfrak{m} = (\varpi)$, residue field k and fraction field K, and we assume that $\mathrm{char}(K) = 0$.

We want to construct a commutative triangle

$$\begin{array}{ccc} \mathrm{P}_k(G) & \xrightarrow{\quad c \quad} & \mathrm{R}_k(G) \\ & {\scriptstyle e} \searrow \quad \nearrow {\scriptstyle d} & \\ & \mathrm{R}_K(G) & \end{array}$$

The map c is the obvious map, and e is the composition

$$\mathrm{P}_k(G) \overset{\psi^{-1}}{\to} \mathrm{P}_\Lambda(G) \overset{\cdot \otimes_\Lambda K}{\to} \mathrm{P}_K(G) = \mathrm{R}_K(G).$$

Let's now define d. Let V be a $K[G]$-module. Let $M_1 \subset V$ be a Λ-lattice, i.e. a finite type Λ-submodule of V such that $KM_1 = V$. After replacing M_1 by $\sum_{g\in G} gM_1$, we may assume that M_1 is stable by G. Then $\overline{M}_1 := M_1 \otimes_\Lambda k$ is a $k[G]$-module. This $k[G]$-module obviously depends on the choice of M_1, but we have the following :

Theorem III.4.1. *With notation as in the paragraph above, $[\overline{M}_1] \in R_k(G)$ only depends on V.*

Hence we can define d by setting $d([V]) = [\overline{M}_1]$ (notation as above) and extending by additivity.

Remark III.4.2. Suppose that $G \subset G'$, with G' another finite group. If M_1 is a G-stable Λ-lattice in V, then $M_1' := \mathrm{Ind}_G^{G'} M_1$ is a G'-stable Λ-lattice in $\mathrm{Ind}_G^{G'} V$, and so we have

$$d([\mathrm{Ind}_G^{G'} V]) = [(\mathrm{Ind}_G^{G'} M_1' \otimes_\Lambda k)] = [\mathrm{Ind}_G^{G'}(M_1' \otimes_\Lambda k)] = \mathrm{Ind}_G^{G'} d([V]).$$

In other words, d is compatible with induction.

Proof of the theorem. Let M_2 be another G-stable Λ-lattice of V.

Case where $\varpi M_1 \subset M_2 \subset M_1$: Let $N = M_1/M_2$. This is a $k[G]$-module (because $\varpi M_1 \subset M_2$). From $\varpi M_2 \subset \varpi M_1 \subset M_2 \subset M_1$, we get an exact sequence of $k[G]$-modules

$$0 \to N \to \overline{M}_2 := M_2 \otimes_\Lambda k = M_2/\varpi M_2 \to \overline{M}_1 \to N \to 0.$$

So, in $\mathrm{R}_k(G)$, we get $[N] - [\overline{M}_2] + [\overline{M}_2] - [N] = 0$, hence $[\overline{M}_1] = [\overline{M}_2]$.

General case : Multiplying M_2 by a high enough power of ϖ, we may assume that $M_2 \subset M_1$. There also exists $n \geq 1$ such that $\varpi^n M_1 \subset M_2$. We prove the result by induction on n. We already did the case $n = 1$, so suppose that $n \geq 2$, and let $M_3 = \varpi^{n-1} M_1 + M_2$. Then :

- $\varpi^{n-1} M_1 \subset M_3 \subset M_1$, so $[M_3 \otimes_\Lambda k] = [\overline{M}_1]$ by the induction hypothesis;
- $\varpi M_3 \subset M_2 \subset M_3$, so $[M_3 \otimes_\Lambda k] = [M_2 \otimes_\Lambda k]$ by the case $n = 1$.

Putting these two together, we are done.

□

III.5 Representations over a field of characteristic $p \not| |G|$

We keep the notation of the previous section, and we also assume that $p := \mathrm{char}(k)$ does not divide $|G|$.

Theorem III.5.1. *1. Every $k[G]$-module is semisimple.*

2. Every $\Lambda[G]$-module that is projective as a Λ-module is projective as a $\Lambda[G]$-module.

3. *The map $d : \mathrm{R}_K(G) \to \mathrm{R}_k(G)$ is an isomorphism, and so are c and e. Also, d induces a bijection $S_K(G) \xrightarrow{\sim} S_k(G)$.*

Proof. Point (i) follows from theorem I.3.2 of chapter I, and point (ii) follows from (i) and theorem III.3.4. By (i), $\mathrm{P}_k(G) = \mathrm{R}_k(G)$, and c is just the identity morphism. Obviously, $de = \mathrm{id}_{\mathrm{R}_k(G)}$ and $e([S_k(G)]) \subset S_K(G)$.

So we just need to show that d is injective. Let V, V' be two $K[G]$-modules such that $d([V] - [V']) = 0$, let $M \subset V$, $M' \subset V'$ be G-stable Λ-lattices, and let $\overline{M} = M \otimes_\Lambda k$, $\overline{M}' = M' \otimes_\Lambda k$. We have $d([V] - [V']) = [\overline{M}] - [\overline{M}'] = 0$ in $\mathrm{R}_k(G)$. As $p \not| |G|$, this implies that $\overline{M} \simeq \overline{M}'$ as $k[G]$-modules, hence $M \simeq M'$ as $\Lambda[G]$ by theorem III.3.4, and so $V \simeq V'$ and $[V] - [V'] = 0$.

□

III.6 Brauer's theorem in positive characteristic

Keep the notation and assumptions of section III.4, and assume that k is algebraically closed and K contains all $|G|$th roots of 1 in $\overline{K}$. (See section 17.2 of Serre's book [29] for a version of this theorem that doesn't assume k algebraically closed.)

By theorem II.6.1 of chapter II, the map $R_K(H) \to R_{\overline{K}}(H)$ is an isomorphism for every subgroup H of G. We do not assume anymore that $p := \mathrm{char}(k)$ is prime to $|G|$.

Theorem III.6.1. *The maps*

$$\mathrm{Ind} := \bigoplus_{\ell \text{prime}} \bigoplus_{H \in X(\ell)} \mathrm{Ind}_H^G : \bigoplus_{\ell \neq p \text{ prime}} \bigoplus_{H \in X(\ell)} \mathrm{R}_k(H) \to \mathrm{R}_k(G)$$

and

$$\mathrm{Ind} := \bigoplus_{\ell \text{prime}} \bigoplus_{H \in X(\ell)} \mathrm{Ind}_H^G : \bigoplus_{\ell \neq p \text{ prime}} \bigoplus_{H \in X(\ell)} \mathrm{P}_k(H) \to \mathrm{P}_k(G)$$

are both surjective.

(See section II.5 of chapter II for the definitions of all the terms.)

Proof. Let $\mathbb{1}_K$ (resp. $\mathbb{1}_k$) be the unit element in $\mathrm{R}_K(G)$ (resp. $\mathrm{R}_k(G)$). Obviously, $d(\mathbb{1}_K) = \mathbb{1}_k$.

By Brauer's theorem (theorem II.5.2 of chapter II, which applies thanks to theorem II.6.1 of the same chapter), we can write

$$\mathbb{1}_K = \sum_{\ell \text{ prime}} \sum_{H \in X(\ell)} \mathrm{Ind}_H^G x_H,$$

for some $x_H \in \mathrm{R}_K(H)$. Applying d to this equality gives

$$\mathbb{1}_k = \sum_{\ell \text{ prime}} \sum_{H \in X(\ell)} \mathrm{Ind}_H^G x'_H,$$

with $x'_H = d(x_H) \in \mathrm{R}_k(H)$. So if $y \in \mathrm{R}_k(G)$ (resp. $y \in \mathrm{P}_k(G)$), then corollary I.5.6.2 of chapter I gives

$$y = y\mathbb{1}_k = \sum_{\ell \text{ prime}} \sum_{H \in X(\ell)} \mathrm{Ind}_H^G(x'_H \,\mathrm{Res}_H^G y),$$

and $x'_H \,\mathrm{Res}_H^G(y)$ is in $\mathrm{P}_k(H)$ if $y \in \mathrm{P}_k(G)$, because Res_H^G sends $\mathrm{P}_k(G)$ to $\mathrm{P}_k(H)$.

□

III.7 Surjectivity of d

We keep the notation and assumptions of section III.4.

Corollary III.7.1. *If k is algebraically closed and K contains all the $|G|$th roots of 1 in $\overline{K}$, then $d : \mathrm{R}_K(G) \to \mathrm{R}_k(G)$ is surjective.*

Remark III.7.2. This result is actually true without the hypothesis on k and K, see section 16.1 of Serre's book [29].

Lemma III.7.3. *Suppose that $G = P \times H$, with P a p-group and H of order prime to p. Then P acts trivially on every semisimple $k[G]$-module.*

Proof. We just need to show that P acts trivially on every simple $k[G]$-module. Let M be a simple $k[G]$-module. As P is a p-group, its only irreducible representation over k is the trivial representation, by problem VII.1.11. so $M^P \neq 0$. (Choose a minimal nonzero $k[P]$-submodule of M, it has to be a simple $k[P]$-module, hence it is the trivial $k[P]$-module, and so it is included in M^P.) As $G = P \times H$ (so P is normal in G), the action of G preserves M^P. As M is simple, $M^P = M$, and so P acts trivially on M.

□

Proof of the corollary. By theorem III.6.1, we may assume that G is ℓ-elementary, for some prime ℓ. Write $G = C \times G'$, with C cyclic of order prime to ℓ and G' a ℓ-group. If $\ell = p$, let $P = G'$ and $H = C$. If $\ell \neq p$, write $C = C_p \times C^p$ with C_p a p-group and C^p of order prime to p, and let $P = C_p$ and $H = C^p \times G'$. In both cases, we have written $G = H \times P$, with P a p-group and H of order prime to p.

Let M be a simple $k[G]$-module. By the lemma, P acts trivially on M. So M is a simple $k[H]$-module. By theorem III.5.1, there exists a simple $K[H]$-module V such that $d([V]) = [M]$. We see V as a $K[G]$-module by making P act trivially, and then we still have $d([V]) = [M]$.

□

III.8 Injectivity of c

Let k be a field of characteristic $p > 0$, and G be a finite group.

Proposition III.8.1. *If G is a p-group, then $\mathrm{P}_k(G) \simeq \mathbb{Z}$, $\mathrm{R}_k(G) \simeq \mathbb{Z}$, and the map $\mathrm{P}_k(G) \to \mathrm{R}_k(G)$ corresponds to multiplication by $p^n := |G|$.*

Proof. We have seen in problem VII.1.11 that the only simple $k[G]$-module is $\mathbb{1}$, so the $\mathbb{Z}$-rank of $\mathrm{P}_k(G)$ and $\mathrm{R}_k(G)$ is 1, the isomorphism $\mathrm{R}_k(G) \xrightarrow{\sim} \mathbb{Z}$ sends $[M]$ to $\dim_k M$, and also $k[G]/\operatorname{rad}(k[G]) = k$. So $k[G]$ is local (and left Artinian), and every element of $k[G]$ is nilpotent or invertible by corollary III.1.2.5. In particular, the only idempotents of $k[G]$ are 0 and 1.

We know that $k[G]$ is a projective $k[G]$-module, let's show that it is indecomposable. Suppose that $k[G] = M_1 \oplus M_2$, with MI_1 and I_2 two left ideals of $k[G]$. By remark I.1.3.16 of chapter I, we get two idempotents e_1, e_2 if $k[G]$ such that $I_1 = k[G]e_1$ and $I_2 = k[G]e_2$. As the only idempotents of $k[G]$ are 0 and 1, this implies that $I_1 = 0$ or $I_2 = 0$.

So we have found the unique projective indecomposable finite length $k[G]$-module that surjects to $\mathbb{1}$, and it is $k[G]$ itself. Now the last assertion follows from the fact that, in $\mathrm{R}_k(G)$, $[k[G]] = \dim_k(k[G])\mathbb{1} = p^n\mathbb{1}$.

□

Theorem III.8.2. *Assume that k is algebraically closed.[5] Then $c : \mathrm{P}_k(G) \to \mathrm{R}_k(G)$ is injective, and its image contains $p^n\mathrm{R}_k(G)$, where p^n is the biggest power of p dividing $|G|$.*

Proof. In the proof, we will use a complete discrete valuation ring Λ with residue field k and algebraically closed characteristic zero fraction field K. The existence of such a Λ is almost proved in problem VII.3.4.

We first prove that $\operatorname{Im}(c) \supset p^n\mathrm{R}_k(G)$. By theorem III.6.1, we may assume that G is ℓ-elementary for some prime number ℓ. Then, as in the proof of corollary III.7.1, we can write $G = H \times P$ with P a p-group and H of order prime to p. The trivial $k[H]$-module is projective (because $k[H]$ is a semisimple ring), so $k[P]$ (with trivial action of H) is a projective $k[G]$-module, and $k[P] = |P|\mathbb{1} \in \mathrm{R}_k(G)$ (by proposition III.8.1) is in the image of c. As $\operatorname{Im}(c)$ is an ideal of $\mathrm{R}_k(G)$, this gives the conclusion.

Now let's prove that c is injective. We already know that $\mathrm{R}_k(G)/\operatorname{Im}(c)$ is a torsion group. As $\mathrm{P}_k(G)$ and $\mathrm{R}_k(G)$ are free $\mathbb{Z}$-modules of the same finite rank, this forces c to be injective.

□

Corollary III.8.3. *The map $e : \mathrm{P}_k(G) \to \mathrm{R}_K(G)$ of section III.4 is injective.*

[5] This is not necessary, see section 16.1 of Serre's book [29].

III.9 Image of the map e

We keep the notation and assumptions of section III.4, and we assume that k is algebraically closed and that K contains all the $|G|$th roots of 1 in $\overline{K}$.[6]

Definition III.9.1. An element $x \in G$ is called *p-singular* if x is not p-regular, i.e. if p divides the order of x.

Remember that we have an injective morphism of rings $\mathrm{R}_K(G) \to \mathscr{C}(G, K)$, (by corollary II.1.2.9 of chapter II), and use it to identify $\mathrm{R}_K(G)$ to a subring of $\mathscr{C}(G, K)$.

Theorem III.9.2. *An element $\chi \in \mathrm{R}_K(G)$ is in the image of e if and only if $\chi(g) = 0$ for every p-singular element g of G.*

Lemma III.9.3. *Suppose that $G = H \times P$, with P a p-group and H of order prime to p. Then :*

1. $k[G] = k[H] \otimes_k k[P]$.
2. $\mathrm{rad}(k[G]) = k \otimes_k \mathrm{rad}(k(P)) = k \otimes_k I_\varepsilon$, *where I_ε is the augmentation ideal of $k[P]$ (see I.3.1 of chapter I).*
3. *A $k[G]$-module $\overline{M}$ is projective if and only if $\overline{M} \simeq \overline{N} \otimes_k k[P]$, with $\overline{N}$ a $k[H]$-module.*
4. *A $\Lambda[G]$-module M is projective if and only if $M \simeq N \otimes \Lambda[P]$, with N a $\Lambda[H]$-module that is free (of finite type) over Λ.*

Proof. 1. Obvious.

2. Let $I = k \otimes_k I_\varepsilon$. Then $k[G]/I \simeq k[H]$ is semisimple, so $I \subset \mathrm{rad}(k[G])$. Also, I acts trivially on every simple $k[G]$-module by lemma III.7.3, so $I \subset \mathrm{rad}(k[G])$.

3. If $\overline{M} = \overline{N} \otimes_k k[P]$ with $\overline{N}$ a $k[H]$-module, then $\overline{N}$ is a projective $k[H]$-module because $k[H]$ is semisimple, so it is a direct summand of a free $k[H]$-module, so $\overline{M}$ is a direct summand of a free $k[G]$-module, hence projective.

   To prove the converse, we may assume (by theorem III.1.4.1) that $\overline{M}$ is a projective indecomposable $k[G]$-module. Then $\overline{N} := \overline{M}/\mathrm{rad}(k[G])\overline{M}$ is a simple $k[G]$-module, so P acts trivially on $\overline{N}$ by lemma III.7.3, so $\overline{N}$ is also a simple $k[H]$-module. Let $\overline{M}' = \overline{N} \otimes_k k[P]$, then $\overline{M}'$ is a projective $k[G]$-module by what we just saw, it is indecomposable because $\overline{M}'/\mathrm{rad}(k[G])\overline{M}' = \overline{N}$ is simple (use corollary III.1.5.2), and so it is isomorphic to $\overline{M}$ by proposition III.1.5.5.

4. If $M \simeq N \otimes_\Lambda \Lambda[P]$ as in the statement, then M is a free Λ-module and $M \otimes_\Lambda k$ is a projective $k[G]$-module by (iii), so M is a projective $\Lambda[G]$-module by theorem III.3.4.

   Conversely, let M be a projective $\Lambda[G]$-module. Then $\overline{M} := M \otimes_\Lambda k$ is a projective $k[G]$-module, so, by (iii), we have $\overline{M} \simeq \overline{N} \otimes_k k[P]$ with $\overline{N}$ a $k[H]$-module. By theorem III.3.4

[6] Again, this is not necessary, as explained in section 16.1 of Serre's book [29].

again, there exists a projective $\Lambda[H]$-module N such that $N \otimes_\Lambda k$. Then $M' := N \otimes_\Lambda \Lambda[P]$ is a projective $\Lambda[G]$-module by what we just saw, and $M' \otimes_\Lambda k \simeq M \otimes_\Lambda k$, so $M \simeq M'$ by theorem III.3.4.

□

Proof of the theorem. Let's prove that any element in the image of e satisfies the condition of the theorem. Let M be a projective $\Lambda[G]$-module, let $V = M \otimes_\Lambda K$. We want to show that $\chi_V(g) = 0$ if $g \in G$ is p-singular. Fix a p-singular $g \in G$. We may assume that $G = \langle g \rangle$, so G is cyclic, so we can write $G = H \times P$ with P a p-group and H of order prime to p. Then by the lemma, $M = N \otimes_\Lambda \Lambda[P]$, with N a $\Lambda[H]$-module that is free over Λ. If we write $g = (g_1, g_2)$ with $g_1 \in H$ and $g_2 \in P$, then we have $\chi_V(g) = \chi_{M \otimes_\Lambda K}(g_1)\chi_{K[P]}(g_2)$. As g is p-singular, $g_2 \neq 1$, so $\chi_{K[P]}(g_2) = 0$ and hence $\chi_V(g) = 0$.

Now let's prove that any element of $\mathrm{R}_K(G)$ satisfying the condition of the theorem is in the image of e. So let $\chi \in R_K(G)$ be such that $\chi(g) = 0$ for every p-singular $g \in G$. By Brauer's theorem (theorem II.5.2 of chapter II, which applies thanks to theorem II.6.1 of the same chapter), we can write $\mathbb{1} = \sum_H \mathrm{Ind}_H^G(\psi_H)$, where we take the sum over elementary (=ℓ-elementary for some prime ℓ) subgroups H of G and $\psi_H \in \mathrm{R}_K(H)$. Using corollary I.5.6.2 of chapter I, we get

$$\chi = \chi\mathbb{1} = \sum_H \mathrm{Ind}_H^G \mathrm{Ind}_H^G(\chi_H),$$

with $\chi_H = \psi_H \mathrm{Res}_H^G \chi$. Clearly, for every H and every $g \in H$, $chi_H(g) = 0$ if g is p-singular. It suffices to show that χ_H is in the image of $e : \mathrm{P}_k(H) \to \mathrm{R}_K(H)$ for every H. In other words, we may assume that G is elementary.

As in the proof of corollary III.7.1, write $G = H \times P$, with P a p-group and H of order prime to p. If $g_1 \in H$ and $g_2 \in P - \{1\}$, we have $\chi(g_1, g_2) = 0$ by assumption. So there is a function $f \in \mathscr{C}(H, K)$ such that $\chi(g_1, g_2) = f(g_1)\chi_{K[P]}(g_2)$ for every $(g_1, g_2) \in H \times P = G$. If W is a representation of H over K, then

$$\mathbb{Z} \ni \langle \chi, \chi_{W \otimes \mathbb{1}_P} \rangle_G = \langle f, \chi_W \rangle_H \langle \chi_{K[P]}, \chi_{\mathbb{1}_P} \rangle_P = \langle f, \chi_W \rangle_H,$$

so $\langle f, \chi_W \rangle_H \in \mathbb{Z}$, hence f is in $\mathrm{R}_K(H)$ and not just $\mathscr{C}(H, K)$. (Because $f = \sum_{W \in S_K(H)} \langle f, \chi_W \rangle_H \chi_{W^*}$ by corollaries II.1.2.5 and II.1.2.6 and theorem II.1.3.1 of chapter II.) Write $f = \sum_{W \in S_K(H)} n_W[W]$, with $n_W \in \mathbb{Z}$. By theorem III.5.1, for every $W \in S_K(H)$, there exists a projective $\Lambda[H]$-module M_W such that $W \simeq M_W \otimes_\Lambda K$. So if $x = \sum_{W \in S_K(H)} n_W[M_W \otimes_\Lambda \Lambda[P]] \in \mathrm{P}_\Lambda(G)$, we have $d(x) = \chi$.

□

IV Irreducible representations of the symmetric group $\mathfrak{S}_n$ over $\mathbb{C}$

The goal of this chapter is to explain the classical description of irreducible representations of $\mathfrak{S}_n$ in terms of partitions of n and to give a formula for their characters.

IV.1 Partitions

Fix a positive integer n.

Definition IV.1.1. A *partition* of n is a finite sequence $\lambda = (\lambda_1, \dots, \lambda_r)$ such that

- $\lambda_1 \geq \cdots \geq \lambda_r$;
- $\lambda_1 + \cdots + \lambda_r = n$.

We write $\mathscr{P}(n)$ for the set of partitions of n.

Definition IV.1.2. The *lexicographic order* on $\mathscr{P}(n)$ is the total order relation given by : $(\lambda_1, \dots, \lambda_r) > (\mu_1, \dots, \mu_s)$ if and only if there exists $i \leq \min(r, s)$ such that $\lambda_j = \mu_j$ for $1 \leq j < i$ and $\lambda_i > \mu_i$. *lexicographic order on partitions*

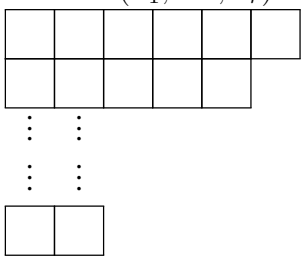
The following result is clear.

Proposition IV.1.3. *Using the decomposition into cycles with disjoint supports of elements of $\mathfrak{S}_n$, we get a bijection*

$$\begin{array}{ccc} \{\textit{conjugacy classes in } \mathfrak{S}_n\} & \xrightarrow{\sim} & \mathscr{P}(n) \\ \sigma & \longmapsto & \textit{sequence of the lengths of the cycles in the decomposition of } \sigma\textit{, ordered in decreasing order.} \end{array}$$

In particular, there is a bijection $\mathscr{P}(n) \simeq S_{\mathbb{C}}(\mathfrak{S}_n)$, where $S_{\mathbb{C}}(\mathfrak{S}_n)$ is the set of isomorphism classes of irreducible representations of $\mathfrak{S}_n$ over $\mathbb{C}$. We will see that there is actually a *canonical* bijection $\mathscr{P}(n) \simeq S_{\mathbb{C}}(\mathfrak{S}_n)$, so we get a canonical bijection between $S_{\mathbb{C}}(\mathfrak{S}_n)$ and the set of conjugacy classes in $\mathfrak{S}_n$; this is special to $\mathfrak{S}_n$ and is not the case for a general finite group.

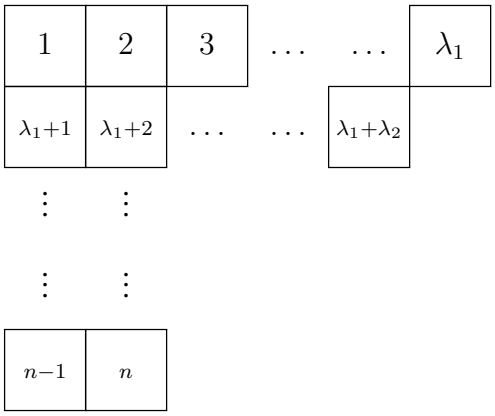
IV.2 Young tableaux and Young projectors

Definition IV.2.1. Let $\lambda = (\lambda_1, \ldots, \lambda_r) \in \mathscr{P}(n)$. The *Young diagram* Y_λ attached to λ is the following diagram  where there are r rows and the ith row has λ_i cases.

A *Young tableau* T_λ corresponding to λ is a filling of the cases of the Young diagram Y_λ with the number $1, \ldots, n$ without repetitions. (Note that Y_λ has exactly n cases.)

If T_λ is a Young tableau corresponding to λ, the *row subgroup* P_{T_λ} (resp. the *column subgroup* Q_{T_λ}) of $\mathfrak{S}_n$ is the subgroup of σ such that σ maps every element of $\{1 \ldots, n\}$ to an element in the same row (resp. in the same column) of T_λ.

Example IV.2.2. If T_λ is equal to



then $P_{T_\lambda} = \mathfrak{S}_\lambda$ is the subgroup of $\sigma \in \mathfrak{S}_n$ stabilizing the sets $\{1, \ldots, \lambda_1\}, \{\lambda_1 + 1, \ldots, \lambda_1 + \lambda_2\}, \ldots, \{\lambda_1 + \cdots + \lambda_{r-1} + 1, \ldots, \lambda_1 + \ldots, \lambda_r\}$.

Remark IV.2.3. We have $P_{T_\lambda} \cap Q_{T_\lambda} = \{1\}$.

Indeed, if $\sigma \in P_{T_\lambda} \cap Q_{T_\lambda}$, then for every $i \in \{1, \ldots, n\}$, $\sigma(i)$ is in the same row and in the same column of T_λ as i, which forces $\sigma(i) = i$.

Remark IV.2.4. Let $\lambda \in \mathscr{P}(n)$. If T_λ is a Young tableau corresponding to λ and $\sigma \in \mathfrak{S}_n$, then, applying σ to all the entries of T_λ, we get another Young tableau corresponding to λ. This gives an action of $\mathfrak{S}_n$ on Young tableaux corresponding to λ. If $T'_\lambda = \sigma T_\lambda$, then $P_{T'_\lambda} = \sigma P_{T_\lambda} \sigma^{-1}$ and $Q_{T'_\lambda} = \sigma Q_{T_\lambda} \sigma^{-1}$

Definition IV.2.5. Let $\lambda \in \mathscr{P}(n)$, and let T_λ be a Young tableau corresponding to λ. We define two elements $a_\lambda, b_\lambda \in \mathbb{Q}[\mathfrak{S}_n]$ by

$$a_\lambda = \frac{1}{|P_{T_\lambda}|} \sum_{\sigma \in P_{T_\lambda}} \sigma$$

and

$$b_\lambda = \frac{1}{|Q_{T_\lambda}|} \sum_{\sigma \in P_{T_\lambda}} \mathrm{sgn}(\sigma)\sigma.$$

The *Young projector* corresponding to T_λ is

$$c_\lambda = a_\lambda b_\lambda.$$

Note that these elements depend on T_λ, but we only indicate the dependence on λ.

Remark IV.2.6. An easy calculation shows that $a_\lambda^2 = a_\lambda$ and $b_\lambda^2 = b_\lambda$. Also, because $P_{T_\lambda} \cap Q_\lambda = \{1\}$ (see remark IV.2.3), we have no cancellations in the sum

$$c_\lambda = \sum_{\sigma \in P_{T_\lambda}, \tau \in Q_{T_\lambda}} \mathrm{sgn}(\tau)\sigma\tau,$$

so $c_\lambda \neq 0$.

Now we prove some basic properties of these elements.

Proposition IV.2.7. *Let $\lambda, \mu \in \mathscr{P}(n)$, and choose corresponding Young tableaux T_λ and T_μ.*

Then, for every $\sigma \in \mathfrak{S}_n$, we have

$$a_\lambda \sigma b_\mu = \begin{cases} 0 & \text{if } \lambda > \mu \\ 0 & \text{if } \lambda = \mu \text{ and } \sigma \notin P_{T_\lambda} Q_{T_\lambda} \\ \mathrm{sgn}(q) a_\lambda b_\lambda & \text{if } \lambda = \mu \text{ and } \sigma = pq, \text{ with } p \in P_{T_\lambda} \text{ and } q \in Q_{T_\lambda}. \end{cases}$$

Proof. We start with the following easy observation (that was already used implicitely in remark IV.2.6): if $p \in P_{T_\lambda}$ and $q \in Q_{T_\mu}$, then

$$a_\lambda p = \sum_{\sigma \in P_{T_\lambda}} (\sigma p) = a_\lambda$$

and

$$q b_\mu = \sum_{\sigma \in Q_{T_\mu}} \mathrm{sgn}(\sigma)(q\sigma) = \mathrm{sgn}(q) \sum_{\sigma \in Q_{T_\mu}} \mathrm{sgn}(q\sigma)(q\sigma) = \mathrm{sgn}(q) b_\mu.$$

In particular, if $s = pq$ with $p \in P_{T_\lambda}$ and $q \in Q_{T_\lambda}$, then

$$a_\lambda s b_\lambda = (a\lambda p)(q b_\lambda) = \mathrm{sgn}(q) a_\lambda b_\lambda.$$

Now suppose that we can prove that $P_{T_\lambda} \cap s Q_{T_\mu} s^{-1}$ contains a transposition τ. Then we have

$$a_\lambda s b_\mu = (a_\lambda \tau) s (s^{-1} \tau s b_\mu) = a_\lambda s \mathrm{sgn}(s^{-1}\tau s) b_\mu = -a_\lambda s b_\mu,$$

hence $a_\lambda s b_\mu = 0$. So we just need to prove that $P_{T_\lambda} \cap s Q_{T_\mu} s^{-1}$ contains a transposition if $\mu < \lambda$, or if $\lambda = \mu$ and $s \notin P_{T_\lambda} Q_{T_\lambda}$.

First note that, thanks to remark IV.2.4, we can find a Young tableau T'_μ corresponding to μ such that $Q_{T'_\mu} = s Q_{T_\mu} s^{-1}$. Also, if we make an element of $Q_{T'\mu}$ act on the Young tableau T'_μ, this won't change its column group. Now let's suppose that $P_{T_\lambda} \cap s Q_{T_\mu} s^{-1} = P_{T_\lambda} \cap Q_{T'_\mu}$ doesn't contain a transposition. Write a_{ij} for the entries of T_λ, where a_{ij} means the entry on the ith row and jth column. Let σ be the element of $\mathfrak{S}_n$ such that $\sigma(T_\lambda) = T'_\mu$. Then, for $2 \leq j \leq \lambda_1$, $\sigma(a_{1j})$ is not in the same column as $\sigma(a_{11})$ (otherwise the transposition $\tau = (a_{11}, a_{1j})$ would be in $P_{T_\lambda} \cap Q_{T'_\mu}$). After making an element of $Q_{T'_\mu}$ act on T'_μ (which doesn't change the problem, as we saw), we can also assume that all the $\sigma(a_{1j})$, $1 \leq j \leq \lambda_1$, are in the first row of T'_μ. And so in particular, $\lambda_1 \leq \mu_1$.

Next, for $2 \leq j \leq \lambda_2$, $\sigma(a_{2j})$ is not in the same column as $\sigma(a_{21})$ (otherwise the transposition $\tau = (a_{21}, a_{2j})$ would be in $P_{T_\lambda} \cap Q_{T'_\mu}$). After making an element of $Q_{T'_\mu}$ act on T'_μ, we can assume that all the $\sigma(a_{2j})$, $1 \leq j \leq \lambda_2$, are in the first two rows of T'_μ.

Applying the same reasoning to all the rows of T_λ, we conclude that, after making an element of $Q_{T'_\mu}$ act on T'_μ, the images by σ of the entries in the ith row of T_λ are all in the first i rows of T'_μ, for every i. In particular, $\lambda_1 + \cdots + \lambda_i \leq \mu_1 + \cdots + \mu_i$ for every i.

Suppose that $\lambda = \mu$. Then using the reasoning above, we can actually make it so that the images by σ of the entries in the ith row of T_λ are all in the ith row of T'_μ, for every i. In other words, $\sigma \in P_{T_\lambda}$.

This already proves that $P_{T_\lambda} \cap Q_{T'_\mu}$ contains a transposition if $\lambda > \mu$. Suppose that $P_{T_\lambda} \cap s Q_{T_\lambda} s^{-1} = P_{T_\lambda} \cap Q_{sT_\lambda}$ doesn't contain a transposition. By what we saw above, $s Q_{T_\lambda} \cap P_{T_\lambda} \neq \varnothing$, i.e. $s \in P_{T_\lambda} Q_{T_\lambda}$.

□

Corollary IV.2.8. *Let $\lambda, \mu \in \mathfrak{S}_n$, and choose corresponding Young tableaux T_λ and T_μ.*

1. *We have $a_\lambda \mathbb{C}[\mathfrak{S}_n] b_\mu = 0$ if $\mu < \lambda$.*
2. *Let $\ell : \mathbb{C}[\mathfrak{S}_n] \to \mathbb{C}$ be the $\mathbb{C}$-linear function defined by*

$$\ell(\sigma) = \begin{cases} 0 & \text{if } \sigma \notin P_{T_\lambda} Q_{T_\lambda} \\ \operatorname{sgn}(q) & \text{if } \sigma = pq, \text{ with } p \in P_{T_\lambda} \text{ and } q \in Q_{T_\lambda}. \end{cases}$$

   Then, for every $x \in \mathbb{C}[\mathfrak{S}_n]$, we have

$$a_\lambda x b_\lambda = \ell(x) a_\lambda b_\lambda = \ell(x) c_\lambda.$$

3. *We have*

$$c_\lambda^2 = \frac{n!}{\dim_{\mathbb{C}}(V_\lambda)} c_\lambda,$$

   where $V_\lambda = \mathbb{C}[\mathfrak{S}_n] c_\lambda$.

Proof. Point (i) and (ii) are obvious consequences of the proposition. Let's prove (iii). By (ii), we have

$$c_\lambda^2 = a_\lambda(b_\lambda a_\lambda)b_\lambda = \ell(b_\lambda a_\lambda)c_\lambda.$$

Let $\alpha = c(b_\lambda a_\lambda) \in \mathbb{C}$, and let u be the $\mathbb{C}$-linear endomorphism of $\mathbb{C}[\mathfrak{S}_n]$ given by right multiplication by c_λ. We have $u^2 = \alpha u$ by the calculation above, so the eigenvalues of u are all in $\{0, \alpha\}$, hence

$$\mathrm{Tr}(u) = \alpha \mathrm{rk}(u) = \alpha \dim_{\mathbb{C}}(V_\lambda).$$

On the other hand, we have, by remark IV.2.6,

$$c_\lambda = 1 + \sum_{\sigma \in P_\lambda Q_\lambda - \{1\}} \pm\sigma,$$

so, using the basis of $\mathbb{C}[\mathfrak{S}_n]$ given by $\mathfrak{S}_n$, we see that $\mathrm{Tr}(c_\lambda) = n!$, hence $\alpha = \frac{n!}{\dim_{\mathbb{C}} V_\lambda}$.

□

IV.3 Partitions and irreducible representations

Definition IV.3.1. Let $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathscr{P}(n)$. Let $\mathfrak{S}_\lambda \subset \mathfrak{S}_n$ be the subgroup of elements σ stabilizing the sets $\{1, \ldots, \lambda_1\}, \{\lambda_1+1, \ldots, \lambda_1+\lambda_2\}, \ldots, \{\lambda_1+\cdots+\lambda_{r-1}+1, \ldots, \lambda_1+\ldots, \lambda_r\}$ (as in example IV.2.2).

We also set

$$U_\lambda = \mathrm{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \mathbb{1}_{\mathfrak{S}_\lambda} = \mathbb{C}[\mathfrak{S}_n] \otimes_{\mathbb{C}[\mathfrak{S}_\lambda]} \mathbb{C} = \mathbb{C}[\mathfrak{S}_n \,/\, \mathfrak{S}_\lambda],$$

where $\mathfrak{S}_n$ acts on the last $\mathbb{C}$-vector space through its action by left translations on $\mathfrak{S}_n \,/\, \mathfrak{S}_\lambda$.

Proposition IV.3.2. *Choose a Young tableau T_λ corresponding to λ. Then we have $U_\lambda \simeq \mathbb{C}[\mathfrak{S}_n]a_\lambda$ as $\mathbb{C}[\mathfrak{S}_n]$-modules.*

Proof. By remark IV.2.4 and example IV.2.2, we may assume that $P_{T_\lambda} = \mathfrak{S}_\lambda$. Then, if $\sigma, \sigma' \in \mathfrak{S}_n$, we have $\sigma a_\lambda = \sigma' a_\lambda$ if and only $\sigma\,\mathfrak{S}_\lambda = \sigma'\,\mathfrak{S}_\lambda$. Let $(\sigma_i)_{i\in I}$ be a system of representatives of $\mathfrak{S}_n \,/\, \mathfrak{S}_\lambda$. Then the $\sigma_i a_\lambda$ have support in pairwise disjoint subsets of $\mathfrak{S}_n$, so they are linearly independent over $\mathbb{C}$. (Where the support of an element $x = \sum_{\sigma\in\mathfrak{S}_n} \alpha_\sigma \sigma \in \mathbb{C}[\mathfrak{S}_n]$ is the set of $\sigma \in \mathfrak{S}_n$ such that $\alpha_\sigma \neq 0$.) So the $(\sigma_i a_\lambda)_{i\in I}$ form a $\mathbb{C}$-basis of $\mathbb{C}[\mathfrak{S}_n]a_\lambda$.

In particular, we can define a $\mathbb{C}[\mathfrak{S}_n]$-linear map $u : U_\lambda = \mathbb{C}[\mathfrak{S}_n \,/\, \mathfrak{S}_\lambda] \to \mathbb{C}[\mathfrak{S}_n]a_\lambda$ by sending $\sigma\,\mathfrak{S}_\lambda$ to σa_λ, for every $\sigma \in \mathfrak{S}_n$. This sends the basis $(\sigma_i\,\mathfrak{S}_\lambda)_{i\in I}$ of $\mathbb{C}[\mathfrak{S}_n \,/\, \mathfrak{S}_\lambda]$ to the basis $(\sigma_i a_\lambda)_{i\in I}$ of $\mathbb{C}[\mathfrak{S}_n]a_\lambda$ that we just defined, and so it's an isomorphism.

□

Theorem IV.3.3. *For every $\lambda \in \mathscr{P}(n)$, we choose a Young tableau T_λ corresponding to λ, and we set, as in corollary IV.2.8, $V_\lambda = \mathbb{C}[\mathfrak{S}_n]c_\lambda$.*

Then V_λ is an irreducible representation of $\mathfrak{S}_n$, and we have

$$U_\lambda = V_\lambda \oplus \bigoplus_{\mu > \lambda} V_\mu^{K_{\mu\lambda}},$$

for some $K_{\mu\lambda} \in \mathbb{N}$. (This integers $K_{\mu\lambda}$ are called the Kostka numbers.*)*

Moreover, the isomorphism class of the representation V_λ only depends on λ (and not on the choice of T_λ), and the map $\mathscr{P}(n) \to S_{\mathbb{C}}(\mathfrak{S}_n)$, $\lambda \longmapsto V_\lambda$, is a bijection.

The representations V_λ are called the *Specht modules*.

Example IV.3.4.

- Take $\lambda = (n)$ (the biggest element of $\mathscr{P}(n)$). Then $\mathfrak{S}_\lambda = \mathfrak{S}_n$ and $Q_{T_\lambda} = \{1\}$ for every choice of T_λ, so
$$a_\lambda = c_\lambda = \sum_{\sigma \in \mathfrak{S}_n} \sigma$$
and $U_\lambda = V_\lambda = \mathbb{1}_{\mathfrak{S}_n}$.
- Take $\lambda = (n-1, 1)$. Then $\mathfrak{S}_\lambda = \mathfrak{S}_{n-1} \times \mathfrak{S}_1 \subset \mathfrak{S}_n$, and U_λ is the representation of $\mathfrak{S}_n$ on $\mathbb{C}^n$ that permutes the coordinates. The only element of $\mathscr{P}(n)$ bigger than λ is (n), so $U_\lambda = V_\lambda \oplus \mathbb{1}$, and V_λ is isomorphic to the subrepresentation of $U_\lambda = \mathbb{C}^n$ equal to $\{(x_1, \dots, x_n) \in \mathbb{C}^n | x_1 + \cdots + x_n = 0\}$.
- Take $\lambda = (1, \dots, 1)$ (the smallest element of $\mathscr{P}(n)$). Take $\mathfrak{S}_\lambda = \{1\}$ and $Q_{T_\lambda} = \mathfrak{S}_n$ for every choice of T_λ. So U_λ is equal to the regular representation $\mathbb{C}[\mathfrak{S}_n]$. We know that
$$\mathbb{C}[\mathfrak{S}_n] \simeq \bigoplus_{V \in S_{\mathbb{C}}(\mathfrak{S}_n)} V^{\oplus \dim_{\mathbb{C}} V}$$
as $\mathbb{C}[\mathfrak{S}_n]$-modules. The only irreducible representations that appear with multiplicity 1 are the 1-dimensional representation, that is, $\mathbb{1}$ and sgn. As $\mathbb{1} = V_{(n)}$, we must have $\mathrm{sgn} = V_{(1,\dots,1)}$.

Lemma IV.3.5. *Let R be a ring, and let $e, f \in R$ be two idempotents. (That is, $e^2 = e$ and $f^2 = f$.) Then the map $eRf \to \mathrm{Hom}_R(Re, Rf)$ sending $x \in eRf$ to the R-linear map $f_x : a \longmapsto ax$ is an isomorphism of groups.*

Proof. Let's prove that this map is injective. Let $x, y \in eRf$ such that $f_x = f_y$. Then $0 = f_x(e) - f_y(e) = e(x - y)$. As $x - y \in eRf$ and $e^2 = e$, we have $e(x - y) = x - y$, and so $x - y = 0$.

Now we prove that the map $x \longmapsto f_x$ is surjective. Let $f \in \mathrm{Hom}_R(Re, Rf)$, let $x = f(e)$. Then $x = f(e^2) = ef(e) \in eRf$. Also, for every $a \in Re$, $f(a) = f(ae) = af(e) = ax$. Hence $f = f_x$.

□

Proof of the theorem. Let $\lambda, \mu \in \mathscr{P}(n)$. Then, by the lemma and (iii) of corollary IV.2.8 (that says that βc_λ is idempotent for some $\beta \in \mathbb{C}^\times$), we have

$$\mathrm{Hom}_{\mathbb{C}[\mathfrak{S}_n]}(V_\lambda, V_\mu) = c_\lambda \mathbb{C}[\mathfrak{S}_n] c_\mu.$$

By (i) and (ii) of corollary IV.2.8, this is equal to 0 if $\lambda > \mu$, and, if $\lambda = \mu$, it is equal to $\ell(b_\lambda \mathbb{C}[\mathfrak{S}_n] a_\lambda) c_\lambda \subset \mathbb{C} c_\lambda$. Note also that $c_\lambda^2 = \frac{n!}{\dim_{\mathbb{C}} V_\lambda} c_\lambda \neq 0$ is in $c_\lambda \mathbb{C}[\mathfrak{S}_n] c_\lambda$, so $c_\lambda \mathbb{C}[\mathfrak{S}_n] c_\lambda = \ell(b_\lambda \mathbb{C}[\mathfrak{S}_n] a_\lambda) c_\lambda$ is not equal to $\{0\}$, and so it is equal to $\mathbb{C} c_\lambda$.

In particular, we have shown that $\dim_{\mathbb{C}} \mathrm{End}_{\mathbb{C}[\mathfrak{S}_n]}(V_\lambda)$, and this implies that V_λ si irreducible. Also, the ffirst part of the calculation above implies that $V_\lambda \not\simeq V_\mu$ if $\lambda > \mu$ or $\lambda < \mu$. As the lexicographic order is a total order, this means that $V_\lambda \not\simeq V_\mu$ is $\lambda \neq \mu$.

So we see that the map $\mathscr{P}(n) \to S_{\mathbb{C}}(\mathfrak{S}_n)$, $\lambda \longmapsto V_\lambda$, is an injection. As its source and target have the same cardinality, this map is bijective, and every irreducible representation of $\mathfrak{S}_n$ is isomorphic to one of the V_λ.

Now let's prove the decomposition of U_λ given in the theorem. By what we just saw (and the semsimplicity of $\mathbb{C}[\mathfrak{S}_n]$), we have $U_\lambda = \bigoplus_{\mu \in \mathscr{P}(n)} V_\mu^{\oplus K_{\mu\lambda}}$, for some $K_{\mu\lambda} \in \mathbb{N}$. Using the lemma and corollary IV.2.8 again, we get

$$\mathrm{Hom}_{\mathbb{C}[\mathfrak{S}_n]}(U_\lambda, V_\mu) = a_\lambda \mathbb{C}[\mathfrak{S}_n] c_\mu = a_\mu(\mathbb{C}[\mathfrak{S}_n] a_\lambda) b_\lambda = \begin{cases} 0 & \text{if } \lambda > \mu \\ \mathbb{C} c_\lambda & \text{if } \lambda = \mu. \end{cases}$$

(If $\lambda = \mu$, we have $a_\lambda \mathbb{C}[\mathfrak{S}_n] c_\lambda \neq 0$ because it contains $a_\lambda c_\lambda = c_\lambda \neq 0$.) So $K_{\lambda\lambda} = 1$ and $K_{\mu\lambda} = 0$ if $\mu < \lambda$.

It just remains to show that the isomorphism class of V_λ doesn't depend on the choice of the Young tableau T_λ. Thanks to the decomposition of U_λ that we just proved, we can prove this by descending induction on $\lambda \in \mathscr{P}(n)$. Also, the case of the biggest element (n) of $\mathscr{P}(n)$ is obvious (see example IV.3.4). So we are done. [1]

□

Corollary IV.3.6. *Every irreducible representation of $\mathfrak{S}_n$ is realizable over $\mathbb{Q}$. (See corollary II.6.2 of chapter II.)*

[1] We could also use remark IV.2.4, which shows that changing the Young tableau conjugates c_λ by an element of $\mathfrak{S}_n$.

IV.4 Characters of the irreducible representations V_λ

The strategy to calculate the character of V_λ is similar to the strategy that we will use to calculate characters of irreducible representations of $\mathfrak{sl}_n(\mathbb{C})$ in section VI.14.4-VI.14.6 of chapter VI : First we calculate the chatacter of the induced representation U_λ, which is much easier. Then we deduce tha character of V_λ, using the fact that U_λ is the direct sum of V_λ and some factors V_μ with $\mu > \lambda$, the fact that V_λ is irreducible and some dark magic.

Definition IV.4.1. If $\sigma \in \mathfrak{S}_n$, we write $C(\sigma)$ for the conjugacy class of σ in $\mathfrak{S}_n$ and let $Z_{\mathfrak{S}_n}(\sigma) = \{\tau \in \mathfrak{S}_n \,|\tau\sigma = \sigma\tau\}$ be the centralizer of σ of $\mathfrak{S}_n$.

Proposition IV.4.2. *Let $\sigma \in \mathfrak{S}_n$. Then*

$$|Z_{\mathfrak{S}_n}(\sigma)| = \prod_{r\geq 1} c_r! r^{c_r}$$

and

$$|C(\sigma)| = \frac{n!}{\prod_{r\geq 1} c_r! r^{c_r}},$$

where, for every $r \geq 1$, c_r is the number of cycles of length r in the decomposition of σ into a product of cycles with disjoint supports.

Proof. Let $\tau \in Z_{\mathfrak{S}_n}(\sigma)$. Then τ has to send the support of each cycle of σ to the support of any other cycle of the same length, and it must also respect the cyclical order given by σ on the support of these cycles. This gives an isomorphism

$$Z_{\mathfrak{S}_n}(\sigma) \simeq \prod_{r\geq 1} ((\mathbb{Z}/r\mathbb{Z})^{c_r} \rtimes \mathfrak{S}_r)$$

(where $\mathfrak{S}_r$ acts on $(\mathbb{Z}/r\mathbb{Z})^{c_r}$ by permuting the entries of the r-uples), hence

$$|Z_{\mathfrak{S}_n}(\sigma)| = \prod_{r\geq 1} c_r! r^{c_r}.$$

Now note that $C(\sigma) = \mathfrak{S}_n / Z_{\mathfrak{S}_n}(\sigma)$. So we get

$$C(\sigma) = \frac{n!}{\prod_{r\geq 1} c_r! r^{c_r}}.$$

□

We fix some $N \geq n$. For every $r \geq 0$, let $P_r(T) = P_r(T_1, \ldots, T_N) = T_1^r + \cdots + T_N^r \in \mathbb{Z}[T_1, \ldots, T_N]$.

Theorem IV.4.3. *Let $\lambda = (\lambda_1, \dots, \lambda_d) \in \mathscr{P}(n)$ and $\sigma \in \mathfrak{S}_n$. Set $\lambda_i = 0$ for $d+1 \leq i \leq N$.*

Then $\chi_{U_\lambda}(\sigma)$ is the coefficient of $T^\lambda := \prod_{i=1}^N T_i^{\lambda_i}$ in the polynomial $\prod_{r\geq 1} P_r(T)^{c_r}$, where, for every $r \geq 1$, c_r is the number of cycles of length r in the decomposition of σ as a product of cycles with disjoint supports.

Proof. Remember that $U_\lambda = \mathrm{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \mathbb{1}$. We use the formula for the character of an induced representation (theorem II.3.1.2 of chapter II). It gives :

$$\chi_{U_\lambda}(\sigma) = \frac{1}{|\,\mathfrak{S}_\lambda\,|} \sum_{\tau\in\mathfrak{S}_n\,|\,\tau^{-1}\sigma\tau\in\mathfrak{S}_\lambda} 1 = \frac{1}{|\,\mathfrak{S}_\lambda\,|} |Z_{\mathfrak{S}_n}(\sigma)|\,|\,\mathfrak{S}_\lambda \cap C(\sigma)|,$$

where $Z_{\mathfrak{S}_n}(\sigma)$ and $C(\sigma)$ are as in definition IV.4.1.

First, we have $\mathfrak{S}_\lambda \simeq \mathfrak{S}_{\lambda_1} \times \cdots \times \mathfrak{S}_{\lambda_d}$, so $|\,\mathfrak{S}_\lambda\,| = \prod_{i=1}^d \lambda_i! = \prod_{i=1}^N \lambda_i!$. Second, by proposition IV.4.2,

$$|Z_{\mathfrak{S}_n}(\sigma)| = \prod_{r\geq 1} c_r! r^{c_r}.$$

Finally, we have to calculate $|\,\mathfrak{S}_\lambda \cap C(\sigma)|$. The conjugacy class $C(\sigma)$ is the set of permutations in $\mathfrak{S}_n$ that have c_r cycles of length r for every $r \geq 1$. So its intersection with $\mathfrak{S}_\lambda$ is a finite disjoint union of the following conjugacy classes in $\mathfrak{S}_\lambda \simeq \mathfrak{S}_{\lambda_1} \times \dots \mathfrak{S}_{\lambda_d}$: The product for $i = 1, \dots, d$ of the conjugacy class in $\mathfrak{S}_{\lambda_i}$ of permutations with $c_{i,r}$ cycles of length r for every $r \geq 1$, for every family $(c_{i,r})_{1\leq d\leq r, r\geq 1}$ such that, for every $r \geq 1$, $c_r = \sum_{i=1}^d c_{i,r}$ and for every $i \in \{1, \dots, d\}$, $\lambda_i = \sum_{r\geq 1} r c_{i,r}$. The cardinality of this product of conjugacy classes is

$$\prod_{i=1}^d \frac{\lambda_i!}{\prod_{r\geq 1} c_{i,r}! r^{c_{i,r}}},$$

by proposition IV.4.2. We can actually take i in $\{1, \dots, N\}$ without changing the result, because $\lambda_i = 0$ for $i > d$.

Putting all this together, we get

$$\chi_{U_\lambda}(\sigma) = \frac{1}{\prod_{i=1}^N \lambda_i!} \prod_{r\geq 1} c_r! r^{c_r} \sum_{(c_{r,i})} \prod_{i=1}^N \frac{\lambda_i!}{\prod_{r\geq 1} c_{i,r}! r^{c_{r,i}}},$$

where the sum is over families $(c_{i,r})$ as above. This is equal to

$$\sum_{(c_{r,i})} \prod_{i=1}^N \frac{c_r!}{\prod_{r\geq 1} c_{r,i}}.$$

On the other, for every $r \geq 1$, we have

$$P_r(T)^{c_r} = \left(\sum_{i=1}^N T_i^r\right)^{c_r} = \sum_{c_r = c_{1,r} + \dots + c_{N,r}} \frac{c_r!}{\prod_{i=1}^N c_{i,r}!} \prod_{i=1}^N T_i^{r c_{i,r}}.$$

So the coefficient of $\prod_{i=1}^N T_i^{\lambda_i}$ in $\prod_{r\geq 1} P_r(T)^{c_r}$ is indeed equal to $\sum_{(c_{r,i})} \prod_{i=1}^N \frac{c_r!}{\prod_{r\geq 1} c_{r,i}}$, where the sum is over families $(c_{i,r})$ as above.

□

Let $\Delta(T) = \Delta(T_1, \ldots, T_N) = \prod_{1\leq i<j\leq N}(T_i - T_j)$. This is also equal to the Vandermonde determinant

$$\det \begin{pmatrix} T_1^{N-1} & T_2^{N-1} & \ldots & T_N^{N-1} \\ \vdots & \vdots & \vdots & \\ T_1 & T_2 & \ldots & T_N \\ 1 & 1 & \ldots & 1 \end{pmatrix}.$$

Theorem IV.4.4. *Let $\lambda = (\lambda_1, \ldots, \lambda_d) \in \mathscr{P}(n)$ and $\sigma \in \mathfrak{S}_n$. Set $\lambda_i = 0$ for $d+1 \leq i \leq N$.*

Then $\chi_{V_\lambda}(\sigma)$ is the coefficient of $\prod_{i=1}^N T_i^{\lambda_i+N-i}$ in the polynomial $\Delta(T)\prod_{r\geq 1} P_r(T)^{c_r}$, where, for every $r \geq 1$, c_r is the number of cycles of length r in the decomposition of σ as a product of cycles with disjoint supports.

Lemma IV.4.5. *Let $\lambda = (\lambda_1, \ldots, \lambda_N) \in \mathbb{Z}^N$ be such that $\lambda_1 \geq \cdots \geq \lambda_N$. Let $\tau \in \mathfrak{S}_N$, and let μ be the N-uple of integers $(\lambda_1+\tau(1)-1, \ldots, \lambda_N+\tau(N)-N)$, rearranged to be in non-increasing order. Then $\mu \geq \lambda$, and we have $\mu = \lambda$ if and only if $\tau = 1$.*

Proof. Let $i_0 \in \{0, \ldots, N\}$ be an integer such that $\lambda_i = \mu_i$ for every $1 \leq i \leq i_0$. Let's show that $\tau(i) = i$ for every $1 \leq i \leq i_0$ and that, if $i_0 \leq N-1$, then $\mu_{i_0+1} \geq \lambda_{i_0+1}$. This clearly implies the lemma (applying the result to the biggest i_0 with the above property.)

We reason by induction on i_0. If $i_0 = 0$, then the first statement is obvious, and the second statement is true because μ_1 is the biggest of all the $\lambda_i + \tau(i) - i$, so $1 \leq i \leq N$, so $\mu_1 \geq \lambda_1 + \tau(1) - 1 \geq \lambda_1$. Suppose that $i_0 \geq 1$ and that we know the result for $i_0 - 1$. First we have to prove that $\tau(i_0) = i_0$. We have $\mu_{i_0} \geq \lambda_i+\tau(i)-i$ for $i_0 \leq i \leq N$, so $\mu_{i_0} \geq \lambda_{i_0}+\tau(i_0)-i_0$. As $\mu_{i_0} = \lambda_{i_0}$, this gives $\tau(i_0) \leq i_0$. But $\tau(i_0) \in \{i_0, \ldots, N\}$ (because $\tau(i) = i$ for $1 \leq i < i_0$, so $\tau(i_0) = i_0$. Next, if $i_0 \leq N-1$, then

$$\mu_{i_0+1} = \sup_{i_0+1\leq i\leq N} (\lambda_i + \tau(i) - i) \geq \lambda_{i_0+1} + \tau(i_0+1) - (i_0+1) \geq \lambda_{i_0+1},$$

because $\tau(i_0+1) \in \{i_0+1, \ldots, N\}$ as $\tau_{|\{1,\ldots,i_0\}} = \mathrm{id}$.

□

Lemma IV.4.6 (Cauchy determinant)**.** *Consider the $N \times N$ matrix A_N with coefficients in $\mathbb{Q}(x_1, \ldots, x_N, Y_1, \ldots, y_N)$ given by $A_N = (\frac{1}{x_i - y_j})_{1\leq i,j\leq N}$.*

Then we have

$$\det(A_N) = \frac{\prod_{1\leq i<j\leq n}(x_i - x_j)(y_j - y_i)}{\prod_{i,j=1}^n (x_i - y_j)}.$$

The matrix A_N is called the *Cauchy matrix*, and its determinant is called the *Cauchy determinant*.

Proof. We prove the result by induction on N. It's obvious for $N = 1$, so suppose that $N \geq 2$ and that we know the result for $N - 1$.

We have $\det(A_N) = \sum_{\tau\in\mathfrak{S}_n} \operatorname{sgn}(\tau) \prod_{i=1}^N (x_i - y_{\tau(i)})^{-1}$, so $f_N(x_i, y_j) := \det(A) \prod_{i,j=1}^N (x_i - y_j)$ is a homogeneous polynomial of degree $N(N-1)$. Also, if we have $x_r = x_s$ (resp. $y_r = y_s$) for two distinct $r, s \in \{1, \ldots, N\}$, then A has two equal rows (resp. columns), so $\det(A) = 0$. Hence $f_N(x_i, y_j) = c \prod_{1\leq i<j\leq N} (x_j - x_i)(y_j - y_i)$, for some $c \in \mathbb{Q}^\times$. Now multiply the first column of A_N by $x_1 - y_1$ and set $x_1 = y_1$. We get a matrix B whose determinant is equal to $(x_1 - y_1) \det(A_N)\big|_{x_1=y_1}$ and also to $f_{N-1}(x_2, \ldots, x_N, y_2, \ldots, y_N)$. The first expression for this determinant is equal to

$$c \frac{\left(\prod_{j=2}^N (x_j - x_1)(y_j - x_1)\right)\left(\prod_{2\leq i<j\leq N} (x_j - x_i)(y_j - y_i)\right)}{\prod_{i=2}^N (x_i - x_1) \prod_{j=2}^N (x_1 - y_j) \prod_{i,j=2}^N (x_i - y_j)}$$

By the induction hypothesis, this is equal to

$$c(-1)^{N-1}(-1)^{(N-1)/(N-2)/2} f_N(x_2, \ldots, x_N, y_2, \ldots, y_N) = c(-1)^{N(N-1)/2} f_N(x_2, \ldots, x_N, y_2, \ldots, y_N),$$

and so $c = (-1)^{N(N-1)/2}$, which finishes the proof.

□

Proof of the theorem. Write $\chi_\lambda = \chi_{V_\lambda}$, and let χ'_λ be the function on conjugacy classes in $\mathfrak{S}_n$ defined in the statement of the theorem. First we want to show that

$$\chi'_\lambda = \chi_\lambda + \sum_{\mu>\lambda} b_{\lambda\mu} \chi_\mu,$$

for some integers $b_{\lambda\mu} \in \mathbb{Z}$.

Let σ and the c_r be as in the statement of the theorem. By definition, $\chi'_\lambda(\sigma)$ is the coefficient of $\prod_{i=1}^N T_i^{\lambda_i+N-i}$ in $\Delta(T) \prod_{r\geq 1} P_r(T)^{c_r}$. As $\Delta(T)$ is equal to the Vandermonde determinant written above, we have

$$\Delta(T) = \sum_{\tau\in\mathfrak{S}_N} \operatorname{sgn}(\tau) \prod_{i=1}^N T_i^{N-\tau(i)}.$$

So $\chi'_\lambda(\sigma)$ is equal to the sum over $\tau \in \mathfrak{S}_N$ of $\operatorname{sgn}(\tau)$ times the coefficient of $\prod_{i=1}^N T_i^{\lambda_i+N-i}$ in $\prod_{i=1}^N T_i^{N-\tau(i)} \prod_{r\geq 1} P_r(T)^{c_r}$, i.e. of $\operatorname{sgn}(\tau)$ times the coefficient of $\prod_{i=1}^N T_i^{\lambda_i-i+\tau(i)}$ in $\prod_{r\geq 1} P_r(T)^{c_r}$.

Let $\tau \in \mathfrak{S}_N$, and denote by $\mu_\tau = (\mu_{\tau,1}, \ldots, \mu_{\tau,N})$ the N-uple of integers $(\lambda_1 + \tau(1) - 1, \cdots + \lambda_N + \tau(N) - N)$, rearranged to be in non-increasing order. Observe

that the polynomial $\prod_{r\geq 1} P_r(T)^{c_r}$ is symmetric in the variables T_i, because all the $P_r(T)$ are. So the coefficient of $\prod_{i=1}^N T_i^{\lambda_i - i + \tau(i)}$ in $\prod_{r\geq 1} P_r(T)^{c_r}$ is equal to the coefficient of $\prod_{i=1}^N T_i^{\mu_{\tau,i}}$. Also, if one of $\mu_{\tau,i}$ is negtaive, then this coefficient is 0, because there are no negative powers of the T_i in $\prod_{r\geq 1} P_r(T)^{c_r}$. Note that saying that none of the $\mu_{\tau,i}$ is negative is the same as saying that μ_τ is a partition of n (because of course $\sum_{i=1}^N \mu_{\tau,i} = \sum_{i=1}^N (\lambda_i + \tau(i) - i) = \sum_{i=1}^N \lambda_i$).

So we get that $\chi'_\lambda(\sigma)$ is equal to the sum over all $\tau \in \mathfrak{S}_N$ such that μ_τ is a partition of n of $\mathrm{sgn}(\tau)$ times the coefficient of $\prod_{i=1}^N T_i^{\mu_{\tau,i}}$ in $\prod_{r\geq 1} P_r(T)^{c_r}$. By theorem IV.4.3, $\chi'_\lambda(\sigma)$ is equal to the sum over all $\tau \in \mathfrak{S}_N$ such that μ_τ is a partition of n of $\chi_{U_\mu}(\sigma)$. Note also that, by lemma IV.4.5, for every $\mu \in \mathfrak{S}_N$, we have $\mu_\tau \geq \lambda$, and that $\mu_\tau = \lambda$ if and only if $\tau = 1$. Hence $\chi'_\lambda = \chi_{U_\lambda} + \sum_{\mu>\lambda} a_{\lambda\mu}\chi_{U_\mu}$, for some integers $a_{\lambda\mu} \in \mathbb{Z}$. Using the decomposition $U_\mu = V_\mu \oplus \bigoplus_{\nu>\mu} V_\nu^{\oplus K_{\nu\mu}}$ of theorem IV.3.3, we get that

$$\chi'_\lambda = \chi_\lambda + \sum_{\mu>\lambda} b_{\lambda\mu}\chi_\mu,$$

for some integers $b_{\lambda\mu} \in \mathbb{Z}$.

Remember the Hermitian inner product $\cdot$ on $\mathscr{C}(\mathfrak{S}_n, \mathbb{C})$ defined in corollary II.1.4.3 of chapter II. By that same corolllary,

$$\chi'_\lambda \cdot \chi'_\lambda = 1 + \sum_{\mu>\lambda} b_{\lambda\mu}^2.$$

So, to finish the proof, we just need to show that $\chi'_\lambda \cdot \chi'_\lambda$, i.e. (by definition of $\cdot$) that

$$\sum_{\sigma\in\mathfrak{S}_n} |\chi'_\lambda(\sigma)|^2 = n!.$$

By the decomposition of elements of $\mathfrak{S}_n$ into products of cycles of disjoint supports, conjugacy classes in $\mathfrak{S}_n$ are given by family $(c_r)_{\geq 1}$ of nonnegative integers such that $n = \sum_{r\geq 1} r c_r$ (c_r is the number of cycles of length r in the decomposition of any element of the conjugacy class). If $\underline{c} = (c_r)_{r\geq 1}$ is any such family and $C_{\underline{c}}$, then, by proposition IV.4.2,

$$|C_{\underline{c}}| = \frac{n!}{\prod_{r\geq 1} c_r! r^{c_r}}.$$

So $\chi'_\lambda \cdot \chi'_\lambda$ is the sum over all such families $(c_r)_{r\geq 1}$ of $\frac{1}{\prod_{r\geq 1} c_r! r^{c_r}}$ times the square of the coefficient of $\prod_{i=1}^N T_i^{\lambda_i + N - i}$ in $\Delta(T)\prod_{r\geq 1} P_r(T)^{c_r}$. Note that, if we take an arbitrary family $(c_r)_{r\geq 0}$ of nonnegative integers that are almost all 0, then the coefficient of $\prod_{i=1}^N T_i^{\lambda_i+N-i}$ in $\Delta(T)\prod_{r\geq 1} P_r(T)^{c_r}$ is 0 unless $\sum_{r\geq 1} r c_r = \sum_{i=1}^N \lambda_i$ for degree reasons. So, in the formula for $\chi'_\lambda \cdot \chi'_\lambda$ that we just got, we can take the sum over all families $(c_r)_{r\geq 0}$ of nonnegative integers that are almost all 0, and we get that $\chi'_\lambda \cdot \chi'_\lambda$ is the coefficient of $\prod_{i=1}^N \prod_{j=1}^N T_i^{\lambda_i+N-i} U_j^{\lambda_j+N-j}$ in $\Delta(T)\Delta(U)S(T,U)$, where

$$S(T,U) = \sum_{(c_r)_{r\geq 1}\in\mathbb{N}^{\mathbb{Z}_{\geq 1}}} \prod_{r\geq 1} \frac{P_r(T)^{c_r} P_r(U)^{c_r}}{c_r! r^{c_r}},$$

where the sum is over families $(c_r)_{\geq 1}$ of nonnegative integers that are almost all 0.

But we have

$$S(T,U) = \sum_{(c_r)} \prod_{r\geq 1} \frac{1}{c_r!} \left(\sum_{i,j=1}^{N} \frac{T_i U_j}{r} \right)^{c_r} = \prod_{r\geq 1} \exp\left(\sum_{i,j=1}^{N} \frac{T_i U_j}{r} \right) = \exp\left(\sum_{r\geq 1} \sum_{i,j=1}^{N} \frac{T_i U_j}{r} \right) =$$

$$\exp\left(-\sum_{i,j=1}^{N} \log(1 - T_i U_j) \right) = \prod_{i,j=1}^{N} \frac{1}{1 - T_i U_j}.$$

So by lemma IV.4.6, $\Delta(T)\Delta(U)S(T,U)$ is the determinant of the $N \times N$ matrix $(\frac{1}{1-T_i U_j})_{1\leq i,j\leq N}$, and we have

$$\Delta(T)\Delta(U)S(T,U) = \sum_{\tau \in \mathfrak{S}_N} \operatorname{sgn}(\tau) \prod_{i=1}^{N} \frac{1}{1 - T_i U_{\sigma(i)}}.$$

Remember that $\chi'_\lambda \cdot \chi'_\lambda$ is the coefficient of $\prod_{i=1}^{N} \prod_{j=1}^{N} T_i^{\lambda_i + N - i} U_j^{\lambda_j + N - j}$ in this formal series. If $\tau \neq 1$, then there exists $r \in \{1, \ldots, N\}$ such that $s := \tau(r) > r$. In the formal power series expansion of $\prod_{i=1}^{N} \frac{1}{1-T_i U_{\sigma(i)}}$, T_r and U_s must have the same exponent in each term. In particular, $\prod_{i=1}^{N} \prod_{j=1}^{N} T_i^{\lambda_i + N - i} U_j^{\lambda_j + N - j}$ does not appear in this expansion, because the exponent $\lambda_r + N - r$ of T_r in this product is greater than the exponent $\lambda_s + N - s$ of U_s. So $\chi'_\lambda \cdot \chi'_\lambda$ is the coefficient of $\prod_{i=1}^{N} \prod_{j=1}^{N} T_i^{\lambda_i + N - i} U_j^{\lambda_j + N - j}$ in $\prod_{i=1}^{N} \frac{1}{1-T_i U_i}$, i.e. 1, and we are done.

□

V Representations of compact groups

V.1 Topological groups, Haar measures, representations

Definition V.1.1. A *topological group* is a topological space G with a group structure such that the maps $G^2 \to G$, $(x, y) \longmapsto xy$, and $G \to G$, $X \longmapsto x^{-1}$, are both continuous.

When we talk about a measure on a topological group G, we will always mean a measure on the σ-algebra of Borel sets in G, i.e. the σ-algebra generated by the open subsets of G.

Theorem V.1.2. [1] *Let G be a compact Hausdorff topological group. Let $\mathscr{C}(G, \mathbb{C})$ be the $\mathbb{C}$-algebra of continuous functions from G to $\mathbb{C}$, with the norm $\|.\|_\infty$ given by $\|f\|_\infty = \sup_{x \in G} |f(t)|$.*

Then there exists a unique $\mathbb{C}$-linear map $\lambda : \mathscr{C}(G, \mathbb{C}) \to \mathbb{C}$ such that :

1. *λ is positive, i.e. $\lambda(f) \geq 0$ if $f(G) \subset \mathbb{R}_{\geq 0}$.*
2. *λ is left invariant, i.e. $\lambda(f) = \lambda(f(g.))$, for every $f \in \mathscr{C}(G, \mathbb{C})$ and every $g \in G$ (where $f(g.)$ is the function $x \longmapsto f(gx)$).*
3. *λ is right invariant, i.e. $\lambda(f) = \lambda(f(.g))$, for every $f \in \mathscr{C}(G, \mathbb{C})$ and every $g \in G$ (where $f(g.)$ is the function $x \longmapsto f(xg)$).*
4. *$\lambda(1) = 1$.*

Moreover, λ is continuous, there exists a unique probability measure dg on G such that, for every $f \in \mathscr{C}(G, \mathbb{C})$,

$$\lambda(f) = \int_G f(g) dg,$$

and this measure also satisfies

$$\int_G f(g) dg = \int_G f(g^{-1}) dg,$$

for every measurable function $f : G \to \mathbb{C}$.

[1] See theorem 5.14 of Rudin's book [26] for compact groups and chapter VI of Loomis's book [21] for the general case. See also problems VII.5.1, VII.5.2 and VII.5.3.

We say that dg is a bi-invariant (or left and right-invariant) probability Haar measure on G.

In the rest of this chapter, if we have a compact Hausdorff topological group G, the notation dg will always mean a bi-invariant probability Haar measure on G.

Remark V.1.3. If we only assume that G is locally compact, then we can find nonzero $\mathbb{C}$-linear maps satisfying (i) and (ii) (resp. (i) and (iii)), and they are unique up to multiplication by a nonnegative real number. These functions are also continuous, and the corresponding measures on G are called left-invariant (resp. right-invariant) Haar measures. If $d_l g$ is a left-invariant Haar measure on G, then there is a unique right-invariant Haar measure $d_r g$ on G such that

$$\int_G f(g) d_r g = \int_G f(g^{-1}) d_l g,$$

for every measurable function $f : G \to \mathbb{C}$.

Example V.1.4.

- A finite group G with the discrete topology is a topological group. A left and right-invariant Haar measure on G is given by $dg(A) = |A|/|G|$.
- Let $G = \mathrm{U}(1) := \{z \in \mathbb{C} || z| = 1\}$, with the topology induced by that of $\mathbb{C}$. This is a topological group, and we have an isomorphism of topological groups $\mathbb{R}/\mathbb{Z} \xrightarrow{\varphi} \mathrm{U}(1)$, $t \longmapsto e^{2i\pi t}$ (where $\mathbb{R}/\mathbb{Z}$ is given the quotient topology). We get a Haar measure on G by taking, for every measurable $f : G \to \mathbb{C}$,

  $$\int_G f(g) dg = \int_0^1 (f \circ \varphi)(t) dt = \int_0^1 f(e^{2i\pi t}) dt,$$

  where dt is the usual Lebesgue measure on $\mathbb{R}$. By the way, not that dt itself is a Haar measure on the topological group $(\mathbb{R}, +)$.

Definition V.1.5. Let G be a topological group and V be a normed $\mathbb{C}$-vector space. Then a *(continuous) representation* of G on V is an abstract representation of G on V such that the action map $G \times V \to V$, $(g, v) \longmapsto gv$, is continuous.

Remark V.1.6.

- The definition makes sense if V is any topological vector (over a topological field).
- With notation as in the definition, let $\mathrm{End}(V)$ be the $\mathbb{C}$-algebra of continuous endomorphisms of V. We put the operator norm on $\mathrm{End}(V)$, and consider a continuous representation of G on V. Then the action of every $g \in G$ on V is a continuous endomorphism of V, so we get a map $G \to \mathrm{End}(V)$. But this map is not continuous in general. (See problem VII.5.7 for a counterexample.)
- With notation as in the previous remark, if $\rho : G \to \mathrm{End}(V)$ is an abstract representation of G on V that is continuous for the weak* topology on $\mathrm{End}(V)$, then it is not

necessarily a continuous representation. (For example, take for G the group of invertible elements of $\mathrm{End}(V)$, with the topology induced by the weak* topology on $\mathrm{End}(V)$, and for $\rho : G \to \mathrm{End}(V)$ the inclusion. This is not a continuous representation of G on V.)

So we see that we have to be a bit careful with the notion of continuous representation in general. In the following two sections, we will see what happens in the particular case of finite-dimensional vector spaces, and in that of unitary representations on Hilbert spaces.

V.2 Finite-dimensional representations

Definition V.2.1. If V is a normed $\mathbb{C}$-vector space, we denote by $\mathrm{End}(V)$ the $\mathbb{C}$-algebra of continuous endomorphisms of V, and we put on it the topology given by the operator norm. We write $\mathrm{GL}(V)$ for $\mathrm{End}(V)^\times$, with the topology induced by that of $\mathrm{End}(V)$.

Remember that, if V is a finite-dimensional $\mathbb{C}$-vector space, then all norms on V are equivalent. So V has a canonical topology, and so does $\mathrm{End}(V)$ (as another finite-dimensional vector space).

Proposition V.2.2. *Let V be a normed $\mathbb{C}$-vector space and $\rho : G \to \mathrm{GL}(V)$ be a morphism of groups. Consider the following conditions.*

(i) The map $G \times V \to V$, $(g, v) \longmapsto \rho(g)(v)$, is continuous (i.e. ρ is a continuous representation of G on V).

(ii) For every $v \in V$, the map $G \to V$, $g \longmapsto \rho(g)(v)$, is continuous.

(iii) The map $\rho : G \to \mathrm{GL}(V)$ is continuous.

Then we have (iii)⇒(i)⇒(ii). If moreover V is finite-dimensional, then all three conditions are equivalent.

Proof.

(i)⇒(ii) is obvious.

(ii)⇒(iii) : Suppose that V is finite-dimensional, and let $(e_1, \dots, e_n)$ be a basis of V, and let $\|.\|$ be the norm on V defined by $\| \sum_{i=1}^{x} x_i e_i \| = \sup_{1 \leq i \leq n} |x_i|$. We use the corresponding operator norm on $\mathrm{End}(V)$ and still denote it by $\|.\|$. Let $g_0 \in G$ and let $\varepsilon > 0$; we are looking for a neighborhood U of $g_0 \in G$ such that : $g \in U \Rightarrow \|\rho(g) - \rho(g_0)\| \leq \varepsilon$.

For every $i \in \{1, \dots, n\}$, the function $G \to V$, $g \longmapsto \rho(g)(e_i)$, is continuous by assumption, so there exists a neighborhood U_i of g_0 in G such that : $g \in U \Rightarrow \|\rho(g)(e_i) - \rho(g_0)(e_i)\| \leq \varepsilon/n$. Let $U = \bigcap_{i=1}^{n} U_i$. Then if $g \in U$, for ev-

ery $v = \sum_{i=1}^{n} x_i e_i \in V$, we have

$$\|\rho(g)(v) - \rho(g_0)(v)\| \leq \sum_{i=1}^{n} \|x_i\| \|\rho(g)(e_i) - \rho(g_0)(e_i)\| < \sum_{i=1}^{n} |x_i| \varepsilon/n \leq \varepsilon \|v\|,$$

which means that $\|\rho(g) - \rho(g_0)\| \leq \varepsilon$.

(iii)⇒(i) : Let $g_0 \in G$, $v_0 \in V$, and $\varepsilon > 0$. We want to find a neighborhood U of g and G and a $\delta > 0$ such that : $g \in U$ and $\|v - v_0\| < \delta \Rightarrow \|\rho(g)(v) - \rho(g_0)(v_0)\| < \varepsilon$.

Choose a δ such that $0 < \delta \leq \frac{\varepsilon}{2\|\rho(g_0)\|}$, and let U be a neighborhood of g_0 in G such that : $g \in G \Rightarrow \|\rho(g) - \rho(g_0)\| < \frac{\varepsilon}{2(\|v_0\|+\delta)}$. Then, if $g \in U$ and $\|v - v_0\| < \delta$, we have $\|v\| \leq \|v_0\| + \delta$, and hence

$$\begin{aligned} \|\rho(g)(v) - \rho(g_0)(v_0)\| &\leq \|\rho(g)(v) - \rho(g_0)(v)\| + \|\rho(g_0)(v) - \rho(g_0)(v_0)\| \\ &\leq \|\rho(g) - \rho(g_0)\| \|v\| + \|\rho(g_0)\| \|v - v_0\| \\ &< \frac{\varepsilon}{2(\|v_0\| + \delta)} (\|v_0\| + \delta) + \|\rho(g_0)\| \delta \\ &\leq \varepsilon/2 + \varepsilon/2 = \varepsilon. \end{aligned}$$

□

For finite-dimensional continuous representations of continuous groups, we can define subrepresentations, irreducible representations, direct sums, semisimple representations, tensor products, Homs and duals just as in the case of finite groups. Also, Schur's lemma (in the form of theorem II.1.2.1 of chapter II) still holds with exactly the same proof. We will see in the next section what happens to Schur orthogonality (theorem II.1.2.2 of chapter II), after we introduce unitary representations : in the case of compact Hausdorff groups, once we formulate it correctly, it still holds.

V.3 Unitary representations

V.3.1 Definition and first properties

Remember that a *(complex) Hilbert space* is a $\mathbb{C}$-vector space V with a Hermitian inner product[2] such that V is complete for the corresponding norm. If V is a finite-dimensional $\mathbb{C}$-vector space with a Hermitian inner product, then it is automatically complete, hence a Hilbert space.

Notation V.3.1.1. Let V and W be Hermitian inner product spaces. For every continuous $\mathbb{C}$-linear map $T : V \to W$, we write $T^* : W \to V$ for the adjoint of T, if it exists. (It always does if V and W are Hilbert spaces.) If V' is a subspace of V, we write $V'^{\perp}$ for the orthogonal of V'. Finally, we write $\mathrm{U}(V)$ for the group of unitary endomorphisms of V.

[2]We will always assume Hermitian inner products to be $\mathbb{C}$-linear in the first variable.

Proposition V.3.1.2. *If V is a Hilbert space and $\rho : G \to \mathrm{U}(V)$ is a morphism of groups, then the following are equivalent :*

1. *The map $G \times V \to V$, $(g, v) \longmapsto \rho(g)(v)$, is continuous.*
2. *For every $v \in V$, the map $G \to V$, $g \longmapsto \rho(g)(v)$, is continuous.*

Definition V.3.1.3. If V is a Hilbert space, a *unitary representation* of G on V is a morphism of groups $\rho : G \to \mathrm{U}(V)$ satisfying the conditions of the proposition above.

Remark V.3.1.4. Note that ρ is not a continuous map in general. (Unless $\dim_{\mathbb{C}} V < +\infty$, in which case ρ is continuous by proposition V.2.2.)

Also, note that we don't need the completeness of V in the proof, so the proposition is actually true for any Hermitian inner product space.

Proof of the proposition. We already seen in proposition V.2.2 that (i) implies (ii). Let's prove that (ii) implies (i). Let $g_0 \in G$, $v_0 \in V$, and $\varepsilon > 0$. We want to find a neighborhood U of g in G and a $\delta > 0$ such that : $g \in U$ and $\|v - v_0\| < \delta \Rightarrow \|\rho(g)(v) - \rho(g_0)(v_0)\| < \varepsilon$.

Choose a neighborhood U of g in G such that : $g \in U \Rightarrow \|\rho(g)(v_0) - \rho(g_0)(v_0)\| < \varepsilon/2$, and take $\delta = \varepsilon/2$. Then, if $g \in U$ and $\|v - v_0\| < \delta$, we have

$$\begin{array}{rcl} \|\rho(g)(v) - \rho(g_0)(v_0)\| & \leq & \|\rho(g)(v) - \rho(g)(v_0)\| + \|\rho(g)(v_0) - \rho(g_0)(v_0)\| \\ & < & \|\rho(g)\|\|v - v_0\| + \varepsilon/2 \\ & < & \varepsilon/2 + \varepsilon/2 = \varepsilon, \end{array}$$

because $\rho(g) \in \mathrm{U}(V)$, so $\|\rho(g)\| = 1$.

□

Remark V.3.1.5. If $\rho : G \to \mathrm{U}(V)$ is a unitary representation of G on a Hilbert space $(V, \langle ., . \rangle)$, then, for every G-invariant subspace W of V, the subspace $W^\perp$ is also G-invariant. Indeed, if $w \in V^\perp$ and $g \in G$, then, for every $v \in W$,

$$\langle v, \rho(g)w \rangle = \langle \rho(g)^{-1}v, w \rangle = 0,$$

so $\rho(g)w \in V^\perp$.

In particular, if W is a *closed* G-invariant subspace of V, then we have $V = W \oplus W^\perp$ with $W^\perp$ a closed G-invariant subspace. If V is finite-dimensional, then every subspace is closed and this shows that every unitary representation of G on V is semisimple.

Theorem V.3.1.6. *Assume that the group G is compact Hausdorff. Let V be a finite-dimensional $\mathbb{C}$-vector space and $\rho : G \to \mathrm{GL}(V)$ be a continuous representation of G on V. Then there exists a Hermitian inner product on V that makes ρ a unitary representation.*

Remark V.3.1. If V is irreducible, we can also prove that this inner product is unique up to a constant. See problem VII.5.9.

Proof of the theorem. Let $\langle .,. \rangle_0$ a Hermitian inner product on V. We define $\langle .,. \rangle : V \times V \to \mathbb{C}$ by the following formula : for all $v, w \in V$,

$$\langle v, w \rangle = \int_G \langle \rho(g)v, \rho(g)w \rangle_0 dg$$

(remember that dg is a bi-invariant probability Haar measure on G). We clearly have $\langle \rho(g)v, \rho(g)w \rangle = \langle v, w \rangle$ for every $v, w \in V$ and $g \in G$, so we just need to show that $\langle .,. \rangle$ is a Hermitian product on V. It's clearly a Hermitian product, so we just need to show that it is positive definite. Let $v \in V - \{0\}$. Then the function $G \to \mathbb{R}$, $g \longmapsto \langle \rho(g)v, \rho(g)v \rangle_0$, is continuous and takes positive values. As G is compact, there exists $\varepsilon > 0$ such that $\langle \rho(g)v, \rho(g)v \rangle_0 > \varepsilon$ for every $g \in G$, and then we have $\langle v, v \rangle \geq \varepsilon > 0$.

□

Corollary V.3.1.7. *If G is compact Hausdorff, then every finite-dimensional continuous representation of G is semisimple.*

V.3.2 The operators $T^0_{v,w}$

The following construction will be used several times in proofs below.

Definition V.3.2.1. Let V and W be Hermitian inner product spaces; we denote both inner products by $\langle .,. \rangle$. If $v \in V$ and $w \in W$, we define the $\mathbb{C}$-linear map $T^0_{v,w} : V \to W$ by $T^0_{v,w}(x) = \langle x, v \rangle w$.

Note that the map $V \to \mathrm{Hom}_{\mathbb{C}}(V, W)$, $v \longmapsto T^0_{v,w}$, is semi-linear.

Proposition V.3.2.2. *1. For every $v \in V$ and $w \in W$, ${T^0_{v,w}}^* = T^0_{w,v}$.*

From now, we suppose that V and W are finite-dimensional.

2. The $T^0_{v,w}$, for $v \in V$ and $w \in W$, generate $\mathrm{Hom}_{\mathbb{C}}(V, W)$ as a $\mathbb{C}$-vector space.

3. If $V = W$, then for every $v, w \in V$, $\mathrm{Tr}(T^0_{v,w}) = \langle w, v \rangle$.

4. For every $v_1, v_2 \in V$ and $w_1, w_2 \in W$, we have

$$\mathrm{Tr}(T^0_{v_1,w_1} {T^0_{v_2,w_2}}^*) = \langle w_1, w_2 \rangle \langle v_2, v_1 \rangle.$$

Proof. 1. Let $x \in V$ and $y \in W$. Then

$$\langle T^0_{v,w}(x), y \rangle = \langle \langle x, v \rangle w, y \rangle = \langle x, v \rangle \langle w, y \rangle = \langle x, v \rangle \overline{\langle y, w \rangle} = \langle x, \langle y, w \rangle v \rangle = \langle x, T^0_{w,v}(y) \rangle.$$

2. If we choose orthonomal bases $(e_i)_{1 \leq i \leq n}$ of V and $(f_j)_{1 \leq j \leq m}$ of W, then, for every $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, m\}$, the matrix of $T^0_{e_i,f_j}$ in these bases is the $m \times n$ with (i, j)-entry equal to 1 and all other entries equal to 0. These clearly generate the $\mathbb{C}$-vector space $M_{mn}(\mathbb{C})$.

3. Let $(v_1, \dots, v_n)$ be an orthogonal basis of V such that $v_1 = v$. Then

$$T^0_{v,w}(v_i) = \begin{cases} \langle v_1, v_1\rangle w & \text{if } i = 1 \\ 0 & \text{otherwise.} \end{cases}$$

   As $w = \sum_{i=1}^n \frac{\langle w, v_i\rangle}{\langle v_i, v_i\rangle} v_i$, this gives the result.

4. For every $y \in W$,

$$T^0_{v_1,w_1} {T^0_{v_2,w_2}}^*(y) = T^0_{v_1,w_1} T^0_{w_2,v_2}(y) = T^0_{v_1,w_1}(\langle y, w_2\rangle v_2) = \langle y, w_2\rangle\langle v_2, v_1\rangle w_1.$$

   Choose an orthogonal basis $(y_1, \dots, y_n)$ of W such that $y_1 = w_2$. Then

$$T^0_{v_1,w_1} {T^0_{v_2,w_2}}^*(y_i) = \begin{cases} \langle y_1, y_1\rangle\langle v_2, v_1\rangle w_1 & \text{if } i = 1 \\ 0 & \text{otherwise.} \end{cases}$$

   As $w_1 = \sum_{i=1}^n \frac{\langle w_1, y_i\rangle}{\langle y_i, y_i\rangle} y_i$, this gives the result.

□

V.3.3 Schur orthogonality

In this section, G is a compact Hausdorff group, and we will only use finite-dimensional complex continuous representations of G.

Definition V.3.3.1. Let $\rho : G \to \mathrm{GL}(V)$ be a continuous representation of G on a finite-dimensional complex vector space. Remember that the map $\chi_V : G \to \mathbb{C}$, $g \longmapsto \mathrm{Tr}(\rho(g))$, is called the *character* of the representation (V, ρ).

By proposition V.2.2, $\chi_V : G \to \mathbb{C}$ is a continuous map.

Theorem V.3.3.2. *Let V, W be continuous representations of G on finite-dimensional complex vector spaces. Assume that V and W are both irreducible. Choose G-invariant Hermitian inner products on V and W, that will both be denoted by $\langle ., .\rangle$. Then for every $v_1, v_2 \in V$ and $w_1, w_2 \in W$,*

$$\int_G \langle gv_1, v_2\rangle \overline{\langle gw_1, w_2\rangle} dg = \begin{cases} \frac{1}{\dim_{\mathbb{C}} V}\langle v_1, w_1\rangle\overline{\langle v_2, w_2\rangle} & \text{if } V \simeq W \\ 0 & \text{otherwise.} \end{cases}$$

Proof. If $v \in V$ and $w \in W$, let

$$T_{v,w} = \int_G g T^0_{v,w} g^{-1} dg \in \mathrm{Hom}_{\mathbb{C}}(V, W),$$

where $T^0_{v,w}$ is as in definition V.3.2.1. Then $T_{v,w}$ is G-equivariant, so, by Schur's lemma, $T_{v,w} = 0$ if $V \not\simeq W$, and, if $V = W$, then $T_{v,w} = c(v, w)\mathrm{id}_V$ with $c(v, w) \in \mathbb{C}$..

Suppose that $V = W$, and let's calculate $c(v, w)$. We have

$$c(v, w) \dim_{\mathbb{C}} V = \mathrm{Tr}(T_{v,w}) = \int_G \mathrm{Tr}(g T^0_{v,w} g^{-1}) dg = \mathrm{Tr}(T^0_{v,w}),$$

so, by proposition V.3.2.2, $c(v, w) = \frac{1}{\dim_{\mathbb{C}} V} \langle w, v \rangle$.

Now let $I = \int_G \langle gv_1, v_2 \rangle \overline{\langle gw_1, w_2 \rangle} dg$. We have

$$\begin{aligned} I &= \int_G \langle gv_1, v_2 \rangle \langle g^{-1} w_2, w_1 \rangle dg \\ &= \int_G \langle \langle g^{-1} w_2, w_1 \rangle gv_1, v_2 \rangle dg \\ &= \int_G \langle g T^0_{w_1,v_1} g^{-1} w_2, v_2 \rangle dg \\ &= \langle T_{w_1,v_1}(w_2), v_2 \rangle \end{aligned},$$

so $I = 0$ if $V \simeq W$, and

$$I = c(w_1, v_1) \langle w_2, v_2 \rangle = \frac{1}{\dim_{\mathbb{C}} V} \langle v_1, w_1 \rangle \overline{\langle v_2, w_2 \rangle}$$

if $V \simeq W$.

□

Corollary V.3.3.3 (Schur orthogonality)**.** *Let V, W be continuous representations of G on finite-dimensional complex vector spaces. Assume that V and W are both irreducible. Then*

$$\int_G \chi_V(g) \overline{\chi_W(g)} dg = \begin{cases} 0 & \text{if } V \simeq W \\ 1 & \text{if } V \simeq W. \end{cases}$$

Proof. Choose G-invariant Hermitian inner products on V and W, and fix orthonomal bases $(v_1, \ldots, v_n)$ of V and $(w_1, \ldots, w_m)$ of W, that we should to be equal if $V = W$. Then we know that, for any $\mathbb{C}$-linear endomorphism u of V (resp. W), we have $\mathrm{Tr}(u) = \sum_{i=1}^n \langle u(v_i), v_i \rangle$ (resp. $\mathrm{Tr}(u) = \sum_{j=1}^m \langle u(w_i), w_i \rangle$). In particular, for every $g \in G$,

$$\chi_V(g) = \sum_{i=1}^n \langle gv_i, v_i \rangle$$

and

$$\chi_W(g) = \sum_{j=1}^m \langle gw_j, w_j \rangle.$$

Hence

$$\int_G \chi_V(g) \overline{\chi_W(g)} dg = \sum_{i=1}^n \sum_{j=1}^m \int_G \langle gv_i, v_i \rangle \overline{\langle gw_j, w_j \rangle} dg.$$

By the theorem, this is equal to 0 if $V \not\simeq W$, and, if $V = W$, it's equal to

$$\frac{1}{\dim_{\mathbb{C}} V} \sum_{i=1}^{n} \sum_{j=1}^{m} \langle v_i, v_j \rangle \overline{\langle v_i, v_j \rangle} = 1.$$

□

Corollary V.3.3.4. *If V is a finite-dimensional continuous representation of G, write $V = \bigoplus_{i \in I} V_i^{\oplus n_i}$, where the V_i are irreducible. (This is possible by corollary V.3.1.7.) Then*

$$\int_G |\chi_V(g)|^2 dg = \sum_{i \in I} n_i^2.$$

In particular, V is irreducible if and only if $\int_G |\chi_V(g)|^2 dg = 1$.

V.4 The space $L^2(G)$

From now on, we take G to be a compact Hausdorff group and dg to be a bi-invariant probability Haar measure on G.

V.4.1 Definition and actions of G

Definition V.4.1.1. We write $L^2(G)$ for the quotient

$$\{f : G \to \mathbb{C} \text{ measurable} | \int_G |f(g)|^2 dg < +\infty\} / \{f : G \to \mathbb{C} \text{ measurable} | \int_G |f(g)|^2 dg = 0\}.$$

This is a Hilbert space for the Hermitian inner product given by

$$\langle f_1, f_2 \rangle = \int_{g \in G} f_1(g) \overline{f_2(g)} dg.$$

If $f \in L^2(G)$, we write $\|f\|_2 = \sqrt{\langle f, f \rangle}$.

Definition V.4.1.2. If $x \in G$ and f if a function from G to $\mathbb{C}$, define $R_x f, L_x f : G \to \mathbb{C}$ by $L_x(g) = L(xg)$ and $R_x(g) = R(gx)$.

Proposition V.4.1.3. *Let $x, y \in G$.*

1. *The operators defined above induced endomorphisms L_x and R_x of $L^2(G)$, with the following properties :*

2. $L_x R_y = R_y L_x$, $R_x R_y = R_{xy}$ *and* $L_x L_y = L_{yx}$.
3. $R_x^* = R_x^{-1} = R_{x^{-1}}$, *and* $L_x^* = L_x^{-1} = L_{x^{-1}}$. *In particular,* R_x *and* L_x *are unitary endomorphisms of* $L^2(G)$.

Proof.

Let $f_1, f_2 : G \to \mathbb{C}$ be measurable functions. Then, using the right and left invariance of dg, we get

$$\langle R_x f_1, R_x f_2 \rangle = \int_G \langle f_1(gx)\overline{f_2(gx)}dg = \int_G f_1(g)\overline{f_2(g)}dg = \langle f_1, f_2 \rangle$$

and

$$\langle L_x f_1, L_x f_2 \rangle = \int_G \langle f_1(xg)\overline{f_2(xg)}dg = \int_G f_1(g)\overline{f_2(g)}dg = \langle f_1, f_2 \rangle.$$

So R_x and L_x preserve both spaces in the quotient defining $L^2(G)$, and hence they induce endomorphisms of $L^2(G)$. Also, the equalities above show that these endomorphisms are unitary, which gives half of (iii). The other half of (iii) will follow from (ii), so let's prove (ii). Let $f \in L^2(G)$ and $g \in G$. Then we have

$$(L_x R_y f)(g) = (R_y f)(xg) = f(xgy) = (L_x f)(gy) = (R_y L_x f)(g),$$

$$(L_x L_y f)(g) = (L_y f)(xg) = f(yxg) = (L_{xy} f)(g)$$

and

$$(R_x R_y f)(g) = (R_y f)(gx) = f(gxy) = (R_{xy} f)(g),$$

which proves (ii).

□

Definition V.4.1.4. If f is a function $G \to \mathbb{C}$, we define a function $\widetilde{f} : G \to \mathbb{C}$ by $\widetilde{f}(g) = \overline{f(g^{-1})}$.

Proposition V.4.1.5. *This defines a unitary endomorphism* $f \longmapsto \widetilde{f}$ *of* $L^2(G)$.

Proof. Let $f : G \to \mathbb{C}$ be a measurable function. Then

$$\int_G |\overline{f(g^{-1})}|^2 dg = \int_G |f(g)|^2 dg$$

(because the group G is compact, so the Haar measure dg is equal to its own pullback by $g \longmapsto g^{-1}$, see theorem V.1.2), which implies every statement in the proposition.

□

V.4.2 The convolution product

Definition V.4.2.1. If $f_1, f_2 \in L^2(G)$, define $f_1 * f_2 : G \to \mathbb{C}$ by

$$(f_1 * f_2)(g) = \int_G f_1(gx) f_2(x^{-1}) dx.$$

This function is called the *convolution product* of f_1 and f_2.

Remark V.4.2.2. If we want this definition to make sense, we should be more careful. First, if $f_1, f_2 : G \to \mathbb{C}$ are measurable and L^2, then we can define $f_1 * f_2 : G \to \mathbb{C}$ by the formula above. Then, if one of f_1 and f_2 happens to be negligible (i.e. if $\int_G |f_1(g)|^2 dg = 0$ or $\int_G |f_2(g)| dg = 0$), then it is easy to see that the function $f_1 * f_2$ is identically zero. So the convolution indeed goes to the quotient and makes sense for $f_1, f_2 \in L^2(G)$. Note that, unlike f_1 and f_2, the function $f_1 * f_2$ is well-defined everywhere.

Proposition V.4.2.3. *1. For any $f_1, f_2 \in L^2(G)$, $f_1 * f_2$ is a bounded function on G, and we have $\|f_1 * f_2\|_\infty \leq \|f_1\|_2 \|f_2\|_2$, where $\|.\|_\infty$ is the supremum norm.*

In particular, $f_1 * f_2 \in L^2(G)$, and $\|f_1 * f_2\|_2 \leq \|f_1\|_2 \|f_2\|_2$.

2. The convolution product is associative, and distributive with respect to addition.

In other words, it makes $L^2(G)$ into an associative $\mathbb{C}$-algebra (with no unit, unless G is finite).

Proof. 1. Let $g \in G$. Then

$$|f_1 * f_2(g)| = |\int_G f_1(gx) f_2(x^{-1}) dx| = |\int_G (L_g f_1) \overline{\widetilde{f_2}(x)} dx| \leq \|R_g f_1\|_2 \|\widetilde{f_2}\|_2 = \|f_1\|_2 \|f_2\|_2$$

by the Cauchy-Schwarz inequality and propositions V.4.1.3 and V.4.1.5. This gives (i). (The second part of (i) follows from the fact that dg is a probability measure.)

2. The distributivity is obvious. Let $f_1, f_2, f_3 \in L^2(G)$, and let $g \in G$. Then :

$$((f_1 * f_2) * f_3)(g) = \int_G (f_1 * f_2)(gx) f_3(x^{-1}) dx = \int_G (\int_G f_1(gxy) f_2(y^{-1}) dy) f_3(x^{-1}) dx,$$

while

$$(f_1 * (f_2 * f_3))(g) = \int_G f_1(gz)(f_2 * f_3)(z^{-1}) dz = \int_G f_1(gz) (\int_G f_2(z^{-1}t) f_3(t^{-1}) dt) dz.$$

Using Fubini's theorem, the change of variables $z \longmapsto y$ and $t \longmapsto y^{-1}x$, and the fact that dg is a bi-invariant Haar measure on G, we see that these two expressions are equal.

□

Remark V.4.2.4. It is also true that $f_1 * f_2$ is continuous for every $f_1, f_2 \in L^2(G)$, but we will not need it here. To prove this, note that, by (i) of the proposition (and the bilinearity of $*$), if f_1 (resp. f_2) is the limit in $L^2(G)$ of a sequence $(f_{1,n})_{n\geq 0}$ (resp. $(f_{2,n})_{n\geq 0}$), then the sequence $(f_{1,n} * f_{2,n})_{n\geq 0}$ converges to $f_1 * f_2$ in $L^\infty(G)$. Now use the fact that continuous functions are dense in $L^2(G)$ (see theorem 3.14 of Rudin's book [25]), and that the convolution of two continuous functions is continuous.

Definition V.4.2.5. Let $f \in L^2(G)$. We define endomorphisms L_f, R_f of $L^2(G)$ by $L_f(f_1) = f * f_1$ and $R_f(f_1) = f_1 * f$.

By proposition V.4.2.3, we get :

Corollary V.4.2.6. *1. These operators are well-defined and continuous, and we have* $\|L_f\| \leq \|f\|_2$, $\|R_f\| \leq \|f\|_2$. [3]

2. For any $f_1, f_2 \in L^2(G)$, $L_{f_1}R_{f_2} = R_{f_2}L_{f_1}$, $R_{f_1}R_{f_2} = R_{f_1 * f_2}$ *and* $L_{f_1}L_{f_2} = L_{f_2 * f_1}$.

Moreover :

Proposition V.4.2.7. *1. For every* $f \in L^2(G)$, $R_f^* = R_{\widetilde{f}}$ *and* $L_f^* = L_{\widetilde{f}}$.

2. For every $f \in L^2(G)$ *and* $x \in G$, $R_x L_f = L_f R_x$ *and* $R_f L_x = L_x R_f$.

Proof. 1. We only prove the first equality, the second one is similar. Let $f_1, f_2 \in L^2(G)$. Then :

$$\langle R_f(f_1), f_2\rangle = \int_G (f_1 * f)(g)\overline{f_2(g)}dg = \int_G\int_G f_1(gx^{-1})f(x)\overline{f_2(g)}dxdg.$$

After the change of variables $y = gx^{-1}$, we see that this is equal to

$$\int_G\int_G f_1(y)\overline{f_2(yx)}f(x)dxdy.$$

After the change of variables $z = x^{-1}$, we see that this is equal to

$$\int_G\int_G f_1(y)\overline{f_2(yz^{-1})}\overline{\widetilde{f}(z)}dzdy = \int_G f_1(y)\overline{(f_2 * \widetilde{f})(y)}dy = \langle f_1, R_{\widetilde{f}}f_2\rangle.$$

2. We only prove the first equality, the second one is similar. Let $h \in L^2(G)$ and $g \in G$. Then

$$(R_x L_f(h))(g) = (f * h)(gx) = \int_G f(gxy)h(y^{-1})dy,$$

while

$$(L_f R_x(h))(g) = (f * (R_x h))(g) = \int_G f(gz)h(z^{-1}x)dz.$$

[3] In fact these are equalities, see remark V.4.2.9.

The change of variables $z = xy$ (and the fact that the measure on G is right-invariant) show that these two expressions are equal.

□

Theorem V.4.2.8. *For every $f \in L^2(g)$, there exists a sequence $(f_n)_{n\geq 0}$ of functions in $L^2(G)$ such that :*

1. $\widetilde{f_n} = f_n$ *for every* $n \geq 0$.
2. $\|f_n\|_2 = 1$ *for every* $n \geq 0$.
3. $f * f_n \to f$ *in* $L^2(G)$ *as* $n \to +\infty$.

Proof. Let $\varepsilon > 0$. Choose a neighborhood U of 1 in G such that $U = U^{-1}$ and that, for any $x, y \in U$, $\|R_x f - f\|_2 < \varepsilon$. (This is possible by corollary V.4.2.6.) Let $h_\varepsilon = \frac{1}{\mathrm{vol}(U)}\mathbb{1}_U$. Obviously, $\widetilde{h_\varepsilon} = h_\varepsilon$ and $\|h_\varepsilon\|_2 = 1$. Moreover, we have

$$(f * h_\varepsilon)(g) = \int_G f(gx^{-1})h_\varepsilon(x)dx = \frac{1}{\mathrm{vol}(U)}\int_U f(gx^{-1})dx$$

for every $g \in G$, so

$$(f * h_\varepsilon)(g) - f(g) = \frac{1}{\mathrm{vol}(U)}\int_U (R_{x^{-1}}(f)(g) - f(g))dx,$$

and

$$\begin{aligned}
\|(f * h_\varepsilon) - f\|_2^2 &= \int_G |(f * h_\varepsilon)(g) - f(g)|^2 dg \\
&= \frac{1}{\mathrm{vol}(U)^2}\int_{G\times U\times U} (R_{x^{-1}}f - f)(g)\overline{(R_{y^{-1}}f - f)(g)}dxdydg \\
&= \frac{1}{\mathrm{vol}(U)^2}\int_{U\times U} \langle R_{x^{-1}}f - f, R_{y^{-1}}f - f\rangle dxdy \\
&\leq \frac{1}{\mathrm{vol}(U)^2}\int_{U\times U} \|R_{x^{-1}}f - f\|_2\|R_{y^{-1}}f - f\|_2 dxdy \\
&\leq \frac{1}{\mathrm{vol}(U)^2}\int_{U\times U} \varepsilon^2 dxdy = \varepsilon^2
\end{aligned}$$

(by the Cauchy-Schwarz inequality and the choice of U).

Now take $f_n = h_{1/2^n}$.

□

Remark V.4.2.9. In particular, this implies that $\|R_f\| = \|f\|_2$. We have a similar result where we take the convolution by f on the left (the proof is exactly the same), and it implies that $\|L_f\| = \|f\|_2$.

Theorem V.4.2.10. *For every $f \in L^2(G)$, the endomorphisms R_f and L_f of $L^2(G)$ are compact.* [4]

Lemma V.4.2.11. *If $f_1, f_2 : G \to \mathbb{C}$, define a function $f_1 \otimes f_2 : G \times G \to \mathbb{C}$ by $(f_1 \otimes f_2)(g_1, g_2) = f_1(g_1)f_2(g_2)$. This induces a $\mathbb{C}$-linear map $L^2(G) \otimes_{\mathbb{C}} L^2(G) \to L^2(G \times G)$, which is injective with dense image.*

Proof. Take a Hilbert basis $(e_i)_{i\in I}$ of $L^2(G)$. Then the family $(e_i \otimes e_j)_{i,j\in I}$ of $L^2(G \times G)$ is clearly orthonormal, and we see easily that any function on $L^2(G \times G)$ that is orthogonal to all the $e_i \otimes e_j$ is 0 almost everywhere, hence 0.

□

Lemma V.4.2.12. *For every $K \in L^2(G \times G)$, define a $T_K : L^2(G) \to L^2(G)$ by $(T_K(h))(x) = \int_G K(x, y)f(y)dy$. Then $\|T_K\| \leq \|K\|_2$.*

Proof. This is an easy calculation.

□

Proof of the theorem. We prove the result for L_f; the proof for R_f is similar.

Consider the function $K : G \times G \to \mathbb{C}$, $(x, y) \longmapsto f(xy^{-1})$. Then $K \in L^2(G \times G)$, so, by the lemma, there exists families of functions $(g_i)_{i\in I}$ and $(h_i)_{i\in I}$ in $L^2(G)$ such that the sum $\sum_{i\in I} g_i \otimes h_i$ converges to K in $L^2(G \times G)$. Note also that $T_K = L_f$.

For every finite subset J of I, let $S_J = \sum_{i,j\in J} g_i \otimes h_j \in L^2(G \times G)$ and $T_J = T_{S_J}$. Then, for every $h \in L^2(G)$, for every $x \in G$,

$$(T_J h)(x) = \sum_{(i,j)\in J^2} \int_G h(y)g_i(x)h_j(y)dy = \sum_{i\in J}(\sum_{j\in J} \int_G h(y)h_j(y)dy)g_i(x).$$

In other words, for every $h \in L^2(G)$, $T_J h$ is in the finite-dimensional subspace of $L^2(G)$ spanned by the g_i, $i \in J$. Hence the operator T_J has finite rank.

To show that L_f is compact, it suffices by theorem V.6.2 to show that it is the limit of the operators T_J as J becomes bigger. But this follows from the second lemma and from the fact that K is the limit of the S_J.

□

V.4.3 The regular representations

Definition V.4.3.1. We make $G \times G$ act on $L^2(G)$ by $(x, y) \cdot f = L_{x^{-1}} R_y f = R_y L_{x^{-1}} f$, i.e. $((x, y) \cdot f)(g) = f(x^{-1}gy)$ for every $g \in G$.

[4] See definition V.6.1.

The restriction of this action to the first (resp. second) factor of $G \times G$ is called the left (resp. right) regular representation of G.

Proposition V.4.3.2. *The representation of $G \times G$ on $L^2(G)$ defined above is continuous, and so it's a unitary representation.*

Proof. We already know that $L_{x^{-1}}R_y$ is a unitary endomorphism of $L^2(G)$ for any $x, y \in G$ by proposition V.4.1.3, so we just need to show that the representation is continuous, and to check this, by proposition V.3.1.2, it suffices to show that, for every $f_0 \in L^2(G)$, the map $G \times G \to L^2(G)$, $(x, y) \longmapsto L_{x^{-1}}R_y f_0$, is continuous.

So fix $f_0 \in L^2(G)$, and let $x_0, y_0 \in G$ and $\varepsilon > 0$. Choose $f : G \to \mathbb{C}$ continuous such that $\|f_0 - f\|_2 \leq \varepsilon/3$. [5] As G is compact and f is continuous, f is uniformely continuous, so their exists a neighborhood U of 1 in G such that, for any $x, y \in U$, for any $g \in G$, $|f(x^{-1}gy) - f(g)| \leq \varepsilon/3$.

If $x \in x_0U$ and $y \in y_0U$, then

$$\|L_{x^{-1}}R_y f - L_{x_0^{-1}}R_{y_0} f\|_2 = \sqrt{\int_G |f(x^{-1}gy) - f(x_0^{-1}gy_0)|dg} \leq \varepsilon/3,$$

so

$$\begin{array}{rcl} \|L_{x^{-1}}R_y f_0 - L_{x_0^{-1}}R_{y_0} f_0\|_2 & \leq & \|L_{x^{-1}}R_y f_0 - L_{x^{-1}}R_y f\|_2 + \\ & & \|L_{x^{-1}}R_y f - L_{x_0^{-1}}R_{y_0} f\|_2 + \|L_{x_0^{-1}}R_{y_0} f - L_{x_0^{-1}}R_{y_0} f_0\|_2 \\ & \leq & \|f - f_0\|_2 + \frac{\varepsilon}{3} + \|f - f_0\|_2 \\ & \leq & \varepsilon. \end{array}$$

□

V.5 The Peter-Weyl theorem

We still assume that G is a compact Hausdorff group.

Definition V.5.1. We write $\widehat{G}$ for the set of isomorphism classes of irreducible continuous representations of G on finite-dimensional complex vector spaces. [6]

Let $(V_\rho, \rho) \in \widehat{G}$. We fix once and for all a G-invariant Hermitian inner product $\langle ., . \rangle$ on V_ρ. Note that $G \times G$ acts on $\mathrm{End}_{\mathbb{C}}(V_\rho) \simeq V_\rho^* \otimes_{\mathbb{C}} V_\rho$ [7] in the usual way,[8] that is, by $(g_1, g_2) \cdot u = \rho(g_1)u\rho(g_2)^{-1}$.

[5] See theorem 3.14 of Rudin's book [25].
[6] If G is a finite group, this was denoted by $S_{\mathbb{C}}(G)$ in chapters I and II.
[7] By proposition II.1.1.10 of chapter II.
[8] See section II.2 of chapter II.

The *Hilbert-Schmidt* inner product on $\mathrm{End}_{\mathbb{C}}(V_\rho)$ is given by the formula $\langle u, v\rangle_{HS} = \mathrm{Tr}(uv^*)$, where v^* is the adjoint of v for the chosen G-invariant inner product $\langle .,. \rangle$ on V_ρ. It's a Hermitian inner product, and we check easily that it is $G \times G$-invariant (because $\rho(g)$ is unitary for every $g \in G$).

We define $\iota_\rho : \mathrm{End}_{\mathbb{C}}(V_\rho) \to L^2(G)$ by $\iota_\rho(u)(g) = \mathrm{Tr}(\rho(g)^{-1} \circ u)$, for any $u \in \mathrm{End}_{\mathbb{C}}(V_\rho)$ and $g \in G$. For every $u \in \mathrm{End}_{\mathbb{C}}(V_\rho)$, the function $g \longmapsto \mathrm{Tr}(\rho(g)^{-1} \circ u)$ is continuous on G, hence L^2 because G is compact. So ι_ρ is well-defined.

Theorem V.5.2. *1. For every $(V_\rho, \rho) \in \widehat{G}$, $\mathrm{End}_{\mathbb{C}}(V_\rho)$ is an irreducible representation of $G \times G$. Also, if $(V_\rho, \rho) \not\simeq (V_{\rho'}, \rho')$, then $\mathrm{End}_{\mathbb{C}}(V_\rho) \not\simeq \mathrm{End}_{\mathbb{C}}(V_{\rho'})$.*

2. *For every $(V_\rho, \rho) \in \widehat{G}$, the map $\iota_\rho : \mathrm{End}_{\mathbb{C}}(V_\rho) \to L^2(G)$ is $G \times G$-equivariant, the map $\dim(V_\rho)^{1/2}\iota_\rho$ is an isometry, and $(\dim_{\mathbb{C}} V_\rho)\iota_\rho$ sends the composition in $\mathrm{End}_{\mathbb{C}}(V_\rho)$ to the convolution product in $L^2(G)$.*
3. *The map $\iota = \bigoplus_{\rho\in\widehat{G}} \iota_\rho : \bigoplus_{\rho\in\widehat{G}} \mathrm{End}_{\mathbb{C}}(V_\rho) \to L^2(G)$ is injective and has dense image.*

In other words, as a representation of $G \times G$ and as $\mathbb{C}$-algebra (without a unit),

$$L^2(G) \simeq \widehat{\bigoplus_{\rho\in\widehat{G}}} \mathrm{End}_{\mathbb{C}}(V_\rho),$$

where $\widehat{\ }$ means "completion".

Lemma V.5.3. *Let V be a Hilbert space wih a unitary representation of G, and let $V_1, V_2 \subset V$ be finite-dimensional irreducible subrepresentations of G. Then either $V_1 \simeq V_2$ as representations of G, or $\langle v_1, v_2\rangle = 0$ for any $v_1 \in V_1$ and $v_2 \in V_2$ (i.e. V_1 and V_2 are orthogonal in V).*

Proof. We know that $V_2^\perp$ is stable by G and that $V = V_2 \oplus V_2^\perp$ (remark V.3.1.5). So the orthogonal projection $\pi : V \to V_2$ is G-equivariant, hence the composition $V_1 \subset V \xrightarrow{\pi} V_2$ is G-equivariant. If $V_1 \not\simeq V_2$, then this G-equivariant map $V_1 \to V_2$ has to be 0 by Schur's lemma, which means that $V_1 \subset V_2^\perp$.

□

Proof of the theorem. Let's prove (i). Let $\rho \in \widehat{G}$. Then $\chi_{\mathrm{End}_{\mathbb{C}}(V_\rho)}(g_1, g_2) = \chi_{V_\rho}(g_1)\overline{\chi_{V_\rho}(g_2)}$, by the definition of the action of $G \times G$ on $\mathrm{End}_{\mathbb{C}}(V_\rho)$, section II.2 of chapter II and corollary II.1.4.2 of the same chapter. By the Schur orthogonality formula (corollary V.3.3.3) and corollary V.3.3.4, we have

$$\int_{G\times G} |\chi_{\mathrm{End}_{\mathbb{C}}(V_\rho)}(g_1, g_2)|^2 dg_1 dg_2 = \left(\int_G |\chi_{V_\rho}(g_1)|^2 dg_1\right)\left(\int_G |\chi_{V_\rho}(g_2)|^2 dg_2\right) = 1,$$

and so $\mathrm{End}_{\mathbb{C}}(V_\rho)$ is irreducible. Moreover, if $(V_{\rho'}, \rho')$ is not isomorphic to (V_ρ, ρ), then, by corollary V.3.3.3 again,

$$\langle \chi_{\mathrm{End}_{\mathbb{C}}(V_\rho)}, \chi_{\mathrm{End}_{\mathbb{C}}(V_{\rho'})}\rangle_{L^2(G\times G)} = \langle \chi_{V_\rho}, \chi_{V_{\rho'}}\rangle_{L^2(G)}\langle \overline{\chi_{V_\rho}}, \overline{\chi_{V_{\rho'}}}\rangle_{L^2(G)} = 0,$$

and so $\mathrm{End}_{\mathbb{C}}(V_\rho)$ and $\mathrm{End}_{\mathbb{C}}(V_{\rho'})$ are not isomorphic.

Let's prove (ii). Fix $(V_\rho, \rho) \in \widehat{G}$. First we prove that ι_ρ is $G \times G$-equivariant. Let $x, y \in G$ and $u \in \mathrm{End}_{\mathbb{C}}(V_\rho)$. Then, for every $g \in G$,

$$\iota_\rho(\rho(x)u\rho(y)^{-1})(g) = \mathrm{Tr}(\rho(g)^{-1}\rho(x)u\rho(y)^{-1}) = \mathrm{Tr}(\rho(x^{-1}gy)^{-1}u) = (L_{x^{-1}}R_y\iota_\rho(u))(g).$$

Then we prove that $(\dim_{\mathbb{C}} V_\rho)^{1/2}$ is an isometry. Let $u_1, u_2 \in \mathrm{End}_{\mathbb{C}}(V_\rho)$. We wan tto show that

$$(\dim_{\mathbb{C}} V_\rho)\langle \iota_\rho(u_1), \iota_\rho(u_2)\rangle_{L^2(G)} = \langle u_1, u_2\rangle_{HS}.$$

Remember the operators $T^0_{v,w}$ from definition V.3.2.1. By proposition V.3.2.2(ii), $\mathrm{End}_{\mathbb{C}}(V_\rho)$ is spanned by the $T^0_{v,w}$, for $v, w \in V_\rho$, so we may assume that $u_1 = T^0_{v_1,w_1}$ and $u_2 = T^0_{v_2,w_2}$, with $v_1, v_2, w_1, w_2 \in V_\rho$. Using proposition V.3.2.2(iii) and the obvious fact that $\rho(g)T^0_{v,w} = T^0_{v,gw}$ for every $g \in G$ and $v, w \in V$, we get

$$\begin{aligned}
\langle \iota_\rho(u_1), \iota_\rho(u_2)\rangle &= \int_G \mathrm{Tr}(\rho(g)^{-1}T^0_{v_1,w_1})\overline{\mathrm{Tr}(\rho(g)^{-1}T^0_{v_2,w_2})}dg \\
&= \int_G \langle g^{-1}v_1, w_1\rangle\overline{\langle g^{-1}w_2, v_2\rangle}dg \\
&= \int_G \langle gv_2, w_2\rangle\overline{\langle gv_1, w_1\rangle}dg \\
&= \frac{1}{\dim_{\mathbb{C}} V_\rho}\langle v_2, v_1\rangle\langle w_1, w_2\rangle,
\end{aligned}$$

where the last equality comes from theorem V.3.3.2. Finally, by proposition V.3.2.2(iv), this equal to $\frac{1}{\dim_{\mathbb{C}} V_\rho}\langle u_1, u_2\rangle_{HS}$.

Now we prove that $(\dim_{\mathbb{C}} V_\rho)\iota_\rho$ sends the composition in $\mathrm{End}_{\mathbb{C}}(V_\rho)$ to the convolution product in $L^2(G)$. Again, by proposition V.3.2.2(ii), we only need to check this for two elements of $\mathrm{End}_{\mathbb{C}}(V_\rho)$ of the form $u_1 = T^0_{v_1,w_1}$ and $u_2 = T^0_{v_2,w_2}$, with $v_1, v_2, w_1, w_2 \in V_\rho$. Let $g \in G$. By proposition V.3.2.2(iv) and (i), we have

$$\iota_\rho(u_1u_2)(g) = \mathrm{Tr}(\rho(g)^{-1}T^0_{v_1,w_1}{T^0_{w_2,v_2}}^*) = \langle \rho(g)^{-1}w_1, v_2\rangle\langle w_2, v_1\rangle.$$

On the other hand, by proposition V.3.2.2(iii) and theorem V.3.3.2, we have

$$\begin{aligned}
(\iota_\rho(u_1) * \iota_\rho(u_2))(g) &= \int_G \mathrm{Tr}(\rho(x)^{-1}\rho(g)^{-1}T^0_{v_1,w_1})\mathrm{Tr}(\rho(x)T^0_{v_2,w_2})dx \\
&= \int_G \langle \rho(x)^{-1}\rho(g)^{-1}w_2, v_1\rangle\langle \rho(x)w_2, v_2\rangle dx \\
&= \int_G \langle \rho(x)w_2, v_2\rangle\overline{\langle \rho(x)v_1, \rho(g)^{-1}w_1\rangle}dx \\
&= \frac{1}{\dim_{\mathbb{C}}(V_\rho)}\langle \rho(g)^{-1}w_1, v_2\rangle\langle w_2, v_1\rangle.
\end{aligned}$$

Let's prove (iii). First we prove that ι is injective. Let $u = (u_\rho)_{\rho\in\widehat{G}} \in \bigoplus_{\rho\in\widehat{G}} \mathrm{End}_{\mathbb{C}}(V_\rho)$ be such that $\iota(u) = 0$. Then $\sum_{\rho\in\widehat{G}} \iota_\rho(u_\rho) = 0$. By lemma V.5.3 and part (i), the subspaces $\iota_\rho(\mathrm{End}_{\mathbb{C}}(V_\rho))$ are pairwise orthogonal to each other, so, for every $\rho \in \widehat{G}$,

$$0 = \langle \iota_\rho(u_\rho), \iota(u)\rangle = \langle \iota_\rho(u_\rho), \iota_\rho(u_\rho)\rangle = \frac{1}{\dim_{\mathbb{C}} V_\rho}\langle u_\rho, u_\rho\rangle,$$

where the last equality comes from (ii). This implies that all the u_ρ are 0, and so $u = 0$.

It remains to show that the image of ι is dense in $L^2(G)$. First we prove that every finite-dimensional G-subrepresentation of $L^2(G)$ is contained in $\mathrm{Im}(\iota)$, where we make G act on $L^2(G)$ by the left regular action. As finite-dimensional representations of G are semisimple (corollary V.3.1.7), it suffices to show that, for every $\rho \in \widehat{G}$ and every G-equivariant map $u : V_\rho \to L^2(G)$, we have $\mathrm{Im}(u) \subset \mathrm{Im}(\iota)$. Fix $\rho \in \widehat{G}$ and $u : V_\rho \to L^2(G)$ a G-equivariant map. Let $v \in V_\rho$. Then for every $f \in L^2(G)$ and every $g \in G$,

$$\begin{aligned}
(u(v) * f)(g) &= \textstyle\int_G u(v)(gx)f(x^{-1})dx \\
&= \textstyle\int_G L_g(u(v))(x)\overline{\widetilde{f}(x)}dx \\
&= \langle L_g(u(v)), \widetilde{f}\rangle_{L^2(G)} \\
&= \langle u(\rho(g)^{-1}v), \widetilde{f}\rangle_{L^2(G)} \\
&= \langle \rho(g)^{-1}v, u^*(\widetilde{f})\rangle_{V_\rho} \\
&= \mathrm{Tr}(\rho(g)^{-1}T^0_{u^*(\widetilde{f}),v}) \\
&= \iota_\rho(T^0_{u^*(\widetilde{f}),v})(g),
\end{aligned}$$

where we used the G-equivariant of u in the 4th equality, and we are using again the operators $T^0_{v,w}$ if definition V.3.2.1. This shows that $u(v) * f \in \mathrm{Im}(\iota_\rho)$ for every $f \in L^2(G)$. By theorem V.4.2.8, we can find a sequence $(f_n)_{n\geq 0}$ in $L^2(G)$ such that $u(v) * f_n \to u(v)$ as $n \to +\infty$. This shows that $u(v) \in \overline{\mathrm{Im}(\iota_\rho)}$. As $\mathrm{Im}(\iota_\rho)$ is a finite-dimensional subspace of $L^2(G)$, it is closed in $L^2(G)$, and so $u(v) \in \mathrm{Im}(\iota_\rho) \subset \mathrm{Im}(\iota)$.

Now we show that $\mathrm{Im}(\iota)$ is dense in $L^2(G)$. This is equivalent to saying that $\mathrm{Im}(\iota)^\perp = 0$. So let $f \in \mathrm{Im}(\iota)^\perp$. By what we just proved, f is orthogonal to every finite-dimensional G-subrepresentation of $L^2(G)$. Let $h \in L^2(G)$ be such that $h = \widetilde{h}$. By proposition V.4.2.7 and theorem V.4.2.10, the endomorphism R_h of $L^2(G)$ is self-adjoint and compact. Hence, by the spectral theorem (theorem V.6.5), $\mathrm{Ker}(R_h)$ is the orthogonal of the closure of the direct sum of the eigenspaces $\mathrm{Ker}(R_h - \lambda\mathrm{id}_{L^2(G)})$, $\lambda \in \mathbb{R}^\times$, which are all finite-dimensional. As R_h is G-equivariant by proposition V.4.2.7, these eigenspaces are all stable by G, and so f is orthogonal to all of them, hence $f \in \mathrm{Ker}(R_h)$, i.e. $f * h = 0$. Now theorem V.4.2.8 gives a sequence $(h_n)_{n\geq 0}$ of elements of $L^2(G)$ such that $\widetilde{h_n} = h_n$ for every $n \geq 0$ and $f * h_n \to f$ as $n \to +\infty$. Applying what we just saw gives $f * h_n = 0$ for every $n \geq 0$, and finally $f = 0$.

□

Corollary V.5.4. *As both the left and the right regular representation of G, $L^2(G)$ is isomorphic*

to

$$\widehat{\bigoplus_{\rho\in\widehat{G}} V_\rho^{\oplus \dim_{\mathbb{C}} V_\rho}}.$$

Remark V.5.5. Let G be a finite. Then, by theorems I.3.2 and I.3.4 of chapter I, we have an isomorphism of $\mathbb{C}$-algebras

$$\mathbb{C}[G] \simeq \prod_{V\in S_{\mathbb{C}}(G)} \mathrm{End}_{\mathbb{C}}(V),$$

where $S_{\mathbb{C}}(G)$ is the set of isomorphisms classes of irreducible representations of G on $\mathbb{C}$-vector spaces. We also have, by the Peter-Weyl theorem, an isomorphism of $\mathbb{C}$-algebras

$$L^2(G) \simeq \prod_{\rho\in\widehat{G}} \mathrm{End}_{\mathbb{C}}(V_\rho).$$

Of course, $\widehat{G} = S_{\mathbb{C}}(G)$, and the two isomorphisms are related by the $\mathbb{C}$-algebra isomorphism

$$L^2(G) \xrightarrow{\sim} \mathbb{C}[G], \quad f \longmapsto \frac{1}{|G|}\sum_{g\in G} f(g)g.$$

V.6 The spectral theorem

This section contains some reminders about compact operators and the spectral theorem.

Definition V.6.1. If V_1, V_2 are two Hilbert spaces over $\mathbb{C}$ and $T : V_1 \to V_2$ is a continuous $\mathbb{C}$-linear map, we say that T is *compact* if the set $\overline{T(B)} \subset V_2$ is compact, where $B = \{v \in V_1 | \|v\| = 1\}$.

Theorem V.6.2. *Let V_1, V_2 be Hilbert spaces. We write $\mathrm{Hom}(V_1, V_2)$ for the space of continuous $\mathbb{C}$-linear maps from V_1 to V_2, and we consider the topology on it defined by the operator norm. Then, for every $T \in \mathrm{Hom}(V_1, V_2)$, the following conditions are equivalent :*

1. *T is compact.*
2. *T is a limit of finite rank elements of $\mathrm{Hom}(V_1, V_2)$.*

Lemma V.6.3. 1. *Every finite-rank operator is compact.*

2. *The space of compact operators $T : V_1 \to V_2$ is closed in $\mathrm{Hom}(V_1, V_2)$.*

Proof. Point (i) follows from the fact that the closed unit ball of a finite-dimensional $\mathbb{C}$-vector space is compact. Let's prove (ii). Let $(T_n)_{n\geq 0}$ be a sequence of compact operators in $\mathrm{Hom}(V_1, V_2)$, and suppose that it converges to $T \in \mathrm{Hom}(V_1, V_2)$. Let $(x_n)_{n\geq 0}$ be a sequence in B. We want to find a subsequence $(y_n)_{n\geq 0}$ such that $(T(y_n))_{n\geq 0}$ converges in V_2; as V_2

is complete, it suffices to find $(y_n)_{n\geq 0}$ such that $(T(y_n))_{n\geq 0}$ is a Cauchy sequence. We construct a sequence $x^{(r)} = (x_n^{(r)})_{n\geq 0}$ of subsequences of $(x_n)_{n\geq 0}$ in the following inductive way : Take $x^{(0)} = (x_n)_{n\geq 0}$. If $r \geq 1$, suppose $x^{(r-1)}$ constructed, and take for $x^{(r)}$ a subsequence of $x^{(r-1)}$ such that $(T_r(x_n^{(r)}))_{n\geq 0}$ is a Cauchy sequence. (This is possible because T_r is a compact operator.) Finally, set $y_n = x_n^{(n)}$. Let's show that $(T(y_n))_{n\geq 0}$ is a Cauchy sequence. Let $\varepsilon > 0$. Choose $m \geq 1$ such that $\|T - T_m\| \leq \varepsilon$, and $N \geq m$ such that, for any $r, s \geq N$, $\|T_m(x_r^{(r)}) - T_m(x_s^{(s)})\| \leq \varepsilon$ (this is possible because $(x_n^{(r)})_{n\geq 0}$ is a subsequence of $(x_n^{(m)})_{n\geq 0}$ for every $r \geq m$). Then, if $r, s \geq N$, we have

$$\|T(y_r) - T(y_s)\| \leq \|(T - T_m)(y_r)\| + \|T_m(y_r - y_s)\| + \|(T_m - T)(y_s)\| \leq 3\varepsilon,$$

because $y_s, y_r \in B$.

□

Proof of the theorem. The implication (ii)⇒(i) follows directly from the lemma. Let's show that (ii) implies (i). Let $T \in \mathrm{Hom}(V_1, V_2)$ be a compact operator. Let $(W_n)_{n\geq 0}$ be a sequence of finite-dimensional subspace of V_2 such that $V_2 = \bigcup_{n\geq 0} W_n$, and, for every $n \geq 0$, let $\pi_n : V_2 \to W_n$ be the orthogonal projection. Set $T_n = \pi_n \circ T$. Then, for every $x \in V_1$, $T_n(x) \to T(y)$ as $n \to +\infty$. Let's show that we actually have $T_n \to T$ in $\mathrm{Hom}(V_1, V_2)$. Let $\varepsilon > 0$. As $\overline{T(B)}$ is compact, we can find $x_1, \ldots, x_r \in B$ such that, for every $y \in B$, there exists $i \in \{1, \ldots, r\}$ such that $\|T(x_i) - T(y)\| \leq \varepsilon$; note that, for every $n \geq 0$, we then have

$$\|T_n(y) - T_n(x_i)\| = \|\pi_n(T(y) - T(x_i))\| \leq \|T(y) - T(x_i)\| \leq \varepsilon.$$

Choose $N \geq 0$ such that, for every $i \in \{1, \ldots, r\}$ and every $n \geq N$, $\|T_n(x_i) - T(x_i)\| \leq \varepsilon$. Let's show that, for $n \geq N$, we have $\|T_n - T\| \leq 3\varepsilon$. Fix $n \geq N$. If $y \in B$, choose $i \in \{1, \ldots, r\}$ such that $\|T(y) - T(x_i)\| \leq \varepsilon$. Then we have

$$\|T(y) - T_n(y)\| \leq \|T(y) - T(x_i)\| + \|T(x_i) - T_n(x_i)\| + \|T_n(x_i) - T_n(y)\| \leq 3\varepsilon.$$

□

Note that the proof of (ii)⇒(i) in the theorem above still works of V_1 and V_2 are general Banach space, but the proof of (i)⇒(ii) does not. (And in fact, it is not true that a compact operator between Banach spaces is always the limit of a sequence of finite rank operators. See Enflo's paper [11] for a counterexample.)

Definition V.6.4. Let V be a Hilbert space, and let T be a continuous endomorphism of V. We say that $\lambda \in \mathbb{C}$ is an *eigenvalue* of T if $\mathrm{Ker}(T - \lambda \mathrm{id}_V) \neq \{0\}$. We denote by $\mathrm{Spec}(T) \subset \mathbb{C}$ the set of eigenvalues of T, and call it the *spectrum* of T.

Theorem V.6.5 (Spectral theorem)**.** *Let V be a Hilbert space over $\mathbb{C}$, and let $T : V \to V$ be a continuous endomorphism of V. Assume that T is compact and self-adjoint, and write $V_\lambda = \mathrm{Ker}(T - \lambda \mathrm{id}_V)$ for every $\lambda \in \mathbb{C}$.*

Then :

1. $\mathrm{Spec}(T) \subset \mathbb{R}$.
2. *If* $\lambda, \mu \in \mathrm{Spec}(T)$ *and* $\lambda \neq \mu$, *then* $V_\mu \subset V_\lambda^\perp$.
3. *If* $\lambda \in \mathrm{Spec}(T) - \{0\}$, *then* $\dim_{\mathbb{C}} V_\lambda < +\infty$.
4. $\mathrm{Spec}(T)$ *is finite or countable, and its only possible limit point is* 0.
5. $\bigoplus_{\lambda \in \mathrm{Spec}(T)} V_\lambda$ *is dense in* V.

Proof. Points (i) and (ii) follow from the fact that T is self-adjoint just as in the finite-dimensional case.

Let $r > 0$. Let $W = \bigoplus_{|\lambda| \geq r} V_\lambda$. Choose a Hilbert basis $(e_i)_{i \in I}$ of W made up of eigenvectors of T, i.e. such that, for every $i \in I$, we have $T(e_i) = \lambda_i e_i$ with $|\lambda_i| \geq r$. If I is infinite, then the family $(T(e_i))_{i \in I}$ cannot have a convergent (non-stationary) subsequence. Indeed, if we had an injective map $\mathbb{N} \to I$, $n \longmapsto i_n$, such that $(T(e_{i_n}))_{n \geq 0}$ converges to some vector v of V, then $\lambda_{i_n} e_{i_n} \to v$, so v is in the closure of $\mathrm{Span}(e_{i_n}, n \geq 0)$. But on the other hand, for every $n \geq 0$, $\langle v, e_{i_n} \rangle = \lim_{m \to +\infty} \langle \lambda_{i_m} e_{i_m}, e_{i_n} \rangle = 0$, so $v \in \mathrm{Span}(e_{i_n}, n \geq 0)^\perp$. This forces $v = 0$. But $\|v\| = \lim_{n \to +\infty} \|\lambda_{i_n} e_{i_n}\| \geq r > 0$, contradiction.

As T is compact, this show that I cannot be infinite, and gives (iii) and (iv).

Let's prove (v). Let $W' = \bigoplus_{\lambda \in \mathrm{Spec}(T)} V_\lambda$, and $W = W'^\perp$. We want to show that $W = 0$. So suppose that $W \neq 0$. As T is self-adjoint and W' is clearly stable by T, $T(W) \subset W$. (If $v \in W$, then for every $w \in W'$, $\langle T(v), w \rangle = \langle v, T(w) \rangle = 0$.) Note that $T_{|W}$ has no eigenvalue, so in particular $\mathrm{Ker}(T) = \{0\}$, hence $\|T_{|W}\| > 0$. Let $B = \{x \in W | \|x\| = 1\}$. As $\|T_{|W}\| = \sup_{x \in B} |\langle T(x), x \rangle|$, there exists a sequence $(x_n)_{n \geq 0}$ of elements of B such that $\langle T(x_n), x_n \rangle \to \lambda$ as $n \to +\infty$, where $\lambda = \pm \|T_{|W}\|$. Then

$$0 \leq \|T(x_n) - \lambda x_n\|^2 = \|T(x_n)\|^2 + \lambda^2 \|x_n\|^2 - 2\lambda \langle T(x_n), x_n \rangle \leq 2\lambda^2 - 2\lambda \langle T(x_n), x_n \rangle$$

converges to 0 as $n \to +\infty$, so $T(x_n) - \lambda x_n$ itself converges to 0. As T is compact, we may assume that the sequence $(T(x_n))_{n \geq 0}$ has a limit in W, say w. Then $T(w) - \lambda w = 0$. But T has no eigenvalue in W, so $w = 0$. But then $T(x_n) \to 0$, so $\langle T(x_n), x_n \rangle \to 0$, so $\lambda = 0 = \|T_{|W}\|$, a contradiction.

□

VI Representations of Lie algebras : the case of $\mathfrak{sl}_n(\mathbb{C})$

The goal of this chapter is to explain the representation theory of the semisimple complex Lie algebra $\mathfrak{sl}_n(\mathbb{C})$, and its relation to the representation theory of the compact group $\mathrm{SU}(n)$. We will start with a few general definitions and quickly specialize.

VI.1 Definitions

Let k be a commutative ring with unit. In this chapter, by an associative k-algebra, we will always mean an associative k-algebra with unit unless otherwise specified.

Definition VI.1.1. A *Lie algebra over* k (or k*-Lie algebra*) is a k-module $\mathfrak{g}$ with a k-bilinear map $[.,.] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfying the following conditions /

1. for every $X \in \mathfrak{g}$, $[X, X] = 0$;
2. (Jacobi identity) for all $X, Y, Z \in \mathfrak{g}$, $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$.

The operation $[.,.]$ is called the *Lie bracket* of $\mathfrak{g}$.

Remark VI.1.2. If we apply condition (i) to $X + Y$, we get $[X, Y] + [Y, X] = 0$, i.e. : (i') $[X, Y] = -[X, Y]$, for any $X, Y \in \mathfrak{g}$. This condition is equivalent to (i) if 2 is invertible in k, but not in general.

Definition VI.1.3. Let $\mathfrak{g}$ be a Lie algebra over k.

1. A (k-)Lie subalgebra of $\mathfrak{g}$ is a k-submodule $\mathfrak{h}$ of $\mathfrak{g}$ such that $[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}$.
2. An *ideal* of $\mathfrak{g}$ is a k-submodule $\mathfrak{a}$ such that $[\mathfrak{g}, \mathfrak{a}] \subset \mathfrak{a}$.
3. If $\mathfrak{h}$ is another Lie algebra over k, a *morphism of (*K*-)Lie algebras* is a k-linear map $u : \mathfrak{g} \to \mathfrak{h}$ such that $u([X, Y]) = [u(X), u(Y)]$ for all $x, Y \in \mathfrak{g}$.

Remark VI.1.4. If $\mathfrak{g}$ is a k-Lie algebra and $\mathfrak{a}$ is an ideal of $\mathfrak{g}$, then condition (i') implies that $[\mathfrak{a}, \mathfrak{g}] = [\mathfrak{g}, \mathfrak{a}] = \subset \mathfrak{a}$. So for Lie algebras, there is not distinction between left and right ideals.

The following proposition is obvious.

Proposition VI.1.5. *1. If $\mathfrak{g}$ is a k-Lie algebra and $\mathfrak{a}$ is an ideal of $\mathfrak{g}$, then the Lie bracket of $\mathfrak{g}$ defines a bracket on the quotient k-module $\mathfrak{g}/\mathfrak{a}$, and this makes $\mathfrak{g}/\mathfrak{a}$ into a k-Lie algebra.*

2. If $u : \mathfrak{g} \to \mathfrak{h}$ is a morphism of k-Lie algebras, then $\operatorname{Ker} u$ is an ideal of $\mathfrak{g}$, $\operatorname{Im} u$ is a Lie subalgebra of $\mathfrak{h}$, and u induces an isomorphism of Lie algberas $\mathfrak{g}/\operatorname{Ker} u \xrightarrow{\sim} \operatorname{Im} u$.

Example VI.1.6.

- Any k-module V becomes a Lie algebra with the trivial Lie bracket $[.,.] = 0$. Such a Lie algebra is called a *commutative Lie algebra*.
- If A is an associative k-algebra (with or without unit), we define the *commutator bracket* $[.,.] : A \times A \to A$ by $[a, b] = ab - ba$, for every $a, b \in A$. This makes A into a Lie algebra (condition (i), condition (ii) follows from the associativity of A).

  Note that the Lie algebra $(A, [.,.])$ is commutative if and only if the associative algebra A is commutative.
- Let V be a k-module. Then the associative algebra $\operatorname{End}_k(V)$ together with its commutator bracket is a Lie algebra, which will be denoted by $\mathfrak{gl}_k(V)$ (or $\mathfrak{gl}(V)$ if k is clear from the context).

  If $V = k^n$, we write $\mathfrak{gl}_n(k) = \mathfrak{gl}_n(V)$.
- Here are some Lie subalgebras of $\mathfrak{gl}_n(k)$:

$$\mathfrak{sl}_n(k) = \{X \in \mathfrak{gl}_n(k) | \operatorname{Tr}(X) = 0\}$$

$$\mathfrak{b}_n(k) = \{X \in \mathfrak{gl}_n(k) | X \text{ is upper triangular}\}$$

$$\mathfrak{u}_n(k) = \{X \in \mathfrak{gl}_n(k) | X \text{ is strictly upper triangular}\}$$

$$\mathfrak{t}_n(k) = \{X \in \mathfrak{gl}_n(k) | X \text{ is diagonal}\}$$

$$\mathfrak{o}_n(k) = \{X \in \mathfrak{gl}_n(k) | X + {}^tX = 0\}$$

  Note that $\mathfrak{u}_n(K)$ is an ideal of $\mathfrak{b}_n(k)$, and that the quotient $\mathfrak{b}_n(k)/\mathfrak{u}_n(k)$ is isomorphic to the commutative Lie algebra $\mathfrak{t}_n(k)$.

Definition VI.1.7. A *representation* of a k-Lie algebra $\mathfrak{g}$ on a k-module V is a morphism of Lie algebras $u : \mathfrak{g} \to \mathfrak{gl}(V)$. We use V, u of (V, u) to refer to the representation. Sometimes we omit u from the notation and write Xv for $u(X)(v)$ ($X \in \mathfrak{g}$, $v \in V$).

Moreover :

- We say that the representation is *faithful* if u is injective.
- A *subrepresentation* of (V, u) is a submodule W of V such that for every $X \in \mathfrak{g}$, $u(X)(W) \subset W$.
- The representation (V, u) is called *irreducible* if $V \neq 0$ and if the only subrepresentations of V are 0 and V.

- The representation (V, u) is called *semisimple* if for every subrepresentation W of V, there exists another subrepresentation W' of V such that $V = W \oplus W'$.

Remark VI.1.8. By using a Zorn lemma argument as in theorem I.1.3.4 of chapter I, we could show that a representation of $\mathfrak{g}$ is semisimple if and only if it is a sum of irreducible subrepresentations. [1] But in the case of most interest to us, which is the case where k is a field and both $\mathfrak{g}$ and V are finite-dimensional over k, we can prove this fact by an easy induction on $\dim_k V$ and so we don't need Zorn's lemma.

Example VI.1.9.

- The *trivial representation* of $\mathfrak{g}$ is the map $u = 0 : \mathfrak{g} \to k$.
- $\mathrm{Tr} : \mathfrak{gl}_n(k) \to k$ is a nontrivial representation of $\mathfrak{gl}_n(k)$ on k.
- The adjoint representation : Consider the map $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ sending X to the endomorphism $Y \longmapsto [X, Y]$ of $\mathfrak{g}$. Then this is a map of Lie algebras, i.e. a representation of $\mathfrak{g}$ on itself, and we call it the *adjoint representation*.

  Indeed, if $X_1, X_2, Y \in \mathfrak{g}$, then

$$\mathrm{ad}([X_1, X_2])(Y) = [[X_1, X_2], Y] = -[Y, [X_1, X_2]],$$

  while

$$[\mathrm{ad}(X_1), \mathrm{ad}(X_2)](Y) = [X_1, [X_2, Y]] - [X_2, [X_1, Y]] = [X_1, [X_2, Y]] + [X_2, [Y, X_1]],$$

  and these are equal by the Jacobi identity (condition (ii) of definition VI.1.1).

Remark VI.1.10. The adjoint representation has many interesting properties. For example :

(1) For every $X \in \mathfrak{g}$, $\mathrm{ad}(X)$ is actually a *derivation* of $\mathfrak{g}$. That is, $\mathrm{ad}(X)([Y, Z]) = [Y, \mathrm{ad}(X)(Z)] + [\mathrm{ad}(X)(Y), Z]$ for all $Y, Z \in \mathfrak{g}$. (This is just a reformulation of the Jacobi identity.)

(2) The set of derivations of $\mathfrak{g}$ is a Lie subalgebra of $\mathfrak{gl}(\mathfrak{g})$, and the image of ad is an ideal of this Lie subalgebra. [2]

Remark VI.1.11. If k is a field of characteristic 0, then Ado's theorem[3] says that every finite-dimensional Lie algebra over k admits a faithful representation on a finite-dimensional k-vector space.

[1] In fact, once we see in corollary VI.2.2.3 that representations of $\mathfrak{g}$ are the same as modules over its universal enveloping algebra, we can just apply theorem I.1.3.4 of chapter I directly.

[2] See problem VII.6.4.

[3] See Ado's paper [1].

VI.2 Universal enveloping algebra

VI.2.1 The tensor algebra of a k-module

Let V be a k-module.

Definition VI.2.1.1. The *nth tensor power* of V (over k) is

$$T^nV = \underbrace{V \otimes_k \cdots \otimes_k V}_{n \text{ times}},$$

with the convention that $T^0V = k$.

The *tensor algebra* of V is

$$T^*V = \bigoplus_{n\geq 0} T^nV.$$

The k-linear maps $T^nV \otimes_k T^mV \to T^{n+m}V$ sending $(v_1 \otimes \cdots \otimes v_n) \otimes (w_1 \otimes \cdots \otimes w_n)$ to $(v_1 \otimes \cdots \otimes v_n \otimes w_1 \otimes \cdots \otimes w_n)$ gives T^*V the structure of a (graded) associative k-algebra, with unit $1 \in k = T^0V$.

We denote by ι the obvious k-module inclusion $V = T^1V \hookrightarrow T^*V$.

These objects satisfy the following universal properties.

Proposition VI.2.1.2. *1. Let W be another k-module. Then we have a canonical bijection*

$$\mathrm{Hom}_k(T^nV, W) = \{n\text{-linear maps } V^n \to W\}.$$

2. Let A be an associative k-algebra (with unit). Then the map

$$\mathrm{Hom}_{k-\mathrm{algebras}}(T^*V, A) \to \mathrm{Hom}_{k-\mathrm{modules}}(V, A)$$

sending $\varphi : T^*V \to A$ to $\varphi\iota$ is a bijection.

Proof. Point (i) is just the universal property of the tensor product.

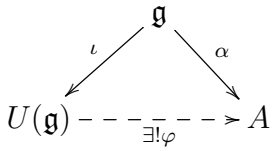
Let's prove point (ii). First, let $\varphi_1, \varphi_2 : T^*V \to A$ be two k-algebras maps such that $\varphi_1\iota = \varphi_2\iota$. As T^1V generates the k-algebra T^*V, this implies $\varphi_1 = \varphi_2$. Next, let $u : V \to A$ be a map of k-modules. For every $n \geq 1$, the k-multilinear map $V^n \to A$, $(v_1, \ldots, v_n) \longmapsto u(v_1) \ldots u(v_n)$ induces a k-linear map $varphi_n : T^nV \to A$. We also write $\varphi_0 : k \to A$ for the structural map. Then $\varphi := \bigoplus_{n\geq 0} \varphi_n : T^*V \to A$ is a k-algebra map (by definition of the product on T^*V), and $\varphi\iota = u$.

□

VI.2.2 The universal enveloping algebra of a Lie algebra

Definition VI.2.2.1. Let $\mathfrak{g}$ be a Lie algebra over k. A *universal enveloping algebra* of $\mathfrak{g}$ is a pair $(\iota, U(\mathfrak{g}))$, where $U(\mathfrak{g})$ is an associative k-algebra with unit and $\iota : \mathfrak{g} \to U(\mathfrak{g})$ is a morphism of Lie algebras from $\mathfrak{g}$ to $U(\mathfrak{g})$ with its commutator bracket, such that :

For every other pair (α, A) with A an associative k-algebra with unit and $\alpha : \mathfrak{g} \to A$ a morphism of Lie algebras from $\mathfrak{g}$ to A with its commutator bracket, there exists a unique morphism of k-algebras $\varphi : U(\mathfrak{g}) \to A$ such that $\alpha = \varphi\iota$.



Theorem VI.2.2.2. *Let $\mathfrak{g}$ be a Lie algebra over k. Then :*

1. *A universal enveloping algebra of $\mathfrak{g}$ exists.*
2. *If $(\iota_1, U_1(\mathfrak{g}))$ and $(\iota_2, U_2(\mathfrak{g}))$ are two universal enveloping algebras of $\mathfrak{g}$, then there exists a unique isomorphism of k-algebras $\varphi : U_1(\mathfrak{g}) \to U_2(\mathfrak{g})$ such that $\varphi\iota_1 = \iota_2$.*

Moreover, if $(\iota, U(\mathfrak{g}))$ is a universal enveloping algebra of $\mathfrak{g}$,then $\iota(\mathfrak{g})$ generates $U(\mathfrak{g})$ as a k-algebra.

Because of point (ii), we usually talk about the universal enveloping algebra of $\mathfrak{g}$. We also often omit ι from the notation.

Proof. Let's prove (ii). With the notation of the theorem, there exist a unique morphisms of k-algebras $\varphi : U_1(\mathfrak{g}) \to U_2(\mathfrak{g})$ and $\psi : U_2(\mathfrak{g}) \to U_2(\mathfrak{g})$ such that $\varphi\iota_1 = \iota_2$ and $\psi\iota_2 = \iota_1$. We just need to show that they isomorphisms. But $\varphi\psi$ and $\mathrm{id}_{U_2(\mathfrak{g})}$ are two morphisms of k-algebra satisfying $\varphi\psi\iota_2 = \mathrm{id}_{U_2(\mathfrak{g})}\iota_2 = \iota_2$, so $\varphi\psi = \mathrm{id}_{U_2(\mathfrak{g})}$. Similarly, $\psi\varphi = \mathrm{id}_{U_1(\mathfrak{g})}$.

Let's prove (i). We have defined in definition VI.2.1.1 the tensor algebra $T^*\mathfrak{g}$ of $\mathfrak{g}$ (seen as a k-module) together with a k-linear map $\iota : \mathfrak{g} \to T^*\mathfrak{g}$, and, by proposition VI.2.1.2, composition with ι induces a bijection, for every associative k-algebra with unit A,

$$\mathrm{Hom}_{k-\mathrm{algebras}}(T^*\mathfrak{g}, A) \to \mathrm{Hom}_{k-\mathrm{modules}}(\mathfrak{g}, A).$$

Let I be the two-sided ideal of $T^* * \mathfrak{g}$ generated by the elements $X \otimes Y - Y \otimes X - [X, Y]$, for $X, Y \in T^1\mathfrak{g} = \mathfrak{g}$. Take $U(\mathfrak{g}) = T^*\mathfrak{g}/I$. We still write ι for the composition $\mathfrak{g} \xrightarrow{\iota} T^*\mathfrak{g} \twoheadrightarrow U(\mathfrak{g})$. Let A be an associative k-algebra with unit. Then, for every k-module map $u : \mathfrak{g} \to A$, if $\varphi : T^*\mathfrak{g} \to A$ is the corresponding k-algebra map, we have

$$\begin{aligned} u \text{ is a Lie algebra map} &\Leftrightarrow \forall X, Y \in \mathfrak{g},\ u([X, Y]) = u(X)u(Y) - u(Y)u(X) \\ &\Leftrightarrow \forall X, Y \in \mathfrak{g},\ \varphi([X, Y]) = \varphi(X)\varphi(Y) - \varphi(Y)\varphi(X) \\ &\Leftrightarrow \varphi(I) = 0 \\ &\Leftrightarrow \varphi \text{ induces a } k\text{-algebra morphism } U(\mathfrak{g}) \to A. \end{aligned}$$

In other words, composition with ι induces a bijection

$$\mathrm{Hom}_{k-\mathrm{Liealgebras}}(U(\mathfrak{g}), A) \to \mathrm{Hom}_{k-\mathrm{modules}}(\mathfrak{g}, A).$$

So $(\iota, U(\mathfrak{g}))$ is a universal enveloping algebra of $\mathfrak{g}$.

The last sentence follows directly from the construction of $U(\mathfrak{g})$.

□

Corollary VI.2.2.3. *Let $\mathfrak{g}$ is a Lie algebra over k, and let $(\iota, U(\mathfrak{g}))$ be its universal enveloping algebra.*

1. *For every k-module V, the map*

$$\mathrm{Hom}_{k-\mathrm{algebras}}(U(\mathfrak{g}), \mathrm{End}_k(V)) \to \mathrm{Hom}_{k-\mathrm{Lie\ algebras}}(\mathfrak{g}, \mathfrak{gl}_k(V))$$

   sending φ to $\varphi \circ \iota$ is a bijection.

   In other words, giving a representation of $\mathfrak{g}$ on V is the same as giving a $U(\mathfrak{g})$-module structure on V (compatible with the k-module structure).

2. *If V and W are representations of $\mathfrak{g}$, hence also $U(\mathfrak{g})$-modules by (i), and if $u : V \to W$ is a k-linear map, then u is a morphism of representations if and only if it is $U(\mathfrak{g})$-linear.*

We can reformulate this corollary informally by saying that representations of $\mathfrak{g}$ are "the same" as $U(\mathfrak{g})$-modules. [4] So all the general results of section I.1 of chapter I apply to representations of Lie algebras.

Proof. 1. This is just the universal property of $U(\mathfrak{g})$.

2. If u is $U(\mathfrak{g})$-linear, then it is a morphism of representations because $\mathfrak{g}$ acts on V and W through $\iota : \mathfrak{g} \to U(\mathfrak{g})$.

   Conversely, suppose that u is a morphism of representations. Then it is linear under the k-subalgebra of $U(\mathfrak{g})$ generated by $\iota(\mathfrak{g})$. But that subalgebra is equal to $U(\mathfrak{g})$ itself.

□

VI.3 The matrix exponential

From now on, we will mostly specialize to the case $k = \mathbb{C}$.

Fix an integer $n \geq 1$, put the usual Euclidian norm $\|.\|$, on $\mathbb{C}^n$ and denote by $\|.\|$ the corresponding operator norm on $M_n(\mathbb{C})$. That is, for every $A \in M_n(\mathbb{C})$, we have

$$\|A\| = \sup_{X \in \mathbb{C}^n,\ \|X\|=1} \|AX\|.$$

[4] In more precise terms, we have an equivalence of categories between the two.

Then we clearly have $\|AB\| \leq \|A\|\|B\|$, for any $A, B \in M_n(\mathbb{C})$.

Definition VI.3.1. The *matrix exponential* is the map $\exp : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ defined by

$$\exp(A) = e^A = \sum_{r \geq 0} \frac{1}{r!} A^r.$$

Proposition VI.3.2. *1. The series defining* $\exp(A)$ *converges absolutely for every* $A \in M_n(\mathbb{C})$*, and* $\exp$ *is a* C^∞ *map.*[5]

2. *For every* $A \in M_n(\mathbb{C})$*, we have*

$$\frac{d}{dt} e^{tA} = Ae^{tA} = e^{tA} A.$$

In particular, the differential of $\exp$ *at* 0 *is given by*

$$d\exp_0 = \mathrm{id}_{M_n(\mathbb{C})}.$$

3. *If* $A, B \in M_n(\mathbb{C})$ *commute, then* $e^{A+B} = e^A e^B = e^B e^A$.

In particular, for every $A \in M_n(\mathbb{C})$*,* $e^A e^{-A} = I_n$*, so* $e^A \in \mathrm{GL}_n(\mathbb{C})$.

4. *For every* $A \in M_n(\mathbb{C})$ *and* $S \in \mathrm{GL}_n(\mathbb{C})$*, we have* $e^{SAS^{-1}} = Se^A S^{-1}$*,* $e^{{}^t A} = {}^t e^A$ *and* $e^{A^*} = (e^A)^*$.

5. *For every* $A \in M_n(\mathbb{C})$*,* $\det(e^A) = e^{\mathrm{Tr} A}$.

Proof. 1. Let $A \in M_n(\mathbb{C})$ and every $r \geq 0$, we have $\|A^r\| \leq \|A\|^r$. So the series defining $\exp$ converges absolutely on every closed ball of $M_n(\mathbb{C})$ centered at 0. This implies that it defines a C^∞ map.

2. Thanks to (i), we can calculate the derivative of $e^{tA} = \sum_{r \geq 0} \frac{t^r}{r!} A^r$ term by term, and then the result is obvious.

   To deduce the formula for $d\exp_0$, note that, for every $A \in M_n(\mathbb{C})$,

$$d\exp_0(A) = \frac{d}{dt} e^{tA}\Big|_{t=0}.$$

3. This is proved just like the similar formula for real (or complex) numbers.

4. This is obvious.

[5] It is even complex analytic.

5. By (iv), we may assume that A is upper triangular. Let $\lambda_1, \ldots, \lambda_n$ be the diagonal entries of A. Then e^A is upper triangular with diagonal entries $e^{\lambda_1}, \ldots, e^{\lambda_n}$, which gives the result.

□

We will need a generalization of point (ii) for later use.

Proposition VI.3.3. *Let $X \in M_n(\mathbb{C})$. Then $d\exp_X$ (the differential of the matrix exponential at X) is given by*

$$d\exp_X(A) = \int_0^1 e^{sX} A e^{(1-s)X} ds$$

for every $A \in M_n(\mathbb{C})$, and $d\exp_X$ is invertible (as a linear map from $M_n(\mathbb{C})$ to itself).

Proof. Let's first prove the formula for $d\exp_X$. Let $A \in M_n(\mathbb{C})$. By definition, we have

$$d\exp_X(A) = \lim_{t\to 0} \frac{1}{t}(e^{X+tA} - e^X).$$

For every $k \geq 0$, define an endomorphism L_k of $M_n(\mathbb{C})$ by

$$L_k(B) = \sum_{i=0}^{k} X^i B X^{k-i}.$$

Then, for every $k \geq 0$,

$$(X + tA)^{k+1} = X + tL_k(A) + O(t^2).$$

So

$$e^{X+tA} = e^X + t\sum_{k\geq 0} \frac{1}{(k+1)!} L_k(A) + O(t^2),$$

which gives

$$d\exp_X(A) = \sum_{k\geq 0} \frac{1}{(k+1)!} L_k(A).$$

On the other hand, using the (easy) fact that

$$\int_0^1 s^{k_1} s^{k_2} ds = \frac{k_1! k_2!}{(k_1 + k_2 + 1)!}$$

for all $k_1, k_2 \geq 0$, we see that

$$\begin{aligned} \int_0^1 e^{sX} A e^{(1-s)X} ds &= \sum_{k_1,k_2\geq 0} \frac{1}{k_1! k_2!} \int_0^1 s^{k_1} X^{k_1} A (1-s)^{k_2} X^{k_2} ds \\ &= \sum_{k_1,k_2\geq 0} \frac{1}{(k_1 + k_2 + 1)!} X^{k_1} A X^{k_2}, \end{aligned}$$

and this is clearly equal to $d\exp_X(A)$.

Now let's show that $d\exp_X$ is invertible. We write $X = Y + N$, with Y diagonalizable, N nilpotent and $YN = NY$. Then, by the first part,

$$d\exp_X = \sum_{k,l\geq 0} \frac{1}{k!l!} T_{k,l},$$

where

$$T_{k,l}(A) = N^k (\int_0^1 s^k (1-s)^l e^{sY} A e^{(1-s)Y} ds) N^l.$$

Using the fact that Y and N commute (and Fubini's theorem), we see that the operators $T_{k,l}$ commute with each other. Because N is nilpotent, the map $T_{k,l}$ is nilpotent as soon as $k \geq 1$ or $l \geq 1$. Putting these two facts together, we see that $d\exp_X$ is invertible if and only if $T_{0,0}$ is invertible. As $T_{0,0} = d\exp_Y$, this means that are reduced to the case where X is diagonalizable.

So let's assume that X is diagonalizable and prove that $d\exp_X$ is invertible. As $d\exp_{SXS^{-1}} = Sd\exp_X S^{-1}$, we may assume that X is diagonal. Let $x_1, \ldots, x_n$ be its diagonal entries. Let $A \in M_n(\mathbb{C})$, and write $A = (a_{ij})$. Then the (i,j)th entry of $d\exp_X(A)$ is

$$\int_0^1 e^{sx_i} a_{ij} e^{(1-s)x_j} ds = \begin{cases} a_{ij} e^{x_j} & \text{if } x_i = x_j \\ a_{ij} \frac{e^{x_j}}{x_i - x_j}(e-1) & \text{otherwise.} \end{cases}$$

So, if $d\exp_X(A) = 0$, then $A = 0$. As $d\exp_X$ is an endomorphism of $M_n(\mathbb{C})$, this suffices to prove that it is invertible.

□

The matrix logarithm : For $B \in M_n(\mathbb{C})$, if $\|B\| < 1$ or B is nilpotent, let

$$\log(I_n + B) = \sum_{r\geq 1} (-1)^{r-1} \frac{1}{r} B^r.$$

This series converges absolutely in every closed ball of $\{B \in M_n(\mathbb{C}) | \|B\| < 1\}$, hence defines a C^∞ function $\log : \mathscr{B} \to M_n(\mathbb{C})$, where $\mathscr{B} = \{A \in M_n(\mathbb{C}) | \|A - I_n\| < 1\}$.

Proposition VI.3.4. *For any $A \in M_n(\mathbb{C})$ such that $\log(A)$ is defined, we have $\exp(\log(A)) = A$.*

Note however that $\log(\exp(A)) = A$ does not hold in general, even if $n = 1$, simply because there are many matrices A such that $\exp(A) = I_n$.

Proof. Consider the formal power series $f(t) = \sum_{r\geq 0} \frac{t^r}{r!}$ and $g(t) = \sum_{r\geq 1} (-1)^{r-1} \frac{t^r}{r}$ in $\mathbb{C}[[t]]$. Then we have $f(g(t)) = t$ in $\mathbb{C}[[t]]$, because $f(g(t)) = t$ because this equality holds for any $t \in \mathbb{C}$ such that $|t| = 1$.

If $A \in M_n(\mathbb{C})$, we can deduce from this that $\exp(\log(A)) = A$ as long as all the double series $\log(\exp(A))$ converges absolutely. If $A = I_n + B$ with B nilpotent, then this double series is a finite sum. If $A = I_n + B$ with $\|B\| < 1$, then the proof is the same as in the case $n = 1$, using the fact that $\|B^r\| \leq \|B\|^r$ for every ≥ 0.

□

VI.4 The Lie algebra of a closed subgroup of $\mathrm{GL}_n(\mathbb{C})$

The open subset $\mathrm{GL}_n(\mathbb{C})$ of $M_n(\mathbb{C})$ (with the induced topology) is a topological group. We will show how to associate a Lie algebra to any closed subgroup G of $\mathrm{GL}_n(\mathbb{C})$. What is really going is that G has a canonical Lie group structure, and the Lie algebra of G will be the tangent space of G at 1. But we can prove this directly, without knowing anything about Lie groups or manifolds.

Definition VI.4.1. Let G be a closed subgroup of $\mathrm{GL}_n(\mathbb{C})$. Then the Lie algebra $\mathrm{Lie}(G)$ of G is the set of $X \in \mathfrak{gl}_n(\mathbb{C})(= M_n(\mathbb{C}))$ such that there exists a C^∞ function $c : \mathbb{R} \to M_n(\mathbb{C})$ with $c(\mathbb{R}) \subset G$, $c(0) = I_n$ and $c'(0) = X$.

Note that $\mathrm{Lie}(\mathrm{GL}_n(\mathbb{C})) = \mathfrak{gl}_n(\mathbb{C})$.

Theorem VI.4.2. *Let G be a closed subgroup of $\mathrm{GL}_n(\mathbb{C})$. Then $\mathrm{Lie}(G)$ is a $\mathbb{R}$-Lie subalgebra of $\mathfrak{gl}_n(\mathbb{C})$ and, for every $g \in G$ and $X \in \mathrm{Lie}(X)$, $gXg^{-1} \in \mathrm{Lie}(G)$.*

Proof. Let $X, Y \in \mathrm{Lie}(G)$, and let $c_X, c_Y : \mathbb{R} \to M_n(\mathbb{C})$ be C^∞ functions such that $c_X(\mathbb{R}), c_Y(\mathbb{R}) \subset G$, $c_X(0) = c_Y(0) = I_n$ and $c'_X(0) = X$, $c'_Y(0) = Y$. Let $\lambda \in \mathbb{R}$ and $g \in G$. Then :

(1) Consider $c_1 : \mathbb{R} \to M_n(\mathbb{C})$, $t \longmapsto c_X(\lambda t)$. Then c is C^∞, $c_1(\mathbb{R}) = c_X(\mathbb{R}) \subset G$, $c_1(0) = c_X(0) = I_n$, and $c'_1(0) = \lambda c'_X(0) = \lambda X$. So $\lambda X \in G$.

(2) Consider $c_2 : \mathbb{R} \to M_n(\mathbb{C})$, $t \longmapsto c_X(t)c_Y(t)$. Then c_2 is C^∞, $c_2(\mathbb{R}) =\subset G$ (because G is a subgroup of $\mathrm{GL}_n(\mathbb{C})$), $c_2(0) = c_X(0)c_Y(0) = I_n$, and $c'_2(0) = c'_X(0)c_Y(0) + c_X(0)c'_Y(0) = X + Y$. So $X + Y \in G$.

(3) Consider $c_3 : \mathbb{R} \to M_n(\mathbb{C})$, $t \longmapsto gc_X(t)g^{-1}$. Then c_3 is C^∞, $c_3(\mathbb{R}) = gGg^{-1} = G$, $c_3(0) = gc_X(0)g^{-1} = I_n$, and $c'_3(0) = gc'_X(0)g^{-1} = gXg^{-1}$. So $gXg^{-1} \in G$.

(4) Finally, consider $c_4 : \mathbb{R} \to M_n(\mathbb{C})$, $t \longmapsto c_Y(t)Xc_Y(t)^{-1}$. Then c_4 is C^∞ and $c_4(\mathbb{R}) \subset \mathrm{Lie}(G)$ by (3). As $\mathrm{Lie}(G)$ is a $\mathbb{R}$-subvector space of $M_n(\mathbb{C})$ by (1) and (2), it is closed in $M_n(\mathbb{C})$, and $c'_4(0) \in \mathrm{Lie}(G)$. Let's calculate $c'_4(0)$. First, using the fact that

$$\frac{d}{dt}\left(c_Y(t)c_Y(t)^{-1}\right) = \frac{d}{dt}I_n = 0,$$

we see that

$$\frac{d}{dt}(c_Y(t)^{-1}) = -c_Y(t)^{-1}c'_Y(t)c_Y(t)^{-1}.$$

So

$$c_4'(t) = c_Y'(t)Xc_Y(t)^{-1} - c_Y(t)Xc_Y(t)^{-1}c_Y'(t)c_Y(t)^{-1},$$

and $\mathrm{Lie}(G) \ni c_4'(0) = YX - XY$.

□

If we want to actually calculate Lie algebras of closed subgroups of $\mathrm{GL}_n(\mathbb{C})$, we need a characterization that's easier to use. This is the goal of the following theorem.

Theorem VI.4.3. *Let G be a closed subgroup of $\mathrm{GL}_n(\mathbb{C})$. Then*

$$\mathrm{Lie}(G) = \{X \in \mathfrak{gl}_n(\mathbb{C}) | \forall t \in \mathbb{R},\ e^{tX} \in G\}.$$

Lemma VI.4.4. *Let $c : \mathbb{R} \to M_n(\mathbb{C})$ be a C^∞ map such that $c(0) = I_n$, and let $X = c'(0)$. Then, for every $t \in \mathbb{R}$,*

$$\lim_{k \to +\infty} c(\tfrac{t}{k})^k = e^{tX}.$$

Proof. Remember from section VI.3 that we have the C^∞ function $\log$, defined in a neighborhood U of I_n in $M_n(\mathbb{C})$, and satisfying $\exp(\log(A)) = A$ for every $A \in U$. Fix $\varepsilon > 0$ such that $c(t) \in U$ if $|t| < \varepsilon$. Let $d :]-\varepsilon, \varepsilon[\to M_n(\mathbb{C})$, $t \longmapsto \log(c(t))$. This is a C^∞ map, and we have

$$d'(0) = d\log_0(c'(0)) = (d\exp_{I_n})^{-1}(X) = X.$$

By Taylor's formula, $d(t) = tX + O(t^2)$.

Now let $t \in \mathbb{R}$. Choose an integer $N \geq 1$ such that $|t/N| < \varepsilon$. Then if $k \geq 1$ and $\ell \in \{0, \ldots, N-1\}$, we have

$$d(\frac{t}{Nk+\ell}) = \frac{tX}{Nk+\ell} + O(\frac{t^2}{(Nk+\ell)^2} = \frac{tX}{Nk+\ell} + O(\frac{1}{k^2}).$$

So $(Nk+\ell)d(\frac{t}{Nk+\ell}) = tX + O(\frac{1}{k})$, and

$$c(\frac{t}{Nk+\ell})^{Nk+\ell} = \exp((Nk+\ell)d(\frac{t}{Nk+\ell})) = \exp(X + O(\frac{1}{k})).$$

Making k tend to $+\infty$, we get $\lim_{r \to +\infty} c(\frac{t}{r})^r = \exp(X)$.

□

Remark VI.4.5. Applying this lemma to the function $c : \mathbb{R} \to \mathbb{R}$, $t \longmapsto 1 + t$, we recover the classical result that

$$\lim_{k \to +\infty} (1 + \tfrac{t}{k})^k = e^t.$$

Proof of the theorem. If $X \in \mathfrak{gl}_n(\mathbb{C})$ is such that $e^{tX} \in G$ for every $t \in R$, then the C^∞ map $c : \mathbb{R} \to M_n(\mathbb{C})$, $t \longmapsto e^{tX}$, has image contained in G, sends 0 to I_n, and satisfies $c'(0) = X$. So $X \in \mathrm{Lie}(G)$.

Conversely, let $X \in \mathrm{Lie}(G)$, and let's prove that $e^{tX} \in G$ for every $t \in \mathbb{R}$. Choose a C^∞ map $c : \mathbb{R} \to M_n(\mathbb{C})$ such that $c(\mathbb{R}) \subset G$, $c(0) = I_n$ and $c'(0) = X$. Let $t \in \mathbb{R}$. By the lemma, $e^{tX} = \lim_{k\to+\infty} c(\frac{t}{k})^r$. We have $c(\frac{t}{k})^k \in G$ for every $k \geq 1$. As G is a closed subgroup of $\mathrm{GL}_n(\mathbb{C})$, this implies that their limit e^{tX} is also in G.

□

Example VI.4.6.

- Take $G = \mathrm{SL}_n(\mathbb{C}) := \{g \in \mathrm{GL}_n(\mathbb{C}) | \det(g) = 1\}$. As $\det(e^{tX}) = e^{\mathrm{Tr}(tX)}$ for every $t \in \mathbb{R}$ and $X \in \mathfrak{gl}_n(\mathbb{C})$, we immediately get

$$\mathrm{Lie}(G) = \mathfrak{sl}_n(\mathbb{C}) = \{X \in \mathfrak{gl}_n(\mathbb{C}) | \mathrm{Tr}(X) = 0\}.$$

- Take $G = \mathrm{GL}_n(\mathbb{R})$. I claim that $\mathrm{Lie}(G) = \mathfrak{gl}_n(\mathbb{R})$. Indeed, if $X \in \mathfrak{gl}_n(\mathbb{R})$, then $e^{tX} \in \mathrm{GL}_n(\mathbb{R})$ for every $t \in \mathbb{R}$. Conversely, let $X \in \mathfrak{gl}_n(\mathbb{C})$ be such that $e^{tX} \in \mathrm{GL}_n(\mathbb{R})$ for every $t \in \mathbb{R}$. Then $\frac{1}{t}(e^{tX} - I_n) \in \mathfrak{gl}_n(\mathbb{R})$ for every $t \in \mathbb{R}$, so

$$X = \frac{d}{dt} e^{tX}\Big|_{t=0} = \lim_{t\in\mathbb{R}, t\to 0} \frac{1}{t}(e^{tX} - I_n) \in \mathfrak{gl}_n(\mathbb{R}).$$

- Take $G = \mathrm{U}(n) := \{g \in \mathrm{GL}_n(\mathbb{C}) | gg^* = I_n\}$. (This is called the *unitary group*.) I claim that

$$\mathrm{Lie}(G) = \mathfrak{u}(n) := \{X \in \mathfrak{gl}_n(\mathbb{C}) | X + X^* = 0\}.$$

  Indeed, if $X \in \mathfrak{u}(n)$, then $X^* = -X$, so X and X^* commute, so, for every $t \in \mathbb{R}$,

$$I_n = e^{t(X+X^*)} = e^{tX} e^{tX^*} = (e^{tX})(e^{tX})^*,$$

  and therefore $e^{tX} \in \mathrm{U}(n)$. Hence $\mathfrak{u}(n) \subset \mathrm{Lie}(G)$.

  Conversely, let $X \in \mathrm{Lie}(G)$. Then $e^{tX} e^{tX^*} = I_n$ for every $t \in \mathbb{R}$, hence

$$0 = \frac{d}{dt}(e^{tX} e^{tX^*})\Big|_{t=0} = X + X^*,$$

  and $X \in \mathfrak{u}(n)$. So $\mathrm{Lie}(G) \subset \mathfrak{u}(n)$.

- Take $G = \mathrm{SU}(n) := \mathrm{U}(n) \cap \mathrm{SL}_n(\mathbb{C})$. (This is called the *special unitary group*.) Then

$$\mathrm{Lie}(G) = \mathfrak{u}(n) \cap \mathfrak{sl}_n(\mathbb{C}) = \{X \in \mathfrak{gl}_n(\mathbb{C}) | X + X^* = 0 \text{ and } \mathrm{Tr}(X) = 0\}.$$

  This Lie algebra is denoted by $\mathfrak{su}(n)$.

- Similarly, if $\mathrm{O}(n) = \{g \in \mathrm{GL}_n(\mathbb{R}) | X^t X = I_n\}$ and $\mathrm{SO}(n) = \mathrm{O}(n) \cap \mathrm{SL}_n(\mathbb{R})$, then

$$\mathrm{Lie}(\mathrm{O}(n)) = \mathrm{Lie}(\mathrm{SO}(n)) = \mathfrak{so}(n) := \{X \in \mathfrak{gl}_n(\mathbb{R}) | X + {}^t X = 0\}.$$

  Note that $\mathrm{O}(n)$ and $\mathrm{SO}(n)$ have the same Lie algebra, unlike $\mathrm{U}(n)$ and $\mathrm{SU}(n)$. This is because $\mathrm{SO}(n)$ is the connected component of I_n in $\mathrm{O}(n)$.

- For every commutative ring k, let $B_n(k)$ (resp. $U_n(k)$, resp. $T_n(k)$) be the group of upper triangular matrices (resp. upper triangular matrices with 1s on the diagonal, resp. diagonal matrices) in $\mathrm{GL}_n(k)$. Then, if $k = \mathbb{R}$ or $\mathbb{C}$,

$$\mathrm{Lie}(B_n(k)) = \mathfrak{b}_n(k)$$

$$\mathrm{Lie}(U_n(k)) = \mathfrak{u}_n(k)$$

$$\mathrm{Lie}(T_n(k)) = \mathfrak{t}_n(k)$$

  (see example VI.1.6 for the notation).

- If G is a finite subgroup of $\mathrm{GL}_n(\mathbb{C})$, then $\mathrm{Lie}(G) = \{0\}$.

The next theorem is the down-to-earth version of the statement "$\mathrm{Lie}(G)$ is the tangent space of G at I_n".

Theorem VI.4.7. *Let G be a closed subgroup of $\mathrm{GL}_n(\mathbb{C})$, and let $\mathfrak{g} = \mathrm{Lie}(G)$. Then there exist neighborhoods V of I_n in $\mathrm{GL}_n(\mathbb{C})$ and U of 0 in $\mathfrak{gl}_n(\mathbb{C})$ such that $\exp : U \to V$ is a diffeomorphism (= C^∞ and bijective with C^∞ inverse) and that $\exp(U \cap \mathfrak{g}) = V \cap G$.*

Proof. Let W be a $\mathbb{R}$-subspace of $\mathfrak{gl}_n(\mathbb{C})$ such that $\mathfrak{g} \oplus W = \mathfrak{gl}_n(\mathbb{C})$. Consider the map $\varphi : \mathfrak{gl}_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$ sending $A + B$ to $e^A e^B$ if $A \in \mathfrak{g}$ and $B \in W$. Then, for $A \in \mathfrak{g}$ and $B \in W$, we have

$$\varphi(A + B) = (I_n + A + O(A^2))(I_n + B + O(B^2)) = I_n + A + B + O(AB, BA, A^2, B^2),$$

so $d\varphi_0 = \mathrm{id}_{M_n(\mathbb{C})}$. By the inverse function theorem, there exists neighborhoods U_1 of 0 in $\mathfrak{g}$, U_2 of 0 in W and V of 1 in $\mathrm{GL}_n(\mathbb{C})$ such that $\varphi : U_1 \times U_2 \to V$ is a diffeomorphism.

Now let's show that, after shrinking U_2 (and consequently V), we have $\exp^{-1}(V \cap G) = U_1$, i.e. $\exp : U_1 \to G \cap V$ is a diffeomorphism. Suppose that this is not the case. Then we can find a decreasing sequence $W_0 = U_2 \supset W_1 \supset W_2 \supset \dots$ of neighborhoods of 0 in W such that every neighborhood of 0 in W contains one of the W_r, and that, for every $r \geq 0$, there exist $A_r \in U$ and $B_r \in W_r - \{0\}$ such that $\varphi(A_r, B_r) = e^{A_r} e^{B_r} \in G$.

As $\bigcap_{r \geq 0} W_r = \{0\}$, $B_r \to 0$ as $r \to +\infty$. Let $Y_r = \frac{1}{\|B_r\|} B_r$; then $\|Y_r\| = 1$. As the unit sphere of W is compact (because W is finite-dimensional), we may assume that the sequence $(Y_r)_{r \geq 0}$ has a limit, say Y. Then $Y \in W$ and $Y \neq 0$ (because $\|Y\| = 1$). We want to show that $e^{tY} \in \mathrm{Lie}(G)$ for $t \in \mathbb{R}$, which will imply that $Y \in \mathfrak{g}$ and contradict the fact that $\mathfrak{g} \cap W = \{0\}$.

Let $t \in \mathbb{R}$. For every $r \geq 0$, if $m_r = \lfloor \frac{t}{\|B_r\|} \rfloor$, then we have $m_r\|B_r\| \leq t \leq (m_r+1)\|B_r\|$. So $m_r\|B_r\| \to t$ as $r \to +\infty$. Now note that, for every $r \geq 0$, $e^{B_r} \in G$ (because $\varphi(A_r, B_r) \in e^{A_r}e^{B_r} \in G$, and $e^{A_r} \in G$), so $e^{m_r B_r} \in G$. As $e^{m_r B_r} = e^{m_r\|B_r\|Y_r}$, we get $e^{m_r\|B_r\|Y_r} \in G$. Making $r \to +\infty$ and using the fact that G is closed in $\mathrm{GL}_n(\mathbb{C})$, we get $e^{tY} \in G$.

□

VI.5 From groups representations to Lie algebra representations

This is the analogue of problem VII.6.5, but for closed subgroups of $\mathrm{GL}_n(\mathbb{C})$ instead of linear algebraic groups.

Everything works thanks to the following proposition, which generalizes problem VII.5.5(2).

Proposition VI.5.1. *Let $c : \mathbb{R} \to \mathrm{GL}_n(\mathbb{C})$ be a continuous morphism of groups. Then there exists a unique $B \in M_n(\mathbb{C})$ such that $c(t) = e^{tB}$, $\forall t \in \mathbb{R}$.*

Proof. The uniqueness of B follows from the fact that $\frac{d}{dt}e^{tB}|_{t=0} = B$.

For the existence, choose $\delta > 0$ such that $\|g - I_n\| \leq \delta\| \Rightarrow \det(g) \neq 0$. As c is continuous, there exists $\varepsilon > 0$ such that $|t| \leq \varepsilon \Rightarrow \|c(t) - c(0)\| \leq \delta$. Let $A = \int_0^\varepsilon c(t)dt$. Then

$$\|\varepsilon I_n - A\| = \| \int_0^\varepsilon (c(0) - c(t))dt\| \leq \int_0 \varepsilon\delta dt = \varepsilon\delta,$$

so $\|I_n - \varepsilon^{-1}A\| \leq \delta$, so $\det(\varepsilon^{-1}A) \neq 0$, so A is invertible.

As c is a morphism of groups, $c(t+s) = c(t)c(s)$ for all $t, s \in \mathbb{R}$. Let $t \in \mathbb{R}$. Then

$$\int_t^{t+\varepsilon} c(s)ds = c(t) \int_0 \varepsilon c(s)ds = c(t)A,$$

so

$$c(t) = \left(\int_t^{t+\varepsilon} c(s)ds \right) A^{-1}.$$

Hence c is C^1, and

$$c'(t) = (c(t+\varepsilon) - c(t))A^{-1} = c(t)B,$$

where $B = (c(\varepsilon) - I_n)A^{-1}$. As $c(0) = I_n$, we finally get $c(t) = e^{tB}$.

□

Theorem VI.5.2. *Let G be a closed subgroup of $\mathrm{GL}_n(\mathbb{C})$ and H be a closed subgroup of $\mathrm{GL}_m(\mathbb{C})$. Let $\rho : G \to H$ be a continuous morphism of groups. For every $X \in \mathrm{Lie}(G)$, the map $c_X : \mathbb{R} \to \mathrm{GL}_m(\mathbb{C})$, $t \longmapsto \rho(e^{tX})$, is a continuous morphism of groups, hence, by proposition VI.5.1, it is of the form $t \longmapsto e^{tY}$ for a uniquely determined $Y \in M_m(\mathbb{C})$. We write $Y = d\rho(X)$. Then*

1. *For every $X \in \mathrm{Lie}\, G$, we have*
$$d\rho(X) = \frac{d}{dt}\rho(e^{tX})|_{t=0}.$$
2. *For every $X \in \mathrm{Lie}\, G$, $d\rho(X) \in \mathrm{Lie}\, H$.*
3. *$d\rho : \mathrm{Lie}\, G \to \mathrm{Lie}\, H$ is a morphism of Lie algebras.*
4. *For every $X \in \mathrm{Lie}\, G$ and $g \in G$, $\rho(e^X) = e^{d\rho(X)}$ and $d\rho(gXg^{-1}) = \rho(g)d\rho(X)\rho(g)^{-1}$.*

Remark VI.5.3. If we knew differential geometry, we would say that $d\rho$ is the differential of the map ρ at $1 \in G$.

Remark VI.5.4. As $\exp(\mathrm{Lie}\, G)$ contains a neighbourhood of 1 in G (by theorem VI.4.7), $d\rho$ determines ρ on G^0 (the connected component of I_n in G).

Proof. Point (i) follows immediately from the definition of $d\rho(X)$, and (ii) follows immediately from the definition of $\mathrm{Lie}\, H$ and the fact that $c_X(\mathbb{R}) \subset H$.

The first assertion of (iv) just follows from the formula $\rho(e^{tX}) = e^{td\rho(X)}$ (which is the definition of $d\rho(X)$), taking $t = 1$. For the second assertion of (iv), consider $c : \mathbb{R} \to H$, $t \longmapsto \rho(g)\rho(e^{tX})\rho(g)^{-1}$. Then

$$c(t) = \rho(ge^{tX}g^{-1}) = \rho(e^{tgXg^{-1}}),$$

so

$$\rho(g)d\rho(X)\rho(g)^{-1} = c'(0) = d\rho(gXg^{-1}).$$

We prove (iii). Let $X, Y \in \mathrm{Lie}\, G$ and $\lambda \in \mathbb{R}$.

(a) We have
$$c_{\lambda X}(t) = \rho(e^{\lambda tX}) = c_X(\lambda t),$$
so
$$e^{td\rho(\lambda X)} = e^{t\lambda d\rho(X)}.$$
Taking derivatives at 0 gives $d\rho(\lambda X) = \lambda d\rho(X)$.

(b) Let $c : \mathbb{R} \to H \subset \mathrm{GL}_m(\mathbb{C})$, $t \longmapsto \rho(e^{tX})\rho(e^{tY})\rho(e^{t(X+Y)})^{-1}$. As $c = c_X c_Y c_{-(X+Y)}$, the map c is C^∞, and we have $c'(0) = d\rho(X) + d\rho(Y) - d\rho(X+Y)$.

By lemma VI.4.4, for every $t \in \mathbb{R}$,

$$\lim_{k\to+\infty} c\left(\frac{t}{k}\right)^k = e^{tc'(0)}.$$

So we just need to prove that, for t fixed,

$$\lim_{k\to+\infty} c\left(\frac{t}{k}\right)^k = I_m.$$

Fix $t \in \mathbb{R}$. Then

$$c\left(\frac{t}{k}\right) = \rho\left(e^{\frac{t}{k}X}e^{\frac{t}{k}Y}e^{-\frac{t}{k}(X+Y)}\right) = \rho(I_n + O(\frac{1}{k^2})),$$

so

$$c\left(\frac{t}{k}\right)^k = \rho((I_n + O(\frac{1}{k^2}))^k) = \rho(I_n + O(\frac{1}{k})),$$

and this tends to $\rho(I_n) = I_m$ as $k \to +\infty$, by continuity of ρ.

(c) Let $c : \mathbb{R} \to \operatorname{Lie} H$, $t \longmapsto d\rho(e^{tX}Ye^{-tX})$. As $d\rho$ is a $\mathbb{R}$-linear map between finite-dimensional $\mathbb{R}$-vector spaces, it is C^∞ and equal to its own differential at every point. So the chain rule gives $c'(0) = d\rho([X, Y])$.

On the other hand, by (iv), $c(t) = \rho(e^{tX})d\rho(Y)\rho(e^{-tX})$, so

$$c'(0) = d\rho(X)d\rho(Y) - d\rho(Y)d\rho(X) = [d\rho(X), d\rho(Y)].$$

□

In particular, every continuous representation $\rho : G \to \mathrm{GL}_m(\mathbb{C})$ of G gives rises to a representation $d\rho : \operatorname{Lie} G \to \mathfrak{gl}_m(\mathbb{C})$ of the Lie algebra of G. Note that $d\rho$ is a much simpler objet than ρ (we don't have to worry about continuity or derivability conditions, for example, as they follow automatically from $\mathbb{R}$-linearity). So we would like to understand the representations of G through the representations of its Lie algebra. This raises two questions :

Question 1 : Does $d\rho$ determine ρ ?

Answer : We have already seen the answer : $d\rho$ determines ρ on G^0. So it's a good tool to understand ρ if G is connected, not so much if G is finite.

Question 2 : Is every Lie algebra map $u : \operatorname{Lie} G \to \mathfrak{gl}_m(\mathbb{C})$ equal to $d\rho$, for some continuous group morphism $\rho : G \to \mathrm{GL}_m(\mathbb{C})$?

Answer : Not in general. Compare the cases $G = S^1 \subset \mathbb{C}^\times$ and $H = \mathbb{R}_{>0} \subset \mathbb{C}^\times$ (both connected). Every representation of $\mathrm{Lie}(H)$ (on a finite-dimensional $\mathbb{C}$-vector space) comes from a representation of H, but this is not true for representations of $\mathrm{Lie}(G)$. For example, the

$\mathbb{R}$-linear map $R \to \mathbb{C}$, $x \longmapsto \frac{x}{2}$, is not the differential of a continuous morphism of groups $\rho : S^1 \to \mathbb{C}^\times$, because every such ρ is of the form $z \longmapsto z^n$ for some $n \in \mathbb{Z}$ by problem VII.5.5(2) (or proposition VI.5.1).

The difference is that H is simply connected and G is not. In general, we can tell which $u : \mathrm{Lie}\, G \to \mathfrak{gl}_m(\mathbb{C})$ lift (or integrate) to G, but the formulation of the answer uses the language of root systems. In the particular case where G is simply connected and connected, every u lifts.

In section VI.8, we will do by hand the case of $G = \mathrm{SU}(n)$ (which is connected and simply connected).

Remark VI.5.5. If $g \in G$ and $X \in \mathrm{Lie}\, G$, let $\mathrm{Ad}(g)$ and $\mathrm{ad}(X)$ be the endomorphisms of $\mathrm{Lie}(G)$ defined by $\mathrm{Ad}(g)(Y) = gYg^{-1}$ and $\mathrm{ad}(X)(Y) = [X, Y]$. Then $\mathrm{Ad} : G \to \mathrm{GL}(\mathrm{Lie}\, G)$ is a continuous morphism of groups.

We have seen that

$$\frac{d}{dt}(e^{tX} Y e^{-tX})|_{t=0} = [X, Y],$$

that is,

$$\frac{d}{dt}(\mathrm{Ad}(e^{tX}))|_{t=0} = \mathrm{ad}(X).$$

In other words, $\mathrm{ad} = d\,\mathrm{Ad}$ (which gives again the fact that $\mathrm{ad} : \mathrm{Lie}\, G \to \mathfrak{gl}(\mathrm{Lie}\, G)$ is a morphism of Lie algebras), and so we get

$$e^{\mathrm{ad}(X)} = \mathrm{Ad}(e^X).$$

Remark VI.5.6. If $\mathfrak{g}$ is a k-Lie algebra and (V_1, u_1) and (V_2, u_1) are representations of $\mathfrak{g}$, we define representations of $\mathfrak{g}$ on $V_1 \otimes_k V_2$ and $\mathrm{Hom}_k(V_1, V_2)$ by :

$$X(v_1 \otimes v_2) = (u_1(X)v_1) \otimes v_2 + v_1 \otimes (u_2(X)v_2)$$

and

$$X\varphi = u_2(X) \circ \varphi - \varphi \circ u_1(X),$$

for $X \in \mathfrak{g}$, $v_1 \in V_1$, $v_2 \in V_2$ and $\varphi \in \mathrm{Hom}_k(V_1, V_2)$. In particular, taking $V_2 = k$ with the trivial representation, we get a representation of $\mathfrak{g}$ on $V_1^* = \mathrm{Hom}_k(V_1, k)$ given by $X\varphi = -\varphi \circ u_1(X)$.

The justification for this is the following (apart from the fact that these formulas do indeed define representations) : If $\mathfrak{g}$ is the Lie algebra of a closed subgroup G of $\mathrm{GL}_n(\mathbb{C})$ and we have $u_1 = d\rho_1$ and $u_2 = d\rho_2$ with $\rho_1 : G \to \mathrm{GL}(V_1)$ and $\rho_2 : G \to \mathrm{GL}(V_2)$ continuous finite-dimensional representations of G on complex vector spaces, then the representation of $\mathfrak{g}$ on $V_1 \otimes_{\mathbb{C}} V_2$ (resp. $\mathrm{Hom}_{\mathbb{C}}(V_1, V_2)$) defined above is obtained by deriving the representation of G on this space. (See problem VII.6.20.) Also, the trivial representation of $\mathfrak{g}$ is clearly the differential of the trivial representation of G.

VI.6 The Baker-Campbell-Hausdorff formula

This is is formula that allows us to express the multiplication in G in terms of the sum and bracket of $\mathrm{Lie}\, G$ (at least in a neighbourhood of 1).

It has many proofs,[6] but here we'll follow a quick and purely algebraic proof due to Eichler. (See the paper [10].)

Here is the setup : Let $\mathbb{Q}\{t_1, \ldots, t_N\} \subset \mathbb{Q}\{\{t_1, \ldots, t_N\}\}$ be the rings of polynomials resp. power series in N noncommuting indeterminates $t_1, \ldots, t_N$. The degree of an element of $\mathbb{Q}\{t_1, \ldots, t_N\}$ is defined in the obvious way.

Let $L \subset \mathbb{Q}\{t_1, \ldots, t_N\}$ be the Lie subalgebra generated by $t_1, \ldots, t_N$. Elements of L are sometimes called *Lie polynomials*. (Note : L is called the free Lie algebra on the set $\{t_1, \ldots, t_N\}$. Also, it's easy to see that $\mathbb{Q}\{t_1, \ldots, t_N\}$ is the universal envelopping algebra of L, but we won't need that fact.) [7]

Proposition VI.6.1. *We have $L = \bigoplus_{n \in \mathbb{Z}_{\geq 0}} L_n$, where L_n is the space of homogeneous degree n polynomials in L (by convention, 0 is in every L_n).*

In other words, if $f \in \mathbb{Q}\{t_1, \ldots, t_N\}$ is a Lie polynomial, then its homogeneous degree n part is also a Lie polynomial.

Proof. Obviously, we have $t_1, \ldots, t_N \in \bigoplus_{n \geq 0} L_n \subset L$, so we just need to show that $\bigoplus_{n \geq 0} L_n$ is a Lie subalgebra of $\mathbb{Q}\{t_1, \ldots, t_N\}$, i.e. is stable by $[.,.]$. But we clearly have $[L_n, L_m] \subset L_{m+m}$.

□

Now let t, s be two noncommuting indeterminates, and consider the formal power series

$$e^t = \sum_{n \geq 0} \frac{t^n}{n!} \in \mathbb{Q}[[t]]$$

$$\log(t) = \sum_{n \geq 1} \frac{(-1)^{n-1}}{n} (t-1)^n \in \mathbb{Q}[[t-1]].$$

Because the constant term of $e^t e^s \in \mathbb{Q}\{\{t, s\}\}$ is equal to 1, the formal power series $\log(e^t e^s) \in \mathbb{Q}\{\{t, s\}\}$ makes sense. Write

$$\log(e^t e^s) = \sum_{n \geq 0} F_n(t, s),$$

with $F_n(t, s)$ homogeneous of degree n. It's easy enough to calculate the first few terms :

$$F_0(t, s) = 0, \quad F_1(t, s) = t + s, \quad F_2(t, s) = \frac{1}{2}[t, s].$$

[6] See for example theorem 7.4 of chapter IV of part I of Serre's book [31].

[7] See theorem 4.2 of chapter IV of part I of Serre's book [31] for a proof.

Theorem VI.6.2 (Baker-Campbell-Hausdorff formula)**.** *For every $n \geq 0$, $F_n(t,s) \in L$, i.e. $F_n(t,s)$ is a Lie polynomial, i.e. $F_n(t,s)$ is obtained from t and s using only vector space operations and the commutator bracket.*

Proof. By induction on n. We have already checked the cases $n = 0, 1, 2$. Let $n \geq 3$, and suppose that we know the theorem for any $m \leq n-1$. Introduce a third noncommuting indeterminate u. We have $(e^t e^s)e^u = e^t(e^s e^u)$ in $\mathbb{Q}\{\{t,s,u\}\}$, hence $\log((e^t e^s)e^u) = \log(e^t(e^s e^u))$, hence

$$(*) \qquad \sum_{i=1}^{+\infty} F_i(\sum_{j=1}^{+\infty} F_j(t,s), u) = \sum_{i=1}^{+\infty} F_i(t, \sum_{j=1}^{+\infty} F_j(s,u)).$$

If $i \geq n+1$, then $F_i(\sum_{j=1}^{+\infty} F_j(t,s), u)$ and $F_i(t, \sum_{j=1}^{+\infty} F_j(s,u))$ only have homogeneous components in degree $\geq n+1$. As for $F_n(\sum_{j=1}^{+\infty} F_j(t,s), u)$ and $F_n(t, \sum_{j=1}^{+\infty} F_j(s,u))$, they only have homogeneous components of degree $\geq n$, and their degree n homogeneous components are $F_n(t+s,u)$ and $F_n(t,s+u)$ respectively.

By the induction hypothesis, if $i \leq n-1$, then $F_i(\sum_{j=1}^{n-1} F_j(t,s), u)$ and $F_i(t, \sum_{j=1}^{n-1} F_j(s,u))$ are Lie polynomials, and their homogeneous components also are Lie polynomials by proposition VI.6.1. Also, if $i \geq 2$, the difference $F_i(\sum_{j=1}^{+\infty} F_j(t,s), u) - F_i(\sum_{j=1}^{n-1} F_j(t,s), u)$ only has homogeneous components in degree $\geq n+1$. If $i = 1$, $F_i(t,s) = t+s$, so the difference above is just $\sum_{j=n}^{+\infty} F_j(t,s)$, it only has homogeneous components in degree $\geq n$, and its degree n homogeneous component is $F_n(t,s)$. Similarly, we see that

$$\sum_{i=1}^{n-1} F_i(t, \sum_{j=1}^{+\infty} F_j(s,u)) - \sum_{i=1}^{n-1} F_i(t, \sum_{j=1}^{n-1} F_j(s,u))$$

only has homogeneous components in degree $\geq n$, and its degree n homogeneous component is $F_n(s,u)$.

And so finally, if we look at the degree n homogeneous components in the equality (*) above, and omit the parts that we know are Lie polynomials by the induction hypothesis, we are left with

$$F_n(t+s,u) + F_n(t,s)$$

on the left-hand side and

$$F_n(t,s+u) + F_n(s,u)$$

on the right-hand side. So these two polynomials are equal modulo L. We also know that $F_n(t,s)$ is homogeneous of degree n, and that $F_n(\lambda t, \mu t) = 0$ for every $\lambda, \mu \in \mathbb{Q}$ (because λt and μt commute, so $\log(e^{\lambda t}e^{\mu t}) = (\lambda+\mu)t$, and because $n \geq 2$).

Now we just have to prove the following fact : Let $f \in \mathbb{Q}\{\{t,s\}\}$ be such that :

(1) $f(t+s,u) + f(t,s) = f(t,s+u) + f(s,u) \mod L$;

(2) $f(\lambda t, \lambda s) = \lambda^n f(t,s)$ for every $\lambda \in \mathbb{Q}$;

(3) $f(\lambda t, \mu t) = 0$ for every $\lambda, \mu \in \mathbb{Q}$.

Then $f \in L$.

Here is the proof of this fact :

Substituting $u = -s$ in (1) (and using (3)) gives :

$$(4) \qquad f(t,s) = -f(t+s,-s) \mod L.$$

Substituting $s = -t$ in (1) (and using (3)) gives

$$(5) \qquad f(s,u) = -f(-s,s+u) \mod L.$$

Hence (5), then (4), then (5), then (2), we get

$$(6) \qquad f(t,s) = -f(-t,t+s) = f(s,-(t+s)) = -f(-s,-t) = (-1)^{n+1} f(s,t) \mod L.$$

Now substituting $u = -\frac{1}{2}s$ in (1) gives :

$$(7) \qquad f(t,s) = f(t,s/2) - f(t+s,-s/2) \mod L.$$

And substituting $t = -\frac{1}{2}s$ in (1) gives :

$$(8) \qquad f(s,u) = f(s/2,u) - f(-s/2,s+u) \mod L.$$

We apply (7) to both terms in the right-hand side of (8) (and use (2)) to get :

$$f(s,u) = 2^{-n} f(s,u) - 2^{-n} f(-s,s+u) - f(s/2+u,-u/2) + f(s/2+u,-s/2-u/2) \mod L.$$

Applying (4) to the last two terms of the right-hand side of the equality above and (5) to the second term gives :

$$f(s,u) = 2^{-n} f(s,u) + 2^{-n} f(s,u) + 2^{-n} f(s+u,u) - 2^{-n} f(u,s+u) \mod L,$$

which by (6) becomes

$$f(s,u) = 2^{1-n} f(s,u) + 2^{-n}(1+(-1)^n) f(s+u,u) \mod L,$$

and finally

$$(9) \qquad f(s,u) = \frac{2^{-n}}{1-2^{1-n}}(1+(-1)^n) f(s+u,u) \mod L.$$

If n is odd, this gives $f(s,u) \in L$. If n is even, replacing t by $t-s$ in (4) and applying (9) gives

$$-f(t,-s) = f(t-s,s) = \frac{2^{-n}}{1-2^{1-n}}(1+(-1)^n) f(t,s) \mod L,$$

hence :

$$(10) \qquad f(t,-s) = -\frac{2^{-n}}{1-2^{1-n}}(1+(-1)^n)f(t,s) \mod L.$$

Now if we use (10) twice, we get :

$$f(t,s) = (\frac{2^{-n}}{1-2^{1-n}}(1+(-1)^n))^2 f(t,s) \mod L,$$

hence $f(t,s) \in L$, because $\frac{2^{-n}}{1-2^{1-n}}(1+(-1)^n) \neq 1$.

□

VI.7 Representations of $\mathfrak{sl}_2(\mathbb{C})$

The following results are proved in problem VII.6.10 :

Theorem VI.7.1. *(i) For every $n \geq 0$, there is exactly one irreducible $(n+1)$-dimensional representation W_{n+1} of $\mathfrak{sl}_2(\mathbb{C})$ (up to isomorphism), and it is given by the following formulas : There is a basis $(v_0, \ldots, v_n)$ of W_{n+1} such that*

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot v_i = \begin{cases} (n+1-i)v_{i-1} & \text{if } i \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \cdot v_i = \begin{cases} (i+1)v_{i+1} & \text{if } i \leq n-1 \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot v_i = (n-2i)v_i.$$

(ii) For every finite-dimensional representation u of $\mathfrak{sl}_2(\mathbb{C})$, $u\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is semisimple (= diagonalizable).

Corollary VI.7.2. *For every $\mathbb{C}$-Lie algebra map $u : \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}_n(\mathbb{C})$, $u\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is a semisimple (= diagonalizable) element of $\mathfrak{gl}_n(\mathbb{C}) = M_n(\mathbb{C})$, and its eigenvalues are in $\mathbb{Z}$.*

Proof. We get the integrality of the eigenvalues of $u\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ by considering a filtration $0 = V_0 \subset V_1 \subset \cdots \subset V_d = \mathbb{C}^n$ by subrepresentations such that each V_i/V_{i-1} is irreducible.

□

Corollary VI.7.3. *For every $\mathbb{R}$-Lie algebra map $u : \mathfrak{su}(2) \to \mathfrak{gl}_n(\mathbb{C})$, $u\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$ is a semisimple (= diagonalizable) element of $\mathfrak{gl}_n(\mathbb{C}) = M_n(\mathbb{C})$, and its eigenvalues are in $i\mathbb{Z}$.*

Proof. Note that $\mathfrak{sl}_2(\mathbb{C}) = \mathbb{C}\mathfrak{su}(2) \subset \mathfrak{gl}_2(\mathbb{C})$, and this induces $\mathfrak{sl}_2(\mathbb{C}) \simeq \mathbb{C} \otimes_{\mathbb{R}} \mathfrak{su}(2)$. Define a $\mathbb{C}$-Lie algebra map $v : \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}_n(\mathbb{C})$ by $v(a \otimes X) = au(X)$ if $a \in \mathbb{C}$, $X \in \mathfrak{su}(2)$, and apply corollary VI.7.2 to v.

□

VI.8 Lifting representations of $\mathfrak{su}(n)$

Notation VI.8.1. If $a_1, \dots, a_n \in \mathbb{C}$, we write $\mathrm{diag}(a_1, \dots, a_n)$ for the diagonal matrix with diagonal entries $a_1, \dots, a_n$.

Theorem VI.8.2. *Let $u : \mathfrak{su}(n) \to \mathfrak{gl}_m(\mathbb{C})$ be a morphism of $\mathbb{R}$-Lie algebras. Then there exists a unique continuous group morphism $\rho : \mathrm{SU}(n) \to \mathrm{GL}_m(\mathbb{C})$ such that $u = d\rho$.*

Remember that this means that, for every $X \in \mathfrak{su}(n)$, $e^{u(X)} = \rho(e^X)$.

Proof. First, we have already seen that ρ is unique (because $\mathrm{SU}(n)$ is connected, hence generated by a neighbourhood of I_n, and $\exp(\mathfrak{su}(n))$ contains a neighbourghood of 1). [8]

(1) Let's show that, for every $X, Y \in \mathfrak{su}(n)$,

$$u(e^X Y e^{-X}) = e^{u(X)} u(Y) e^{-u(X)}.$$

Indeed

$$e^X Y e^{-X} = \mathrm{Ad}(e^X)(Y) = e^{\mathrm{ad}(X)}(Y) = \sum_{n \geq 0} \frac{1}{n!} (\mathrm{ad}(X))^n(Y).$$

As $u \circ \mathrm{ad}(X) = \mathrm{ad}(u(X)) \circ u$ (because u sends brackets to brackets), this show that $u(e^X Y e^{-X})$ is equal to

$$\sum_{n \geq 0} \frac{1}{n!} (\mathrm{ad}(u(X)))^n(u(Y)) = e^{\mathrm{ad}(u(X))}(Y) = \mathrm{Ad}(e^{u(X)})(u(Y)) = e^{u(X)} u(Y) e^{-u(X)}.$$

(2) Now let's show that every element of $\mathrm{SU}(n)$ and every element of $\mathfrak{su}(n)$ is conjugate by an element of $\mathrm{SU}(n)$ to a diagonal matrix. Indeed, if $A \in \mathfrak{su}(n)$ or $\mathrm{SU}(n)$, then A is normal (ie commutes with A^*), so, by the spectral theorem, A is diagonalizable in an orthonormal basis $(e_1, \dots, e_n)$, ie there exists $S \in \mathrm{U}(N)$ (the change of basis matrix) such that SAS^{-1} is diagonal. But after replacing e_n by some $\det(S)^{-1} e_n$ (which doesn't change the fact that the basis is orthonormal, because $|\det(S)| = 1$), we may assume that $\det(S) = 1$, ie $S \in \mathrm{SU}(n)$.

[8] Actually, we'll see during the proof that $\exp(\mathfrak{su}(n)) = \mathrm{SU}(n)$.

(3) The map $\exp : \mathfrak{su}(n) \to \mathrm{SU}(n)$ is surjective. Indeed, let $g \in \mathrm{SU}(n)$. By (2), there exists $S \in \mathrm{SU}(n)$ such that $SgS^{-1} = \mathrm{diag}(a_1, \dots, a_n)$. As $SgS^{-1} \in \mathrm{SU}(n)$, $a_1, \dots, a_n \in \mathrm{U}(1)$ and their product is 1. So we can find $\mu_1, \dots, \mu_n \in i\mathbb{R}$ such that $e^{\mu_i} = a_i$ for every i and $\sum_{i=1}^n \mu_i = 0$. Let $X = \mathrm{diag}(\mu_1, \dots, \mu_n)$, then $X \in \mathfrak{su}(n)$, so $S^{-1}XS \in \mathfrak{su}(n)$, and $e^{S^{-1}XS} = S^{-1}e^X S = g$.

(4) Let $X, Y \in \mathfrak{su}(n)$ such that $e^X = e^Y$. We want to show that $e^{u(X)} = e^{u(Y)}$. By (2), X and Y are diagonalizable. Let $a_1, \dots, a_n$ (resp. $b_1, \dots, b_n$) be the eigenvalues of X (resp. Y). After changing the order of the b_r, we may assume that $e^{a_r} = e^{b_r}$ for every $r \in \{1, \dots, n\}$. As the a_r and b_r are in $i\mathbb{R}$ and $\sum_{r=1}^n a_r = \sum_{r=1}^n b_r = 0$, this implies that $a_r = b_r + 2i\pi k_r$, with $k_r \in \mathbb{Z}$ and $\sum_{r=1}^n k_r = 0$.

For every $r \in \{1, \dots, n-1\}$, let $h_r = \mathrm{diag}(0, \dots, 0, 1, -1, 0, \dots, 0)$, and let $\mathfrak{s}_r$ be the set of elements $A = (a_{ij})$ of $\mathfrak{su}(n)$ such that $a_{ij} = 0$ unless $i, j \in \{r, r+1\}$. Then $\mathfrak{s}_r \simeq \mathfrak{su}(2)$, and $ih_r \in \mathfrak{s}_r$. By corollary VI.7.3 (applied to $u_{|\mathfrak{s}_r}$), $u(ih_r)$ is semisimple (in $M_m(\mathbb{C})$) and all its eigenvalues are in $i\mathbb{Z}$. In particular, $e^{u(2i\pi h_r)} = I_m$.

Now let $g \in \mathrm{SU}(n)$ be such that $gXg^{-1} = \mathrm{diag}(a_1, \dots, a_n)$. By the previous paragraph, for every $l_1, \dots, l_{n-1} \in \mathbb{Z}$,

$$\exp(u(gXg^{-1} + \sum_{r=1}^{n-1} 2il_r\pi h_r)) = \exp(u(gXg^{-1})),$$

so

$$\exp(u(X + \sum_{r=1}^{n-1} 2il_r\pi g^{-1}h_r g)) = \exp(u(X)).$$

(We also have similar equalities withtout the "u", but they are obvious.)

This means that we can, without changing e^X and $e^{u(X)}$, replace $a_1, \dots, a_n$ with $a_1 + 2i\pi c_1, \dots, a_n + 2i\pi c_n$ for any $c_1, \dots, c_n \in \mathbb{Z}$ such that $c_1 + \cdots + c_n = 0$. Also, we have a similar result for Y.

In particular, we may assume that X and Y have the same eigenvalues, ie that they are conjugate. Write $Y = SXS^{-1}$ with $S \in \mathrm{SU}(n)$. We have $Se^XS^{-1} = e^Y = e^X$, so we can write $S = e^Z$ with $Z \in \mathfrak{su}(n)$ and Z centralizing e^X. (Writing $\mathbb{C}^n$ as a sum of eigenspaces of e^X and choosing a basis adapted to that, we may assume that S is a matrix diagonal by blocks, and we just need Z to be diagonal by blocks with the same block sizes, which can clearly be accomplished.)

We have $e^X Z e^{-X} = Z$, so using (1) gives $e^{u(X)}u(Z)e^{-u(X)} = u(Z)$, so $u(Z)$ centralizes $e^{u(X)}$, so $e^{u(Z)}$ also centralizes $e^{u(X)}$.

Next, using $Y = e^Z X e^{-Z}$ and using (1) again gives $u(Y) = e^{u(Z)}u(X)e^{-u(Z)}$, hence $e^{u(Y)} = e^{u(Z)}e^{u(X)}e^{-u(Z)}$. As $e^{u(Z)}$ centralizes $e^{u(X)}$, we get $e^{u(X)} = e^{u(Y)}$.

(5) By (3) and (4), we can define $\rho : \mathrm{SU}(n) \to \mathrm{GL}_m(\mathbb{C})$ using the formula $\rho(e^X) = e^{u(X)}$, for every $X \in \mathfrak{su}(n)$.

(6) Let $X, Y \in \mathfrak{su}(n)$ be small enough (i.e. close enough to 0) so that $\log(e^X e^Y)$ and $\log(e^{u(X)}e^{u(Y)})$ make sense. By the Baker-Campbell-Hausdorff formula (theorem VI.6.2), we have

$$\log(e^X e^Y) = \sum_{n\geq 1} F_n(X,Y),$$

with the F_n Lie polynomials in two noncommuting indeterminates. As u is a map of Lie algebras, $u(F_n(X,Y)) = F_n(u(X),u(Y))$ for every n. So

$$u(\log(e^X e^Y)) = \log(e^{u(X)}e^{u(Y)}),$$

so $\rho(e^X e^Y) = \rho(e^X)\rho(e^Y)$. In other words, if g and h are in a small enough neighbourhood of I_n in $\mathrm{SU}(n)$, then $\rho(gh) = \rho(g)\rho(h)$.

(7) Let $\gamma : \mathbb{R} \to \mathrm{GL}_n(\mathbb{C})$ be an analytic map (i.e. locally given by a converging power series), and assume that $\gamma(\mathbb{R}) \subset \mathrm{SU}(n)$. Then I claim that $\rho \circ \gamma : \mathbb{R} \to \mathrm{GL}_m(\mathbb{C})$ is also analytic.

To prove this, let $t_0 \in \mathbb{R}$, write $g_0 = \gamma(t_0)$ and pick $X_0 \in \mathfrak{su}(n)$ such that $e^{X_0} = g_0$. Then, because $d\exp_{X_0}$ is invertible (see proposition VI.3.3), there exist neighbourhoods U (resp. V) of X_0 (resp. g_0) in $\mathfrak{su}(n)$ (resp. $\mathrm{SU}(n)$) such that $\exp$ induces a bijection $U \xrightarrow{\sim} V$, and because $\exp$ is analytic and has invertible differential everywhere on U, the inverse $\ell : V \xrightarrow{\sim} U$ of $\exp : U \xrightarrow{\sim} V$ is also analytic. Also, by the same proof as in theorem VI.4.7, after shrinking U and V, we have $\ell(V \cap \mathrm{SU}(n)) = U \cap \mathfrak{su}(n)$.

Now choose $\varepsilon > 0$ such that $\gamma(]t_0-\varepsilon, t_0+\varepsilon[) \subset V$. Then on $]t_0-\varepsilon, t_0+\varepsilon[$, $\rho\circ\gamma$ is equal to $\exp\circ u \circ \ell \circ \gamma$, which is analytic.

(8) Let $X, Y \in \mathfrak{su}(n)$. Consider the maps $c_1, c_2 : \mathbb{R} \to \mathrm{GL}_m(\mathbb{C})$ defined by $c_1(t) = \rho(e^{tX}e^{tY})$ and $c_2(t) = \rho(e^{tX})\rho(e^{tY})$. By (7), c_1 and c_2 are both analytic. By (6), $c_1(t) = c_2(t)$ if $|t|$ is small enough. By the identity theorem, $c_1 = c_2$, [9] so $\rho(e^X e^Y) = c_1(1) = c_2(1) = \rho(e^X)\rho(e^Y)$.

Finally, by (3), we see that ρ is a morphism of groups.

(9) In a neighbourhood of I_n in $\mathrm{SU}(n)$, ρ is equal to $\exp\circ u \circ \log$, hence it is continuous. As ρ is a morphism of groups, it is continuous everywhere.

□

Remark VI.8.3. Let G be a closed connected subgroup of $\mathrm{GL}_n(\mathbb{C})$ and $\mathfrak{g} = \mathrm{Lie}(G)$. Let $\rho : G \to \mathrm{GL}_m(\mathbb{C})$ be a continuous representation of G on $\mathbb{C}^m$, and let $u = d\rho : \mathfrak{g} \to \mathfrak{gl}_m(\mathbb{C})$. Let W be a complex subspace of $\mathbb{C}^m$. Then W is stable by G if and only if it is stable by $\mathfrak{g}$.

Proof. Suppose that W is stable by $\mathfrak{g}$. Let $X \in \mathfrak{g}$. Then $u(X)^d(W) \subset W$ for every $d \in \mathbb{Z}_{\geq 0}$, so, as W is closed in $\mathbb{C}^m$, $e^{u(X)}(W) \subset W$, i.e. $\rho(e^X)(W) \subset W$. As $\exp(\mathfrak{g})$ generates G, this shows that W is stable by G.

[9] See corollary 1.2.6 of Krantz and Parks's book [19]. Another way to prove this is to observe that c_1 and c_2, as real analytic functions on $\mathbb{R}$, both extend to complex analytic functions on a neighbourhood of $\mathbb{R}$ in $\mathbb{C}$ and to use the identity theorem for complex analytic functions.

Conversely, suppose that W is stable by G. Let $X \in \mathfrak{g}$. Then $\rho(e^{tX})(W) \subset W$ for every $t \in \mathbb{R}$, as $t^{-1}(\rho(e^{tx}) - I_n)(W) \subset W$ for every $t \neq 0$. As $X = \lim_{t\to 0} t^{-1}(\rho(e^{tX}) - I_n)$ and W is closed in $\mathbb{C}^m$, $u(X)(W) \subset W$.

□

Corollary VI.8.4. *We have bijections*

$$\mathrm{Hom}_{\mathbb{C}\text{–}\textit{Lie alg}}(\mathfrak{sl}_n(\mathbb{C}), \mathfrak{gl}_m(\mathbb{C})) \xrightarrow{\sim} \mathrm{Hom}_{\mathbb{R}\text{–}\textit{Lie alg}}(\mathfrak{su}(n), \mathfrak{gl}_m(\mathbb{C})) \xleftarrow{\sim} \mathrm{Hom}_{Gr,cont}(\mathrm{SU}(n), \mathrm{GL}_m(\mathbb{C})),$$

where the first map is given by restriction along the inclusion $\mathfrak{su}(n) \subset \mathfrak{sl}_n(\mathbb{C})$ and the second map is $\rho \longmapsto d\rho$.

In particular, using corollary V.3.1.7 of chapter V, we get that every representation of $\mathfrak{sl}_n(\mathbb{C})$ or $\mathfrak{su}(n)$ on a finite-dimensional complex vector space is semisimple.

Proof. The only thing that we have not yet proved is the fact that the first map is bijective. This follows from the fact that the obvious map $\mathfrak{su}(n) \otimes_{\mathbb{R}} \mathbb{C} \to \mathfrak{sl}_n(\mathbb{C})$, $X \otimes a \longmapsto aX$, is an isomorphism.

Indeed, this map is surjective because if $X \in \mathfrak{sl}_n(\mathbb{C})$, then $X = \frac{1}{2}(X - X^*) + \frac{i}{2}((X + X^*)/i)$ with $X - X^*, (X - X^*)/i \in \mathfrak{su}(n)$, and then it is an isomorphism because $\dim_{\mathbb{R}}(\mathfrak{su}(n)) = \dim_{\mathbb{C}}(\mathfrak{sl}_n(\mathbb{C})) = n^2 - 1$.

□

VI.9 Some irreducible representations of $\mathrm{SU}(n)$

VI.9.1 The exterior algebra

Definition VI.9.1.1. Let k be a commutative ring and V be a k-module. Remember from definition VI.2.1.1 the tensor algebra $T^*V = \bigoplus_{n\geq 0} T^nV$, where $T^nV = V^{\otimes n}$ for $n \geq 1$ and $T^0V = k$. For every $n \geq 0$, let I_n be the submodule of T^nV generated by all elements of the form $v_1 \otimes \cdots \otimes v_n$ such that there exists $i \in \{1, \ldots, n-1\}$ with $v_i = v_{i+1}$. Let $I = \bigoplus_{n\geq 0} I_n$. It is clear that the product of T^*V sends $I_n \otimes T^mV$ and $T^nV \otimes I_m$ to I_{n+m}, for every $n, m \geq 0$. So I is a two-sided ideal of T^*V, and we can form the quotient $\bigwedge^* V = T^*V/I$, which is called the *exterior algebra* of V. We have $\bigwedge^* V = \bigoplus_{n\geq 0} \wedge^n V$, where $\wedge^n V = T^nV/I_n$. For every n, $\wedge^n V$ is called the nth exterior power of V.

We usually denote the product in $\bigwedge^* V$ by $\wedge$ (instead of $\otimes$).

Remark VI.9.1.2. Note that if $x, y \in V$, then $x \wedge x = y \wedge y = (x + y) \wedge (x + y) = 0$ (in $\bigwedge^* V$), so $x \wedge y = -y \wedge x$. As the symmetric group $\mathfrak{S}_n$ is generated by the transpositions $(i, i + 1)$, we conclude that, for all $v_1, \ldots, v_n \in V$ and $\sigma \in \mathfrak{S}_n$,

$$(*) \qquad v_{\sigma(1)} \wedge \cdots \wedge v_{\sigma(n)} = sgn(\sigma) v_1 \wedge \cdots \wedge v_n.$$

Let W be another k-module.

Definition VI.9.1.3. We say that a multilinear map $u : V^n \to W$ is *alternating* if, for every $v = (v_1, \dots, v_n) \in V^n$, if there exists $i \in \{1, \dots, n-1\}$ such that $v_i = v_{i+1}$, then $f(v) = 0$.

Remember that $\mathrm{Hom}_k(T^nV, W)$ is the k-module of multilinear maps $V^n \to W$. By the very definition of $\wedge^n V$, $\mathrm{Hom}_k(\wedge^n V, W) \subset \mathrm{Hom}_k(T^nV, W)$ is the submodule of alternating maps.

Now suppose that V is free of finite rank over k, and choose a basis $(e_1, \dots, e_d)$ of V.

Proposition VI.9.1.4. *For every $n \geq 0$, the k-module $\wedge^n V$ is free of finite rank, and a basis of $\wedge^n V$ is given by the family $e_{i_1} \wedge \cdots \wedge e_{i_n}$, with $1 \leq i_1 < \cdots < i_n \leq d$.*

In particular, $\wedge^n V = 0$ if $n \geq d+1$.

Proof. Fix $n \geq 0$, let A be the set of $(i_1, \dots, i_n) \in \mathbb{Z}^d$ such that $1 \leq i_1 < \cdots < i_n \leq d$, and, for every $\alpha = (i_1, \dots, i_n) \in A$, let $e_\alpha = e_{i_1} \wedge \cdots \wedge e_{i_n}$.

(A) The family $(e_\alpha)_{\alpha \in A}$ is generating : We know that the family $(e_{i_1} \otimes \cdots \otimes e_{i_n})_{i_1, \dots, i_n \in \{1, \dots, d\}}$ is a basis of T^nV, so the family $(e_{i_1} \wedge \cdots \wedge e_{i_n})_{i_1, \dots, i_n \in \{1, \dots, d\}}$ generates $\wedge^n V$. By formula (*) above, the family $(e_{i_1} \wedge \cdots \wedge e_{i_n})_{i_1, \dots, i_n \in \{1, \dots, d\}, i_1 \leq \cdots \leq i_n}$ also generates $\wedge^n V$. But by the definition of $\wedge^n V$, $e_{i_1} \wedge \cdots \wedge e_{i_n} = 0$ if there exists $r \in \{1, \dots, n-1\}$ such that $i_r = i_{r+1}$. So we are left with only the e_α, $\alpha \in A$.

(B) The family $(e_\alpha)_{\alpha \in A}$ is linearly independent : Let $(e_1^*, \dots, e_d^*)$ be the basis of V^* dual to $(e_1, \dots, e_d)$. Let $\alpha = (i_1, \dots, i_n) \in A$. We define a multilinear map $e_\alpha^* : V^n \to k$ by the following formula : For every $(v_1, \dots, v_n) \in V^n$,

$$e_\alpha^*(v_1, \dots, v_n) = \sum_{\sigma \in \mathfrak{S}_n} sgn(\sigma) \prod_{r=1}^{n} e_{i_r}^*(v_{\sigma(r)}).$$

This is obviously multilinear, and I claim that it is alternating. Indeed, let $(v_1, \dots, v_n) \in V^n$, suppose that we have $r \in \{1, \dots, n-1\}$ such that $v_r = v_{r+1}$, and let τ be the transposition $(r, r+1) \in \mathfrak{S}_n$. Choose a subset S of $\mathfrak{S}_n$ such that $\mathfrak{S}_n = S \sqcup S\tau$. For every $\sigma \in \mathfrak{S}_n$, we have

$$\prod_{s=1}^{n} e_{i_s}^*(v_{\sigma(s)}) = \prod_{s=1}^{n} e_{i_s}^*(v_{\sigma\tau(s)}),$$

hence

$$e_\alpha^*(v_1, \dots, v_n) = \sum_{\sigma \in S} sgn(\sigma) \prod_{s=1}^{n} e_{i_s}^*(v_{\sigma(s)}) - \sum_{\sigma \in S} sgn(\sigma) \prod_{s=1}^{n} e_{i_s}^*(v_{\sigma(s)}) = 0.$$

As e_α^* is alternating, it gives a linear map $\wedge^n V \to k$, that we still denote by e_α^*. Now note that by definition of e_α^*, we have $e_\alpha^*(e_\alpha) = 1$ and $e_\alpha^*(e_\beta) = 0$ for every $\beta \in A - \{\alpha\}$.

This shows that the family $(e_\alpha)_{\alpha \in A}$ is linearly independent.

□

Remark VI.9.1.5. Suppose that $\mathfrak{g}$ is a Lie algebra over k and that we have a representation of $\mathfrak{g}$ on V. Then the induced representation on T^nV (defined in remark VI.5.6) obviously has I_n as a subrepresentation, so we get a representation of $\mathfrak{g}$ on $\wedge^n V$.

VI.9.2 Exterior power representations

Definition VI.9.2.1. The *standard representation* of $\mathfrak{sl}_n(\mathbb{C})$ is by definition the inclusion $\mathfrak{sl}_n(\mathbb{C}) \subset \mathfrak{gl}_n(\mathbb{C})$. The corresponding representation of $\mathrm{SU}(n)$ is the inclusion $\mathrm{SU}(n) \subset \mathrm{GL}_n(\mathbb{C})$, and it is also called the standard representation. For every $d \in \mathbb{Z}_{\geq 0}$, consider the induced representations of $\mathrm{SU}(n)$ and $\mathfrak{sl}_n(\mathbb{C})$ on $E_d := \wedge^d(\mathbb{C}^n)$.

Note that the standard representation of $\mathfrak{sl}_n(\mathbb{C})$ also integrates to a continuous representation of $\mathrm{SL}_n(\mathbb{C})$ (given by the inclusion $\mathrm{SL}_n(\mathbb{C}) \subset \mathrm{GL}_n(\mathbb{C})$), so the representation $\mathrm{SU}(n)$ on E_d extends to a continuous representation of $\mathrm{SL}_n(\mathbb{C})$, whose differential is the representation of $\mathfrak{sl}_n(\mathbb{C})$ that we just defined.

Proposition VI.9.2.2. *For every $0 \leq d \leq n$, the representation E_d is irreducible.*

Remark VI.9.2.3. We know that $E_d = 0$ for $d > n$.

Proof. Let $(e_1, \ldots, e_n)$ be the standard basis of $\mathbb{C}^n$. For $1 \leq i \leq n-1$, we denote by X_i the matrix in $M_n(\mathbb{C})$ defined by

$$X_i e_j = \begin{cases} e_i & \text{if } j = i+1 \\ 0 & \text{otherwise} \end{cases}$$

(that is, X_i is the elementary matrix often denoted by $E_{i,i+1}$). Then $X_i \in \mathfrak{sl}_n(\mathbb{C})$ (because $\mathrm{Tr}(X_i) = 0$). If $1 \leq i_1 < \cdots < i_d \leq n$, then

$$X_i(e_{i_1} \wedge \cdots \wedge e_{i_d}) = \begin{cases} e_{i_1} \wedge \cdots \wedge e_{i_{r-1}} \wedge e_{i_r - 1} \wedge e_{i_{r+1}} \wedge \cdots \wedge e_{i_d} & \text{if } i = i_r - 1 \text{ and } i_{r-1} < i_r - 1 \\ 0 & \text{otherwise.} \end{cases}$$

So if

$$X = (X_d X_{d+1} \ldots X_{n-1})(X_{d+1} X_{d+2} \ldots X_{n-2}) \ldots (X_1 X_2 \ldots X_{n-d}),$$

where we take the product in the universal enveloping algebra U of $\mathfrak{sl}_n(\mathbb{C})$ (which also acts on E_d), then, for $1 \leq i_1 < \cdots < i_d \leq n$,

$$X(e_{i_1} \wedge \ldots e_{i_d}) = \begin{cases} e_1 \wedge \cdots \wedge e_d & \text{if } i_r = n - d + r \text{ for every } r \\ 0 & \text{otherwise.} \end{cases}$$

Now let's prove that E_d is irreducible. Let V be a nonzero subrepresentation of E_d, choose $v \in V - \{0\}$, write

$$v = \sum_{1 \leq i_1 < \cdots < i_d \leq n} a_{i_1, \ldots, i_d} e_{i_1} \wedge \cdots \wedge e_{i_d},$$

and choose $1 \leq j_1 < \cdots < j_d \leq n$ such that $a_{j_1,\ldots,j_d} \neq 0$. Let $g \in \mathrm{SL}_n(\mathbb{C})$ such that, for every $r \in \{1,\ldots,d\}$, $g(e_{j_r}) = \pm e_{n-d+r}$. Then $gv \in V$, and the coefficient of $e_{n-d+1} \wedge \cdots \wedge e_n$ in gv is nonzero, so we may assume that $j_r = n - d + r$. We may also assume that $a_{n-d+r,\ldots,n-d} = 1$. If we apply X to v, we also get an element of V. By the calculation above, $Xv = e_1 \wedge \cdots \wedge e_d$, so $e_1 \wedge \cdots \wedge e_d \in V$. Now take $1 \leq i_1 < \cdots < i_d \leq n$. Pick an element g of $\mathrm{SL}_n(\mathbb{C})$ such that $g(e_r) = \pm e_{i_r}$ for every $r \in \{1,\ldots,d\}$. Then $g(e_1 \wedge \cdots \wedge e_d) = \pm e_{i_1} \wedge \cdots \wedge e_{i_d} \in V$. We have seen that V contains a basis of E_d, so $V = E_d$.

□

VI.10 Characters

We now introduce a new incarnation of the character of a representation.

Definition VI.10.1. If G is a topological group, let $R_c(G)$ be the quotient of the free module on the generators $[V]$, where V is a continuous representation of G on a finite-dimensional $\mathbb{C}$-vector space, by the relations $[V] = [V'] + [V'']$, for every exact sequence $0 \to V' \to V \to V'' \to 0$ of continuous finite-dimensional representations of G. We put a multiplication on $R_c(G)$ by setting $[V][W] = [V \otimes W]$. This is a commutative ring, called the *(continuous) representation ring* of G.

The proof of the following proposition is exactly the same as the proof of the similar proposition I.4.4 of chapter I for finite length modules over a ring.

Proposition VI.10.2. *As in definition V.5.1, let's denote by $\widehat{G}$ the set of isomorphism classes of continuous irreducible representations of G on finite-dimensional $\mathbb{C}$-vector spaces.*

Then $R_c(G)$ is the free $\mathbb{Z}$-module on the $[V]$, $V \in \widehat{G}$.

Let T_c be the subgroup of diagonal matrices in $\mathrm{SU}(n)$. So T_c is canonically isomorphic to $\{(u_1,\ldots,u_n) \in \mathrm{U}(1)^n | u_1 \ldots u_n = 1\}$. We make the group $W := \mathfrak{S}_n$ act on T_c by $\sigma(u_1,\ldots,u_n) = (u_{\sigma(1)},\ldots,u_{\sigma(n)})$.

Let $D = \{(a,\ldots,a) \in \mathbb{Z}^n,\ a \in \mathbb{Z}\}$, and let $X^* = \mathbb{Z}^n/\mathbb{Z}$. We make W act on $\mathbb{Z}^n$ by $\sigma(a_1,\ldots,a_n) = (a_{\sigma(1)},\ldots,a_{\sigma(n)})$. Note that W preserves D, hence we get an action of W on X^*.

Note that, as T_c is a commutative compact group, $\widehat{T_c}$ [10] is just the set of continuous group morphisms $T_c \to \mathbb{C}^\times$ by problem VII.5.5.

Then the usual multiplication of functions makes $\widehat{T_c}$ a commutative (discrete) group.

[10] The set of isomorphism classes of continuous irreducible representations of T_c, see definition V.5.1.

Define a map $\mathbb{Z}^n \to \widehat{T}_c$ by sending $(a_1, \dots, a_n) \in \mathbb{Z}^n$ to $(u_1, \dots, u_n) \longmapsto u_1^{a_1} \dots u_n^{a_n}$. This is a morphism of groups and it sends D to 1, so it descends to a morphism of groups $\iota : X^* \to \widehat{T}_c$, which is clearly an isomorphism. Also, for every $a \in X^*$, $u \in T_c$ and $\sigma \in W$,

$$\iota(\sigma a)(u) = \iota(a)(\sigma u).$$

Set $G = \mathrm{SU}(n)$.

For every continuous finite-dimensional representation V of G, the restriction of V to T_c is a continuous finite-dimensional representation of T_c, hence a direct sum of elements of $\widehat{T}_c$. Using the isomorphism $\widehat{T}_c \simeq X^*$, we can see $\chi_{V|T_c}$ as an element of $\mathbb{Z}[X^*]$ (the group algebra of X^*) with nonnegative coefficients. This construction goes to the quotient in $R_c(G)$ and induces a morphism of groups $\chi : R_c(G) \to \mathbb{Z}[X^*]$.

Theorem VI.10.3. *(i) χ is a morphism of rings.*

(ii) χ is injective.

(iii) The image of χ is contained in $\mathbb{Z}[X^*]^W$, where W acts on $\mathbb{Z}[X^*]$ through its action on X^* (and $\mathbb{Z}[X^*]^W$ is the space of invariants of W).[11]

Proof. Point (i) just follows from the formula $\chi_{V\otimes W} = \chi_V \chi_W$ (see proposition II.1.1.3 of chapter II).

Point (ii) follows from the Schur orthogonality relations (corollary V.3.3.3 of chapter V) as in the case of finite groups. Indeed, these relations imply that $(\chi_V)_{V\in\widehat{G}}$ is an orthonormal family in $L^2(G)$, hence linearly independent. Note that all the elements in this family are in $\mathscr{C}(G)^G$ (where $\mathscr{C}(G)$ is the set of continuous functions $G \to \mathbb{C}$, and G acts on $\mathscr{C}(G)$ by $(g \cdot f)(x) = f(gxg^{-1})$). As every element of G is conjugate in G to an element of T_c,[12] the restriction map $\mathscr{C}(G) \to \mathscr{C}(T_c)$ induces an injection $\mathscr{C}(G)^G \to \mathscr{C}(T_c)$, so the family $(\chi_{V|T_c})_{V\in\widehat{G}}$ is linearly independent in $\mathscr{C}(T_c)$. Now using the linear independence of characters of irreducible representations of T_c, which follows from the Schur orthogonality relations, and the isomorphism $\iota : X^* \xrightarrow{\sim} \widehat{T}_c$, we can identify $\mathbb{Z}[X^*]$ to a subring of $\mathscr{C}(T_c)$, and all the $\chi_{V|T_c}$, $V \in \widehat{G}$, are in this subring, and of course they still form a linearly independent family. As $\chi_{V|T_c} \in \mathbb{Z}[X^*]$ is just $\chi([V])$, and as $([V])_{V\in\widehat{G}}$ is a basis of $R_c(G)$ over $\mathbb{Z}$, this gives the result.

To prove (iii), take a finite-dimensional representation V of G and an element σ of $W = \mathfrak{S}_n$. Then the corresponding permutation matrix $A \in \mathrm{GL}_n(\mathbb{C})$ (defined by $A_{ij} = 1$ if $j = \sigma(i)$ and 0 otherwise) is in $U(n)$, so there exists $c \in \mathrm{U}(1)$ such that $g := cA \in G$. By definition of the permutation matrix, for every $u = \mathrm{diag}(u_1, \dots, u_n) \in T_c$, $gug^{-1} = AuA^{-1} = \sigma u$, hence

$$\chi_V(\sigma u) = \chi_V(gug^{-1}) = \chi_V(u).$$

By the compatibility of $\iota : X^* \xrightarrow{\sim} \widehat{T}_c$ with the action of W (established above), this shows that $\chi([V]) \in \mathbb{Z}[X^*]^W$.

[11] It's actually a subring.

[12] See (2) of the proof of theorem VI.8.2, but this should be a lemma.

□

Example VI.10.4. Let $(e_1, \dots, e_n)$ be the canonical basis of $\mathbb{Z}^n$, and denote by $(\overline{e}_1, \dots, \overline{e}_n)$ its image in X^*. (Not a basis anymore.) To avoid horrible confusions, for every $\lambda \in X^*$, we denote by c_λ the corresponding element of $\mathbb{Z}[X^*]$.

If V is the representation E_d of section 9.2, then

$$\chi([V]) = \sum_{1 \leq i_1 < \dots < i_d \leq n} c_{\overline{e}_{i_1} + \dots + \overline{e}_{i_d}}.$$

This just follows from proposition VI.9.1.4.

VI.11 Weights

In this section, we still take $G = \mathrm{SU}(n)$, and we use the notation of the previous section. All the representations of G are on complex vector spaces.

Definition VI.11.1. The *Bruhat order* on $\mathbb{Z}^n$ is the (partial) order relation defined by : If $a = (a_1, \dots, a_n)$ and $b = (b_1, \dots, b_n)$ are in $\mathbb{Z}^n$, then $a \preccurlyeq b$ if and only if $a_1 \leq b_1$, $a_1 + a_2 \leq b_1 + b_2$, …, $a_1 + \dots a_{n-1} \leq b_1 + \dots + b_{n-1}$ and $a_1 + \dots + a_n = b_1 + \dots + b_n$. (Note the last relation !)

This obviously goes to the quotient and induces an order relation on X^*, still denote by $\preccurlyeq$. Also, for every $\lambda_1, \lambda_2, \mu \in X^*$,

$$\lambda_1 \preccurlyeq \lambda_2 \Rightarrow \lambda_1 + \mu \preccurlyeq \lambda_2 + \mu.$$

Definition VI.11.2. Let V be a continuous finite-dimensional representation of G. Then the set of *weights* of V, denoted by $\Lambda(V)$, is the subset of $\lambda \in X^*$ such that the coefficient of c_λ in $\chi([V])$ is nonzero. That coefficient is called the *multiplicity of the weight λ in V*.

If V is irreducible, we say that $\lambda \in \Lambda(V)$ is a *highest weight of V* if it's maximal in $\Lambda(V)$ for the Bruhat order. [13]

Finally, we let $\Lambda^+ \subset X^*$ be the subset of elements that have a lift $(a_1, \dots, a_n)$ in $\mathbb{Z}^n$ such that $a_1 \geq \dots \geq a_n$. (Then this is true for every lift.)

Remark VI.11.3. The *root system* Φ of G (or $\mathrm{Lie}\, G$) is by definition the set of nonzero weights of the adjoint representation of G on $\mathrm{Lie}\, G$ (i.e. the representation of G corresponding to $\mathrm{ad} : \mathrm{Lie}(G) \to \mathfrak{gl}(\mathrm{Lie}\, G)$). It plays an important role in understanding weights in general and has nice properties. For example, every weight in Φ has multiplicity 1 and Φ contains a basis of X^*.

[13] This is not the correct definition of highest weights for general (possibly infinite-dimensional) representations of $\mathfrak{sl}_n(\mathbb{C})$, but it is equivalent to the correct definition for irreducible finite-dimensional representations. See definition VI.14.1.1 for the correct definition.

Example VI.11.4. If $V = E_d$, then the element $\overline{e}_1 + \cdots + \overline{e}_d = (1, \ldots, 1, 0, \ldots, 0)$ (with d 1's) is the biggest element of $\Lambda(V)$ (for the Bruhat order), so it's the highest weight of V. We denote this element of Λ^+ by ϖ_d. (This is called a *fundamental weight* of G or $\operatorname{Lie} G$.)

Theorem VI.11.5. *(i) For every $\lambda \in \Lambda^+$, there exists a unique (up to isomorphism) irreducible (continuous finite-dimensional) representation W_λ of $\operatorname{Lie} G$ (or G) such that λ is the highest weight of W_λ, and every irreducible representation of $\operatorname{Lie} G$ is of that form. This gives a bijection $\Lambda^+ \xrightarrow{\sim} \widehat{G}$.*

(ii) The morphism $\chi : R_c(G) \to \mathbb{Z}[X^*]^W$ is an isomorphism.

Remark VI.11.6. This theorem (mutatis mutandis) stays true in more general situations (for representations of the Lie algebra), but we have to use things like the action of the universal envelopping algebra of $\operatorname{Lie} G$ on a representation of G (and the Poincaré-Birkhoff-Witt theorem about this universal enveloping algebra) to show that every irreducible representation V has a unique highest weight, which is bigger than all the other weights of V. And then it takes quite a bit of work to construct the W_λ (and especially to show they're finite-dimensional), and then we still have to say for which λ the representation W_λ lifts to a representation of the group. But here, because everything is explicit, we can just cheat (and we've already seen that lifiting representations to the group is automatic in our case).

Proof. First let's prove that, for every $\lambda \in \Lambda^+$, there exists some irreducible representation W_λ of G that has λ as its unique highest weight. Choose a lift $(a_1, \ldots, a_n)$ of λ in $\mathbb{Z}^n$, and set $d_i = a_i - a_{i+1}$ for every $i \in \{1, \ldots, n\}$. Note that the d_i are nonnegative (because $\lambda \in \Lambda^+$) and do not depend on the lift. Consider the representation $V = E_1^{\otimes d_1} \otimes_{\mathbb{C}} \cdots \otimes_{\mathbb{C}} E_{n-1}^{\otimes d_{n-1}}$ of G. Then the weights of V are the $d_1\lambda_1 + \cdots + d_{n-1}\lambda_{n-1}$, where λ_i is a weight of E_{d_i} for every i. In particular, $\Lambda(V)$ has a biggest element, which is $d_1\varpi_1 + \cdots + d_{n-1}\varpi_{n-1} = \lambda$, and this element has multiplicity 1. Now if $V = \bigoplus_{i \in I} V_i$ is the decomposition of V into irreducible representations, then $\Lambda(V) = \bigcup_{i \in I} \Lambda(V_i)$, so one (and only one) of the V_i must have λ as a weight. We take W_λ equal to this V_i. Note that all the weights of W_λ are $\preccurlyeq \lambda$ (because this is true for elements of $\Lambda(V)$, and $\Lambda(W_\lambda) \subset \Lambda(V)$). In particular, if $\lambda \neq \mu$, we cannot have $W_\lambda \simeq W_\mu$.

For every $\lambda \in X^*$, we define an element $d_\lambda \in \mathbb{Z}[X^*]$ by

$$d_\lambda = \sum_{\sigma \in W/W_\lambda} c_{\sigma\lambda} = \sum_{\mu \in W\lambda} c_\mu,$$

where $W_\lambda = \{\sigma \in W | \sigma(\lambda) = \lambda\}$. Then d_λ is obviously in $\mathbb{Z}[X^*]^W$, and the family $(d_\lambda)_{\lambda \in \Lambda^+}$ is a basis of $\mathbb{Z}[X^*]^W$ over $\mathbb{Z}$. (Simply because Λ^+ is a set of representatives of X^*/W.)

Let R be the subgroup of $R_c(G)$ generated by the $[W_\lambda]$, $\lambda \in \Lambda^+$. This is a free group with basis $([W_\lambda])_{\lambda \in \Lambda^+}$, because $W_\lambda \not\simeq W_\mu$ if $\lambda \neq \mu$. Let φ be the restriction of the injection $\chi : R_c(G) \to \mathbb{Z}[X^*]^W$ to R. For every $\lambda \in \Lambda^+$, we have

$$\varphi([W_\lambda]) = d_\lambda + \sum_{\mu \prec \lambda} a_{\lambda\mu} d_\mu,$$

with $a_{\lambda\mu} \in \mathbb{Z}$. This means that, in the given bases of R and $\mathbb{Z}[X^*]^W$, the (infinite) matrix of φ is lower triangular with 1's one the diagonal, and this implies that φ is invertible. In particular, φ is surjective. As φ is the composition of the injective maps χ and $R \subset R_c(G)$, this means that χ is also surjective, hence an isomorphism (giving (ii)), and that $R = R_c(G)$ (giving (i)).

□

Remark VI.11.7. It follows easily from the construction of W_λ given in the theorem that we have, for all $\lambda, \mu \in \Lambda^+$,

$$W_\lambda \otimes W_\mu = W_{\lambda+\mu} \oplus \bigoplus_{\nu \prec \lambda+\mu} W_\nu^{\oplus c_\nu}.$$

Indeed, this is already true for the tensor product of the bigger representations of the form $V = E_1^{\otimes d_1} \otimes_{\mathbb{C}} \cdots \otimes_{\mathbb{C}} E_{n-1}^{\otimes d_{n-1}}$ that are used in the proof.

Remark VI.11.8. In general, it is not so easy to write the irreducible representation of G with highest weight λ explicitely. [14] But if $\lambda = (d, 0, \ldots, 0)$, then W_λ is simply the dth symmetric power of the standard representation. (See problem VII.6.14.)

VI.12 More about roots and weights

VI.12.1 Weights of infinite-dimensional representations

Let $\mathfrak{t}$ denote the subspace of diagonal matrices in $\mathfrak{g} := \mathfrak{sl}_n(\mathbb{C})$; it's a commutative Lie subalgebra, equal to $\{(x_1, \ldots, x_n) \in \mathbb{C}^n | x_1 + \cdots + x_n = 0\}$. We write $\mathfrak{t}^*$ for the dual space of $\mathfrak{t}$. Note that the obvious map $\mathrm{Lie}(T_c) \otimes_{\mathbb{R}} \mathbb{C} \to \mathfrak{t}$ (sending $X \otimes a$ to aX) is an isomorphism. So for every element ρ of $\widehat{T_c}$, the complexification of $d\rho$ is a map of Lie algebras $\mathfrak{t} \to \mathbb{C}$; as $\mathfrak{t}$ is commutative, this is just an element of $\mathfrak{t}^*$. Remember that we identified $\widehat{T_c}$ to X^* (before theorem VI.10.3). Using this, the map above sends the class of $(a_1, \ldots, a_n) \in \mathbb{Z}^n$ in X^* to the linear map $(x_1, \ldots, x_n) \longmapsto a_1x_1 + \ldots a_nx_n$ on $\mathfrak{t}$. In particular, it is injective, and we will use it to identify $\mathfrak{t}^*$ and $X^* \otimes_{\mathbb{Z}} \mathbb{C}$.

Let V be a representation of $\mathfrak{sl}_n(\mathbb{C})$ on a $\mathbb{C}$-vector space. *In this section, we do not automatically assume that representations are finite-dimensional.*

Definition VI.12.1.1. Let $\lambda \in \mathfrak{t}^*$. The *weight space* of λ in V is

$$V(\lambda) = \{v \in V | \forall X \in \mathfrak{t},\ X \cdot v = \lambda(X)v\}.$$

Any nonzero element of $V(\lambda)$ is said to be *of weight* λ. We say that λ is a *weight* of V if $V(\lambda) \neq 0$, and then its *multiplicity* is $\dim_{\mathbb{C}} V(\lambda)$.

Remark VI.12.1.2. If $\lambda \in X^*$ and V is a finite-dimensional, these definitions agree with the ones in definition VI.11.2.

[14] But see problem VII.7.4.

VI.12.2 Roots

Remember that the set Φ of roots of $\mathfrak{g}$ is the set of nonzero weights of $\mathfrak{g}$ in its adjoint representation on itself. An easy calculation shows that, for $\lambda \in \mathfrak{t}^*$,

$$\mathfrak{g}(\lambda) = \begin{cases} \mathfrak{t} & \text{if } \lambda = 0 \\ \mathbb{C}E_{ij} & \text{if } \lambda = \overline{e}_i - \overline{e}_j \text{ with } i \neq j \\ 0 & \text{otherwise,} \end{cases}$$

where $E_{ij} \in M_n(\mathbb{C})$ is the matrix with (i,j)-entry equal to 1 and all the other entries equal to 0.

So $\Phi = \{\overline{e}_i - \overline{e}_j, i \neq j\}$. The set of *positive roots* is by definition $\Phi^+ := \{\overline{e}_i - \overline{e}_j, i < j\}$ (these are the weights with weight space contained in the space of strictly upper triangular matrices), and the set of *simple roots* is $\Delta = \{\overline{e}_i - \overline{e}_{i+1}, 1 \leq i \leq n-1\}$. Note that $\Phi = \Phi^+ \sqcup (-\Phi^+)$ and that Δ is a basis of $\mathfrak{t}^*$.

If $\alpha = \overline{e}_i - \overline{e}_j \in \Phi$, we write $X_\alpha = E_{ij}$ (it's a generator of the weight space of α), $Y_\alpha = X_{-\alpha}$ and $H_\alpha = E_{ii} - E_{jj}$, and we let $\mathfrak{s}_\alpha$ be the $\mathbb{C}$-subspace of $\mathfrak{g}$ generated by X_α, Y_α and H_α. It's clear that s_α is actually a Lie subalgebra, and that it is isomorphic to $\mathfrak{sl}_2(\mathbb{C})$.

VI.13 The Weyl character formula

For every $\lambda \in \Lambda^+$, let $\chi_\lambda = \chi([W_\lambda]) \in \mathbb{Z}[X^*]^W$. We write $X^*_{\mathbb{Q}} = X^* \otimes_{\mathbb{Z}} \mathbb{Q}$ and $\rho = \frac{1}{2}\sum_{\alpha \in \Phi^+} \alpha \in X^*_{\mathbb{Q}}$. Using the isomorphism $X^* = \mathbb{Z}^n/\mathbb{Z}(1,\ldots,1)$ defined above, we get

$$\rho = (\frac{n-1}{2}, \frac{n-3}{2}, \ldots, \frac{1-n}{2}).$$

Definition VI.13.1. The *Weyl denominator* is

$$\Delta = c_\rho \prod_{\alpha \in \Phi^+} (1 - c_{-\alpha}) \in \mathbb{Z}[X^*_{\mathbb{Q}}].$$

It follows easily from the definition that Δ is not a zero divisor in $\mathbb{Z}[X^*_{\mathbb{Q}}]$ (in fact, it is invertible in a suitable "completion" of $\mathbb{Z}[X^*_{\mathbb{Q}}]$, see example VI.14.4.3 below), and that

$$\Delta = \prod_{\alpha \in \Phi^+} (c_{\alpha/2} - c_{-\alpha/2}).$$

Theorem VI.13.2. *In $\mathbb{Z}[X^*_{\mathbb{Q}}]$, we have an equality*

$$\Delta\chi_\lambda = \sum_{\sigma \in W} sgn(\sigma) c_{\sigma(\lambda+\rho)}.$$

VI.14 Proof of the Weyl character formula

This requires a little bit more knowledge of the theory of highest weights.

VI.14.1 Highest weights

Let V be a representation of $\mathfrak{g}$ (not necesarily finite-dimensional). An easy calculation shows that, if $\lambda \in \mathfrak{t}^*$ and $\alpha \in \Phi$, then $X_\alpha \cdot V(\lambda) \subset V(\alpha + \lambda)$.

Definition VI.14.1.1. We say that $v \in V$ is a *highest weight vector* if it's a weight vector of some weight (in particular, $v \neq 0$) and $X_\alpha \cdot v = 0$ for every positive root α. We say that $\lambda \in \mathfrak{t}^*$ is a *highest weight* of V if V has a highest weight vector of weight λ. Finally, we say that V is a *highest weight representation* of $\mathfrak{g}$ if there exists a highest weight vector $v \in V$ such that $V = \mathfrak{g} \cdot v$.

Let $\mathfrak{b}$ (resp. $\mathfrak{n}$) be the subspace of upper triangular (resp. strictly upper triangular) matrices in $\mathfrak{g}$. These are both Lie subalgebras, $\mathfrak{n}$ is an ideal of $\mathfrak{b}$, and the quotient $\mathfrak{b}/\mathfrak{n}$ is canonically identified to $\mathfrak{t}$. Note that the X_α for $\alpha \in \Phi^+$ form a basis of $\mathfrak{n}$. So a weight vector v of V is a highest weight vector if and only $\mathfrak{n} \cdot v = 0$. In other words, a nonwero element v of V is a highest weight vector of weight $\lambda \in \mathfrak{t}^*$ if and only if, for every $X \in \mathfrak{b}$, $X \cdot v = \lambda(X)v$, where we used the isomorphism $\mathfrak{b}/\mathfrak{n} = \mathfrak{t}$ to see λ as a Lie algbera morphism $\mathfrak{b} \to \mathbb{C}$.

This observation (and the fact that $\mathfrak{g}$ is generated by $\mathfrak{t}$ and the X_α, $\alpha \in \Phi$, and that $X_\alpha \cdot V(\lambda) \subset V(\alpha + \lambda)$ for every $\alpha \in \Phi$ and $\lambda \in \mathfrak{t}^*$) immediately implies the following result :

Proposition VI.14.1.2. *If V is a highest weight representation of highest weight λ, then V is generated by weight vectors, the weights of V are all of the form $\lambda - \sum_{\alpha \in \Phi^+} n_\alpha \alpha$ with the $n_\alpha \geq 0$, and the multiplicity of λ in V is* 1.

VI.14.2 The Poincaré-Birkhoff-Witt theorem and the Casimir element

Let $(x_1, \ldots, x_N)$ ($N = n^2 - 1$) be any basis of $\mathfrak{g}$ as a $\mathbb{C}$-vector space. By problem VII.6.15, the monomials $x_1^{r_1} \ldots x_N^{r_N}$, $r_1, \ldots, r_N \in \mathbb{Z}_{\geq 0}$, generate the universal enveloping algebra $U\mathfrak{g}$ as a $\mathbb{C}$-vector space. By problem VII.6.16, these elements are actually linearly independent in $U\mathfrak{g}$ (because their images in $U\mathfrak{gl}_n(\mathbb{C})$ are linearly independent), so they form a basis of $U\mathfrak{g}$. This fact, which is true for a general Lie algebra over any field (or even over a commutative ring, as long as we assume that the Lie algebra is free as a module over this ring) is called the *Poincaré-Birkhoff-Witt theorem* and proved, for example, in theorem 4.3 of chapter III of part I of Serre's book [31].

Let $\mathfrak{n}^-$ be the subspace of strictly lower triangular matrices in $\mathfrak{g}$. If we apply the Poincaré-Birkhoff-Witt theorem to the basis $\{Y_\alpha, \alpha \in \Phi^+\} \cup \{H_\alpha, \alpha \in \Delta\} \cup \{X_\alpha, \alpha \in \Phi^+\}$ of $\mathfrak{g}$, ordered so that the elements of the first set (who form a basis of $\mathfrak{n}^-$) are smaller than those of the two other sets (whose union forms a basis of $\mathfrak{b}$), then we see that we have an isomorphism of $\mathbb{C}$-vector spaces $U\mathfrak{g}/U\mathfrak{b} \simeq U\mathfrak{n}^-$.

Definition VI.14.2.1. The *Casimir element* of $\mathfrak{g}$ is the element c of $U\mathfrak{g}$ defined by

$$c = \frac{1}{2}\sum_{\alpha\in\Delta} H_\alpha^2 + \sum_{\alpha\in\Phi} X_\alpha Y_\alpha.$$

The following fact can be checked by a direct calculation (see problem VII.6.21).

Proposition VI.14.2.2. *The element c is central in $U\mathfrak{g}$.*

In particular, by Schur's lemma, the Casimir element acts by a scalar on every irreducible finite-dimensional representation of $\mathfrak{g}$.

Here is another result that makes the Casimir element very useful :

Proposition VI.14.2.3. *Let V be a representation of $\mathfrak{g}$ (not necessarily finite-dimensional) and v be a highest weight vector of V of weight $\lambda \in \mathfrak{t}^*$. Then*

$$c \cdot v = ((\lambda + \rho, \lambda + \rho) - (\rho, \rho))v,$$

where $\rho = \frac{1}{2}\sum_{\alpha\in\Phi^+} \alpha$ as before and $(.,.)$ is the symmetric bilinear map $\mathfrak{t}^* \times \mathfrak{t}^* \to \mathbb{C}$ corresponding to the quadratic form $(\lambda_1, \ldots, \lambda_n) \longmapsto \frac{1}{2}\sum_{i=1}^{n-1}(\lambda_i - \lambda_{i+1})^2$. (Remember that we have identified $\mathfrak{t}^*$ to the quotient $\mathbb{C}^n/\mathbb{C}(1, \ldots, 1) = X^* \otimes_{\mathbb{Z}} \mathbb{C}$ by making $(\lambda_1, \ldots, \lambda_n)$ correspond to the linear map $(x_1, \ldots, x_n) \longmapsto \lambda_1 x_1 + \ldots \lambda_n x_n$.)

Proof. Let $(\lambda_1, \ldots, \lambda_n) \in \mathbb{C}^n$ be a representative of λ.

If $\alpha = \overline{e}_i - \overline{e}_{i+1} \in \Delta$, then $H_\alpha^2 \cdot v = (\lambda_i - \lambda_{i+1})^2$.

If $\alpha \in -\Phi^+$, then $Y_\alpha \in \mathfrak{n}$, so $Y_\alpha \cdot v = 0$ and $X_\alpha Y_\alpha \cdot v = 0$.

If $\alpha = \overline{e}_i - \overline{e}_j \in \Phi^+$, then $X_\alpha \cdot v = 0$, so

$$X_\alpha Y_\alpha \cdot v = [X_\alpha, Y_\alpha] \cdot v - Y_\alpha X_\alpha \cdot v = H_\alpha \cdot v = (\lambda_i - \lambda_j)v.$$

So we get

$$c \cdot v = \left(\frac{1}{2}\sum_{i=1}^{n-1}(\lambda_i - \lambda_{i+1}) + \sum_{1\le i<j\le n}(\lambda_i - \lambda_j)\right) v,$$

which is the desired result.

□

VI.14.3 Verma modules

Definition VI.14.3.1. Let $\lambda \in \mathfrak{t}^*$. We see λ as a Lie algebra map $\mathfrak{b} \to \mathbb{C}$ by using the isomorphism $\mathfrak{b}/\mathfrak{n} = \mathfrak{t}$ as before. This gives a representation of $\mathfrak{b}$, hence also of $U\mathfrak{b}$, on $\mathbb{C}$, which we 'll denote by $\mathbb{C}_\lambda$. As $U\mathfrak{b}$ is a subalgebra of $U\mathfrak{g}$, we can see $U\mathfrak{g}$ as a right $U\mathfrak{b}$-module in an obvious way. The *Verma module* of highest weight λ is

$$V_\lambda = U\mathfrak{g} \otimes_{U\mathfrak{b}} C_\lambda.$$

It's a left $U\mathfrak{g}$-module, hence also a representation of $\mathfrak{g}$.

Proposition VI.14.3.2. *The representation V_λ is a highest weight representation of $\mathfrak{g}$ of highest weight λ.*

Let $v \neq 0$ be vector of V_λ of a weight λ (we know that v is unique up to scaling by the first sentence). If we chose an ordering $\alpha_1, \ldots, \alpha_m$ of Φ^+, then a basis of V_λ (as a $\mathbb{C}$-vector space) is given by the $Y_{\alpha^1}^{r_1} \ldots Y_{\alpha_m}^{r_m} v$ with $r_1, \ldots, r_m \in \mathbb{Z}_{\geq 0}$, and the vector $Y_{\alpha^1}^{r_1} \ldots Y_{\alpha_m}^{r_m} v$ has weight $\lambda - (r_1\alpha_1 + \ldots r_m\alpha_m)$.

Proof. The vector $1 \otimes 1 \in U\mathfrak{g} \otimes_{U\mathfrak{b}} \mathbb{C}_\lambda = V_\lambda$ is clearly a highest weight vector of weight λ, unless it is 0. It also generates the $U\mathfrak{g}$-module V_λ, so it cannot be 0, because $V_\lambda \neq 0$. This proves the first sentence. The rest follows from the Poincaré-Birkhoff-Witt theorem, applied to the same basis of $\mathfrak{g}$ as in VI.14.2.

□

Proposition VI.14.3.3. *Let V be a highest weight representation of $\mathfrak{g}$ of highest weight λ. Then we have a surjective $\mathfrak{g}$-equivariant map $V_\lambda \to V$.*

It's easy to see that this map is unique up to scaling. (Using the fact that λ has multiplicity 1 in V.) So in a way the Verma module is the universal highest weight representation of highest weight λ.

Proof. Let v be a weight λ vector of V. By the discussion of highest weights in VI.14.1, $\mathfrak{b}$ acts on v through the map $\lambda : \mathfrak{b} \to \mathbb{C}$, so we have a $U\mathfrak{b}$-linear map $\mathbb{C}_\lambda \to V$ sending 1 to v. This extends to a $U\mathfrak{g}$-linear map $V_\lambda \to V$ by the universal property of the tensor product, and this map is surjective because $V = \mathfrak{g} \cdot v$.

□

In particular, if $\lambda \in \Lambda^+$, we get a surjective map $V_\lambda \to W_\lambda$. In fact :

Proposition VI.14.3.4. *(i) If V is a highest weight representation of $\mathfrak{g}$, then it has a unique irreducible quotient.*

(ii) If V is an irreducible highest weight representation of highest weight $\lambda \in \mathfrak{t}^*$, then it is isomorphic to the quotient of V_λ given by (i).

We denote the unique irreducible quotient of V_λ by W_λ, even when $\lambda \notin \Lambda^+$. By (ii), this does not conflict with the notation introduced previously in the case $\lambda \in \Lambda^+$.

In fact, this construction generalizes to other semisimple Lie algebras and gives a way to construct the irreducible highest weight representations. The hard part in general is showing that W_λ is finite-dimensional if and only if $\lambda \in \Lambda^+$.

Proof. Note that, by the previous two propositions, V (hence also all its subquotients) is generated by weight vectors.

Let λ be the highest weight of V, and let W be the sum of all the subrepresentations of V that don't contain a vector of weight λ. If W' is any proper subrepresentations of V, it's generated by weight vectors by the observation above, and none of the weights of W' is λ (because λ has multiplicity 1 in V), and so $W' \subset W$. So W is actually the sum of all the proper subrepresentations of V.

This implies easily that V/W is irreducible. Indeed, let Z be a proper subrepresentation of V/W. Then so the inverse image of Z in V is a proper subrepresentation, hence contained in W, so $Z = 0$.

Now let W' be another subrepresentation of V such that V/W' is irreducible. Then $V/W' \neq 0$, so W' is proper, so $W' \subset W$. As V/W' is irreducible, this implies that $W' = W$ (otherwise W/W' would be a nonzero proper subrepresentation).

Finally, let's prove (ii). If V is as in (ii), then, by the previous proposition, there exists a surjective $\mathfrak{g}$-equivariant map $V_\lambda \to V$, so V is isomorphic to an irreducible quotient of V_λ. Then the uniqueness in (i) implies the conclusion.

□

VI.14.4 Characters of Verma modules

If V is a (possibly infinite-dimensional) representation of $\mathfrak{g}$, we would like to define its character as $\sum_{\lambda \in \mathfrak{t}^*} \dim(V(\lambda)) c_\lambda$, which would recover the definition of $\chi([V])$ for finite-dimensional representations. But that sum won't be in $\mathbb{Z}[X^*]$ or even $\mathbb{Z}[\mathfrak{t}^*]$ in general because the weight spaces $V(\lambda)$ could be infinite-dimensional, and the sum could also be infinite. We can make the first problem go away by imposing conditions on V (for example, that it be a highest weight representation), and we make the second problem go away by enlarging the target ring.

First, let's extend the Bruhat order from X^* to $\mathfrak{t}^*$.

Definition VI.14.4.1. If $\lambda, \mu \in \mathfrak{t}^*$, we say that $\lambda \preccurlyeq \mu$ if $\mu - \lambda = \sum_{\alpha \in \Phi^+} n_\alpha \alpha$, with the n_α in

$\mathbb{Z}_{\geq 0}$. (This is of course a very restrictive condition.) It's easy enough to see that this gives back the previous definition if λ and μ are in X^*.

Definition VI.14.4.2. For every $\lambda \in \mathfrak{t}^*$, let

$$C_\lambda = \{\mu \in \mathfrak{t}^* | \mu \preccurlyeq \lambda\} = \{\lambda - \sum_{\alpha \in \Phi^+} n_\alpha \alpha,\ n_\alpha \in \mathbb{Z}_{\geq 0}\}.$$

We define A to be the set of formal sums $\sum_{\lambda \in \mathfrak{t}^*} a_\lambda c_\lambda$, where $a_\lambda \in \mathbb{Z}$, such that there exists $\lambda_1, \ldots, \lambda_r \in \mathfrak{t}^*$ such that, if $\lambda \notin C_{\lambda_1} \cup \cdots \cup C_{\lambda_r}$, then $a_\lambda = 0$. This contains $\mathbb{Z}[X^*]$, and it's easy to see that the formulas defining the addition and multiplication on $\mathbb{Z}[X^*]$ still make sense for elements of A, and that this makes A into a commutative ring.

Example VI.14.4.3. If $\alpha \in \Phi^+$, then $\sum_{r \geq 0} c_{-r\alpha}$ is an element of A. As it is obviously the inverse of $1 - c_{-\alpha}$, this show that $1 - c_{-\alpha}$ is invertible in A for every $\alpha \in \Phi^+$, hence so is the Weyl denominator Δ introduced in section 13. In particular, this gives a proof of the fact that Δ is not a zero divisor in $\mathbb{Z}[X^*_{\mathbb{Q}}]$.

Proposition VI.14.4.4. *(i) Let V be a highest weight representation of $\mathfrak{g}$. Then $\chi_V := \sum_{\lambda \in \mathfrak{t}^*} \dim_{\mathbb{C}}(V(\lambda)) c_\lambda$ is an element of A.*

(ii) If $V = V_\lambda$, then

$$\chi_V = c_\lambda \prod_{\alpha \in \Phi^+} (1 - c_{-\alpha})^{-1} = \Delta^{-1} c_{\lambda + \rho}.$$

Note that we are asserting in particular that all the weight spaces of V are finite-dimensional.

Definition VI.14.4.5. If V is a highest weight representation of $\mathfrak{g}$, we call the χ_V defined above the *character* of V. (Hence the notation.)

Remark VI.14.4.6. The character χ_V does not determine V in general, because V has no reason to be a semisimple representation.

Proof of the proposition. Let λ be the highest weight of V. We have seen that there exists a surjective $\mathfrak{g}$-equivariant map $V_\lambda \to V$. Since (ii) implies that $\dim_{\mathbb{C}}(V(\mu))$ is finite and equal to 0 unless $\mu \preccurlyeq \lambda$, it is enough to prove (ii).

So let's assume that $V = V_\lambda$. We have seen in proposition VI.14.3.2 that the Poincaré-Birkhoff-Witt theorem gives a basis of V_λ : Choose a highest weight vector v in V_λ, and an ordering $\alpha_1, \ldots, \alpha_m$ of Φ^+. Then the $Y_{\alpha_1}^{r_1} \ldots Y_{\alpha_m}^{r_m} v$ for $r_1, \ldots, r_m \in \mathbb{Z}_{\geq 0}$ form a basis of V_λ, and each $Y_{\alpha_1}^{r_1} \ldots Y_{\alpha_m}^{r_m} v$ is of weight $\lambda - \sum_{i=1}^m r_i \alpha_i$. This means that for every $\mu \in \mathfrak{t}^*$,

$$\dim_{\mathbb{C}}(V(\mu)) = |\{(r_1, \ldots, r_m) \in \mathbb{Z}_{\geq 0}^m | \mu = \lambda - (r_1\alpha_1 + \cdots + r_m\alpha_m)\}|.$$

This is precisely the coefficient of $c_{\lambda - \mu}$ in

$$\prod_{\alpha \in \Phi^+} (1 - c_{-\alpha})^{-1} = \prod_{\alpha \in \Phi^+} \sum_{r \geq 0} c_{-r\alpha},$$

which proves the result. (The second equality follows directly from the definition of Δ.)

□

VI.14.5 Jordan-Hölder series of Verma modules

We would now like to relate the characters of V_λ and W_λ, at least when $\lambda \in X^*$.

Proposition VI.14.5.1. *Let V be a highest weight representation of $\mathfrak{g}$ of highest weight λ. Suppose that $\lambda \in X^*$.*

Then V has a filtration $V = V_0 \supset V_1 \supset V_2 \supset \dots$ such that, for every r, V_r/V_{r+1} is of the form W_μ, with $\mu \preccurlyeq \lambda$ and $(\mu+\rho, \mu+\rho) = (\lambda+\rho, \lambda+\rho)$. (The pairing $(.,.)$ was defined in proposition VI.14.2.3.)

Actually this result is still true without the assumption on λ, but it's a bit harder to prove. [15]

Proof. Let S be the set of $\mu \in X^*$ such that $(\mu+\rho, \mu+\rho) = (\lambda+\rho, \lambda+\rho)$. The last condition defines a compact subset of $X^* \otimes_{\mathbb{Z}} \mathbb{R}$ (since $(.,.)$ is positive definite on $X^* \otimes_{\mathbb{Z}} \mathbb{R}$). Since X^* is discrete in $X^* \otimes_{\mathbb{Z}} \mathbb{R}$, the set S is finite. Let

$$d(V) = \sum_{\mu \in S} \dim_{\mathbb{C}} V(\mu).$$

We have seen that the $V(\mu)$ are finite-dimensional, so $d(V)$ is finite. We prove the proposition by induction on $d(V)$.

If V is irreducible,we are done. Otherwise, it contains a proper nonzero $\mathfrak{g}$-subrepresentation W. Since V is generated by weight vectors, so is W, so W contains at least one highest weight vector v. [16] Let μ be the weight of v. After shrinking W, we may assume that $W = \mathfrak{g} \cdot v$, so that W is a highest weight representations of highest weight μ. Now by proposition VI.14.2.3, the Casimir element $c \in U\mathfrak{g}$ acts by $(\lambda+\rho, \lambda+\rho) - (\rho,\rho)$ on V, and by $(\mu+\rho, \mu+\rho) - (\rho,\rho)$ on W. Since $W \subset V$, $(\lambda+\rho, \lambda+\rho) = (\mu+\rho, \mu+\rho)$. Hence W and V/W are both highest weight representations, and $d(V/W)$ and $d(W)$ are both $< d(V)$. If $d(V) = 1$, this gives a contradiction and shows that V had to be irreducible (and hence we're done). If $d(V) > 1$, this shows that we can conclude by applying the induction hypothesis to V/W and W.

□

Let's write $\chi_\lambda = \chi_{W_\lambda}$ and $\chi'_\lambda = \chi_{V_\lambda}$.

Corollary VI.14.5.2. *There exist integers $a_{\lambda\mu} \in \mathbb{Z}$ such that*

$$\chi_\lambda = \chi'_\lambda + \sum_{\substack{\mu \prec \lambda \\ (\mu+\rho,\mu+\rho)=(\lambda+\rho,\lambda+\rho)}} a_{\lambda\mu} \chi'_\mu,$$

for every $\lambda \in X^*$.

[15] See section 24.2 of Humphreys's book [15].

[16] Take any weight vector v in W, say of weight μ. If $X_\alpha \cdot v = 0$ for every $\alpha \in \Phi^+$, then v is a highest weight vector and we are done. Otherwise, replace v by a nonzero $X_\alpha \cdot v$. This will be a weight vector of weight $\mu + \alpha$, and we apply the same procedure to it. This has to end after a finite number of steps, because all the weights of W are $\preccurlyeq \lambda$.

Proof. Let's write D_λ for the set of $\mu \in \mathfrak{t}^*$ such that $\mu \prec \lambda$ and $(\mu+\rho, \mu+\rho) = (\lambda+\rho, \lambda+\rho)$. We have seen in the proof of proposition VI.14.5.1 that this is a finite set if $\lambda \in X^*$. By this proposition (and the fact that λ has multiplicity 1 in V_λ), we know that there exists nonnegative integers $b_{\lambda\mu}$ such that

$$\chi'_\lambda = \chi_\lambda + \sum_{\mu \in D_\lambda} b_{\lambda\mu}\chi_\mu.$$

Inverting these relations gives the result.

□

VI.14.6 End of the proof of the Weyl character formula

Let $\lambda \in \Lambda^+$. By corollary VI.14.5.2 and the calculation of the χ'_μ in proposition VI.14.4.4, we know that there exists relatives integers a_μ such that $a_\lambda = 1$ and

$$\Delta\chi_\lambda = \sum_{\mu \in D_\lambda} a_\mu c_{\mu+\rho},$$

where D_λ is the set of $\mu \in X^*$ such that $\mu \preccurlyeq \lambda$ sand $(\mu+\rho, \mu+\rho) = (\lambda+\rho, \lambda+\rho)$. As in the proof of proposition VI.14.5.1, D_λ is finite (because it's the intersection of a compact subset and a discrete subset of $X^* \otimes_{\mathbb{Z}} \mathbb{R}$).

Let $\sigma \in W$. We already know that $\sigma(\chi_\lambda) = \chi_\lambda$. On the other hand, if $\alpha = \overline{e}_i - \overline{e}_j$ is a positive root (i.e. if $i < j$), then $\sigma(\alpha) = \overline{e}_{\sigma(i)} - \overline{e}_{\sigma(j)}$ is a root, and it's positive if and only if $\sigma(i) < \sigma(j)$. As

$$\Delta = \prod_{\alpha \in \Phi^+} (c_{\alpha/2} - c_{-\alpha/2}),$$

this shows that $\sigma(\Delta) = sgn(\sigma)\Delta$, hence $\sigma(\Delta\chi_\lambda) = sgn(\sigma)\Delta\chi_\lambda$, hence, for every $\mu \in D_\lambda$, if $\sigma(\mu+\rho) = \mu'+\rho$, then $a_\mu = sgn(\sigma)a_{\mu'}$.

In particular,

$$\Delta\chi_\lambda = \sum_{\sigma \in W} sgn(\sigma)c_{\sigma(\lambda+\rho)} + R,$$

with $R \in \mathbb{Z}[X^*_{\mathbb{Q}}]$. To finish the proof, we have to show that $R = 0$.

If $R \neq 0$, there exists $\mu \in D_\lambda$ such that $a_\mu \neq 0$ and $\mu+\rho \notin W(\lambda+\rho)$. After replacing $\mu+\rho$ by $\sigma(\mu+\rho)$ for some $\sigma \in W$, we may assume that $\mu+\rho$ is dominant and not equal to $\lambda+\rho$. As $\mu \preccurlyeq \lambda$, we can write

$$(\lambda+\rho) - (\mu+\rho) = \sum_{\alpha \in \Phi^+} n_\alpha \alpha,$$

with $n_\alpha \in \mathbb{Z}_{\geq 0}$. As $(\lambda+\rho, \lambda+\rho) = (\mu+\rho, \mu+\rho)$, this gives

$$0 = 2(\mu+\rho, \sum_\alpha n_\alpha\alpha) + (\sum_\alpha n_\alpha\alpha, \sum_\alpha n_\alpha\alpha) = 2\sum_\alpha n_\alpha(\alpha, \mu+\rho) + (\sum_\alpha n_\alpha\alpha, \sum_\alpha n_\alpha\alpha).$$

As $\mu + \rho$ is dominant, $(\mu + \rho, \alpha) \geq 0$ for every $\alpha \in \Phi^+$. So we get $(\sum_\alpha n_\alpha \alpha, \sum_\alpha n_\alpha \alpha) = 0$, hence $\sum_\alpha n_\alpha \alpha = 0$. But then $\lambda = \mu$, a contradiction.

VII Exercises

VII.1 Chapter I exercises

VII.1.1 Review of tensor products

Let M be a right R-module and N be a left R-module. Their *tensor product over* R, denoted by $M \otimes_R N$, is the quotient of the free abelian group with basis $M \times N$ by the subgroup I generated by the elements :

- $(x + x', y) - (x, y) - (x', y)$, for every $x, x' \in M$ and $y \in N$;
- $(x, y + y') - (x, y) - (x, y')$, for every $x \in M$ and $y, y' \in N$;
- $(xa, y) - (x, ay)$, for every $x \in M$, $y \in N$ and $a \in R$.

If $(x, y) \in M \times N$, we write $x \times y$ for its image in $M \otimes_R N$.

If A and B are left (resp. right) R-modules, we write $\mathrm{Hom}_R(A, B)$ for the group of R-linear morphisms from A to B. If $R = \mathbb{Z}$, we write Hom instead of Hom_R. (In that case, A and B are just abelian groups, and $\mathrm{Hom}(A, B)$ is the set of morphisms of groups from A to B.)

(1). Now let M and N be as above, and let P be an abelian group. We see $\mathrm{Hom}(N, P)$ as a right R-module by the formula : $\forall a \in R, \forall f \in \mathrm{Hom}(N, P), \forall x \in N,\ (f \cdot a)(x) = f(ax)$. We define a map $\varphi : \mathrm{Hom}_R(M, \mathrm{Hom}(N, P)) \to \mathrm{Hom}(M \otimes_R N, P)$ by setting, if $f \in \mathrm{Hom}_R(M, \mathrm{Hom}(N, P))$ and $(x, y) \in M \times N$, $\varphi(f)(x \otimes y) = f(x)(y)$.

Show that $\mathrm{Hom}(N, P)$ is indeed a right R-module and that the map φ is well-defined and an isomorphism of abelian groups.

(2). Similarly, if N is a right R-module and M is a left R-module, then $\mathrm{Hom}(N, P)$ has a natural left R-module structure (given by $(a \cdot f)(x) = f(xa)$) and we have an isomorphism $\mathrm{Hom}_R(M, \mathrm{Hom}(N, P)) = \mathrm{Hom}(N \otimes_R M, P)$.

(3). Let M and N be as in (2), let S be another ring, and suppose that N is a (S, R)-bimodule, that is, that there is a left S-module structure on N such that : $\forall a \in S, \forall x \in N, \forall b \in R,\ a(xb) = (ax)b$.

Show that $N \otimes_R M$ has a natural left S-module structure and construct, for every left S-module P, a natural left R-module structure on $\mathrm{Hom}_S(N, P)$ and an isomorphism of

abelian groups $\mathrm{Hom}_R(M, \mathrm{Hom}_S(N, P)) \xrightarrow{\sim} \mathrm{Hom}_S(N \otimes_R M, P)$.

(4). Let K be a field and n a positive integer. Set $R = M_n(K)$ (the ring of $n \times n$ matrices with entries in K) and $M = K^n$. We make $M_n(K)$ act on K^n on the left by seeing K^n as the space of $n \times 1$ matrices (= column vectors) and using matrix multiplication. Similarly, we make $M_n(K)$ act on K^n on the right by seeing K^n as the space of $1 \times n$ matrices (= row vectors) and using matrix multiplication. In that way, M becomes a left R-module and a right R-module.

Calculate $M \otimes_R M$.

Solution.

(1). Let's show that $\mathrm{Hom}(N, P)$ is a right R-module. It is clear that the map $(f, a) \longmapsto f \cdot a$ is additive in $f \in \mathrm{Hom}(N, P)$ and $a \in R$. Let $f \in \mathrm{Hom}(N, P)$ and $a, b \in R$. We have to show that $f \cdot (ab) = (f \cdot a) \cdot b$. But, for every $x \in N$,

$$(f \cdot (ab))(x) = f((ab)x) = f(a(bx)) = (f \cdot a)(bx) = ((f \cdot a) \cdot b)(x).$$

Let's show that φ is well-defined. Let $f \in \mathrm{Hom}_R(M, \mathrm{Hom}(N, P))$. The formula for $\varphi(f)$ above gives a function from $M \times N$ to P, which extends by linearity to a morphism of groups from the free abelian group with basis with basis $M \times N$ to P; let's call it F. We have to show that F is zero on the ideal I defined above. So let $x, x' \in M$, $y, y' \in N$ and $a \in R$. We have :

$$F((x + x', y) - (x, y) - (x', y)) = f(x + x')(y) - f(x)(y) - f(x')(y) = 0$$

by additivity of f,

$$F((x, y + y') - F(x, y) - F(x, y') = f(x)(y + y') - f(x)(y) - f(x)(y') = 0$$

by additivity of $f(x)$, and

$$F(xa, y) - F(x, ay) = f(xa)(y) - f(x)(ay) = (f(x) \cdot a)(y) - f(x)(ay) = 0$$

by R-linearity of f and the definition of the R-module structure on $\mathrm{Hom}(N, P)$.

It's clear that φ is additive in f.

Let's show that φ is an isomorphism by constructing its inverse, which we'll call ψ. If $g \in \mathrm{Hom}(M \otimes_R N, P)$, define $\psi(g) \in \mathrm{Hom}_R(M, \mathrm{Hom}(N, P))$ by setting, for $x \in M$ and $y \in N$, $(\psi(g)(x))(y) = g(x \otimes y)$. The map $\psi(g)(x)$ is a morphism of groups because elements of the form $(x, y + y') - (x, y) - (x, y')$ are in I, and the map $\psi(g)$ is R-linear because elements of the form $(x + x', y) - (x, y) - (x', y)$ and $(xa, y) - (x, ay)$ are in I. Also, ψ is clearly additive in g.

Let's show that φ and ψ are inverses of each other. For every $f \in \mathrm{Hom}_R(M, \mathrm{Hom}(N, P))$, we have $\psi(\varphi(f)) = f$ by definition. If $g \in \mathrm{Hom}(M \otimes_R N, P)$, then g and $\varphi(\psi(g))$ are

both additive, and equal on elements of $M \otimes_R N$ of the form $x \otimes y$; as those elements generated the group $M \otimes_R N$, $\varphi(\psi(g)) = g$.

(2). This is a particular case of part (3).

(3). Let X be the free abelian group with basis $N \times M$, so that $N \otimes_R M = X/I$. We make S act on X by $s \cdot (n, m) = (sn, m)$, if $s \in S$, $m \in M$ and $n \in N$, and extending this by additivity. This is not a left S-module structure, and I is a S-submodule (because N is a (S, R)-bimodule), so we get a left S-module structure on $N \otimes_R M$ such that $s(n \otimes m) = (sn) \otimes m$ for every $s \in S$, $m \in M$ and $n \in N$.

Now let P be a left S-module, and let's put a left R-module structure on $\mathrm{Hom}_S(N, P)$. If $f \in \mathrm{Hom}_S(N, P)$ and $r \in R$, we define $r \cdot f$ by $(r \cdot f)(x) = f(xr)$, for every $x \in N$. This respects sums, $1 \in R$ acts trivially, and, if $r_1, r_2 \in R$, $f \in \mathrm{Hom}_S(N, P)$ and $x \in N$, we have

$$((r_1 r_2) \cdot f)(x) = f(x r_1 r_2) = (r_2 \cdot f)(x r_1) = (r_1 \cdot (r_2 \cdot f))(x).$$

So we do get a left R-module structure on $\mathrm{Hom}_S(N, P)$.

Let's construct inverse isomorphisms

$$\varphi : \mathrm{Hom}_R(M, \mathrm{Hom}_S(N, P)) \xrightarrow{\sim} \mathrm{Hom}_S(N \otimes_R M, P)$$

and

$$\psi : \mathrm{Hom}_S(N \otimes_R M, P) \xrightarrow{\sim} \mathrm{Hom}_R(M, \mathrm{Hom}_S(N, P)).$$

We can use almost the same formulas as in (1). If $f \in \mathrm{Hom}_R(M, \mathrm{Hom}_S(N, P))$ and $(x, y) \in M \times N$, we set $\varphi(f)(y \otimes x) = f(x)(y)$. If $g \in \mathrm{Hom}_S(N \otimes_R M, P)$, define and $(x, y) \in M \times N$, define $\psi(g) \in \mathrm{Hom}_R(M, \mathrm{Hom}_S(N, P))$ by setting $(\psi(g)(x))(y) = g(y \otimes x)$. The verification that these are well-defined and inverses of each other is also almost the same as in (1). We set things up so that everything will be compatible with the S-actions. For example, say that we wanted to check that, for $f \in \mathrm{Hom}_R(M, \mathrm{Hom}_S(N, P))$, $\varphi(f)$ is indeed S-linear. We take $s \in S$, $x \in M$ and $y \in N$, and then

$$\varphi(f)(s(y \otimes x)) = \varphi(f)((sy) \otimes x) = f(x)(sy) = s(f(x)(y)) = s(\varphi(f)(y \otimes x)),$$

as $f(x)$ is S-linear.

(4). Things will be more clear if we write $M \otimes_R M$ as $M_{1n}(K) \otimes_{M_{nn}(K)} M_{n1}(K)$, where $M_n(K)$ acts on both sides by the matrix product. (This is the same, by the definition of the two actions of R on M.)

Let $e = \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix} \in M_{1n}(K)$ and $f = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in M_{n1}(K)$. If $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in M_n(K)$ is such that $b_1 = 1$, then there exists $X \in M_n(K)$ such that $eX = e$ and $Xb = f$, and then

$e \otimes b = e \otimes f$. As these elements b generate the K-vector space $M_{n1}(K)$, we see that $e \otimes b$ is in the line spanned by $e \otimes f$ for every $b \in M_{n1}(K)$. We show similarly that, for every $a \in M_{1n}(K)$, $a \otimes f$ is in the line spanned by $e \otimes f$. Finally, we get $\dim_K(M \otimes_R M) \leq 1$.

Consider the map $\varphi : M_{1n}(K) \otimes_{M_{nn}(K)} M_{n1}(K) \to M_{11}(K) = K$ defined by $\varphi(a \otimes b) = ab$. This map is well defined, because $(a, b) \longmapsto ab$ is additive and a and b, and because, for every $a \in M_{1n}(K)$, $b \in M_{n1}(K)$ and $X \in M_{nn}(K)$, we have $(aX)b = a(XB)$. Note also that $\varphi(e \otimes f) = 1$, so φ is surjective. As $\dim_K(M \otimes_R M) \leq 1$ and $\dim_K(K) = 1$, this implies that φ is an isomorphism.

□

VII.1.2 Some properties of projective modules

(1). If $0 \to M' \to M \to M'' \to 0$ is an exact sequence of left modules over a ring R, with M'' projective, and if N is a right R-module, prove that the sequence $0 \to N \otimes_R M' \to N \otimes_R M \to N \otimes_R M'' \to 0$ is still exact.

(2). If $0 \to M' \to M \to M'' \to 0$ is an exact sequence of left modules over a ring R, and if N is a projective right R-module, prove that the sequence $0 \to N \otimes_R M' \to N \otimes_R M \to N \otimes_R M'' \to 0$ is still exact.

In fancy terms, this is saying that projective modules are flat. [1]

Solution.

(1). As M'' is projective, the exact sequence $0 \to M' \to M \to M'' \to 0$ splits, so it is isomorphic to the exact sequence $0 \to M' \to M' \oplus M'' \to M'' \to 0$ (where the map are $\begin{pmatrix} \mathrm{id}_{M'} & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ \mathrm{id}_{M''} \end{pmatrix}$). When we tensor by N, we get the sequence $0 \to N \otimes_R M' \to (N \otimes_R M') \oplus (N \otimes_R M'') \to N \otimes_R M'' \to 0$ (with similar maps), which is obviously exact.

(2). We only need to check that the map $M' \otimes_R N \to M \otimes_R N$ is injective. (The other exactness properties are a general property of the tensor product, and are true without any condition on N.) Choose a right R-module N' such that $F := N \oplus N'$ is a free R-module. Then we have a commutative square

$$\begin{array}{ccc} M' \otimes_R N & \longrightarrow & M \otimes_R N \\ \downarrow & & \downarrow \\ M' \otimes_R F & \longrightarrow & M \otimes_R F \end{array}$$

[1] See 24.20 of Lam's book [20] for a definition.

where all the arrows except the top horizontal one are known to be injective. This implies that the top horizontal arrow is injective too.

□

VII.1.3 Division rings

We say that a ring R is a *division ring* if it is nonzero and if every nonzero element of R is invertible (that is, for every $a \in R - \{0\}$, there exists $b \in R$ such that $ab = ba = 1$). If K is a commutative ring and R is a ring, we say that R is a *K-algebra* is R is a K-module and : $\forall a, b \in K, \forall x, y \in R,\ (ax)(by) = (ab)(xy)$. If R is a division ring and a K-algebra, we also say that it is a *division algebra* over K.

(1). Let K be an algebraically closed field and R be a division algebra over K that is finite-dimensional as a K-vector space. Show that $R = K$.

(2). Give an example of a finite-dimensional noncommutative division algebra over $\mathbb{R}$.

(3). Give an example of a noncommutative division algebra over $\mathbb{C}$.

Hint : If K is a field and σ is an automorphism of K, let $K((t, \sigma))$ be the ring of Laurent series with coefficients in K, where we twist the multiplication by setting $t^n a = \sigma^n(a)t^n$, for every $n \in \mathbb{Z}$ and every $a \in K$. Show that $K((t, \sigma))$ is always a division ring (and it's a division algebra over the subfield of K composed of σ-invariant elements).

Solution.

(1). As K is a field, the map of rings $K \to R$, $\lambda \longmapsto \lambda \cdot 1$ is injective. Let's show that it is surjective. Let $a \in R$. The map $m_a : R \to R$, $x \longmapsto ax$ is K-linear (because R is a K-algebra), R is a finite-dimensional K-vector space and K is algebraically closed, so m_a has at least one eigenvalue. In other words, there exists $\lambda \in K$ such that $\mathrm{Ker}(m_a - \lambda \cdot \mathrm{id}) = \mathrm{Ker}(m_{a-\lambda}) \not= 0$. This implies that $a - \lambda$ is not invertible. As R is a division algebra, we get $a - \lambda = 0$, ie $a = \lambda \in K$.

(2). See problem VII.1.6.

(3). We take the hint. First, let $L = \{x \in K | \sigma(x) = x\}$. Then L is a subfield of K, and elements of L commute with t, so they commute with every element of $K((t, \sigma))$ (because elements of $K((t, \sigma))$ are of the form $\sum_{r \geq n} a_r t^r$,[2] for some $n \in \mathbb{Z}$ and some $a_r \in K$). So $K((t, \sigma))$ is a L-algebra.

Now we show that $K((t, \sigma))$ is a division ring. Let $f \in K((t, \sigma)) - \{0\}$, and write

[2]This sum, as well as the other sums appearing in this proof, is not assumed to have only a finite number of nonzero terms.

$f = \sum_{r \geq n} a_r t^r$, with $a_n \neq 0$. Then

$$f = \sum_{r \geq 0} a_{n+r} t^{n+r} = t^n \sum_{r \geq 0} \sigma^{-n}(a_{r+n}) t^r = t^n \sigma^{-n}(a_n) \sum_{r \geq 0} \sigma^{-n}(a_n^{-1} a_{r+n}) t^r.$$

As $t^n \sigma^{-n}(a_n)$ is invertible (its inverse is $\sigma^{-n}(a_n)^{-1} t^{-n}$, we just have to show that the second factor is invertible. That is, we may assume that $f = 1 + g$, with $g = \sum_{r \geq 1} b_r t^r$. Just like in the case of usual power series, we can show that $\sum_{m \geq 0} (-1)^m g^m$ makes sense and is the inverse of $1 + g$.

Now to find a noncommutative division algebra over $\mathbb{C}$, we apply the construction above with $K = \mathbb{C}(x, y)$ and σ defined by $\sigma(x) = y$, $\sigma(y) = x$.

□

VII.1.4 Ideals of rings of matrices

Let K be a field, n be a positive integer and $R = M_n(K)$ be the set of $n \times n$ matrices with entries in K.

(1). Give a list of left ideals of R.

(2). Which of these are ideals ?

(3). We say that $a \in R$ is a left (resp. right) zero divisor if there exists $b \in R - \{0\}$ such that $ab = 0$ (resp. $ba = 0$). Show that an element of R is a left zero divisor if and only if it's a right zero divisor.

(4). We say that $a \in R$ is left (resp. right) invertible if there exists $b \in R$ such that $ba = 1$ (resp. $ab = 1$). For $a \in R$, show that the following are equivalent :

- a is left invertible;
- a is right invertible;
- a is not a zero divisor.

(5). Which of the equivalences of (3) and (4) stay true in $M_n(\mathbb{Z})$?

Solution.

(1). For every subspace V of K^n, let

$$I_V = \{A \in M_n(K) | \forall v \in V \; Av = 0\}.$$

This is obviously a left ideal of $M_n(K)$. Moreover, the ideal I_V determines V, as $V = \bigcap_{A \in I_V} \mathrm{Ker}(A)$. Let's show that every left ideal is of that form.

So let I be a left ideal of $M_n(K)$. Let

$$V = \bigcap_{A \in I} \mathrm{Ker}(A),$$

a subspace of $M_n(K)$. By definition of V, we have $I \subset I_V$. Let's show that $I = I_V$, that is, that every matrix with kernel cnotaining V is in I. Suppose that we have shown :
$(*)$ For every $v_0, w_0 \in K^n$ and every subspace W of K^n such that $K^n = V \oplus Kv_0 \oplus W$, there exists $A \in I$ such that $W \subset \mathrm{Ker}(A)$ and $Av_0 = w_0$.

Let's show that this implies the result. Let $A \in I_V$. Let $(v_1, \dots, v_n)$ be a basis of K^n such that $(v_1, \dots, v_i)$ is a basis of V, where $i = \dim V$. By $(*)$, for every $j \geq i+1$, there exists $A_j \in I$ such that $A_j v_j = Av_j$ and $A_j v_k = 0$ for $k \neq j$. Then $A = A_{i+1} + \cdots + A_n$, so $A \in I$.

Now let's prove $(*)$. Fix v_0, w_0, W as in the statement of $(*)$. If we can find $A \in I$ such that $Av_0 \notin A(W)$, then we are done; indeed, in thet case we can find $B \in M_n(K)$ such that $BW = 0$ and $BAv_0 = w_0$, and then $BA \in I$ satisfies the conclusion of $(*)$. So let's assume that, for every $A \in I$, $Av_0 \in AW$. Let $(e_1, \dots, e_n)$ be the canonical basis of K^n and $i = \dim V + 1$. Without loss of generality, we may assume that $(e_1, \dots, e_{i-1})$ is a basis of V, $e_i = v_0$ and $(e_{i+1}, \dots, e_n)$ is a basis of W. Let $A \in I$, and let r be its rank. We can find a invertible matrix $B \in M_n(K)$ such that $(BAe_{i+1}, \dots, BAe_n) = (e_1, \dots, e_r, 0, \dots, 0)$. Write $BAv_0 = BAe_i = \sum_{j=1}^n \lambda_j(A) e_j$. Then $\lambda_j(A) = 0$ for $j > r$ and $Av_0 = \sum_{j=i+1}^n \lambda_{j-i}(A) Ae_j$ (because $BAv_0 = \sum_{j=i+1}^n \lambda_{j-i}(A) BAe_j$, and B is invertible). Now let A' be another element of I, let r' be its rank, choose $B' \in GL_n(K)$ and define the $\lambda_j(A')$ as above. We claim that, for every $s \leq \min(r, r')$, $\lambda_s(A) = \lambda_s(A')$. Indeed, fix such a s, and consider the elementary matrix $E_{s,s}$ (with entries 1 at the coordinates (s,s) and 0 everywhere else). Then

$$E_{s,s} Ae_j = \begin{cases} \lambda_s(A) e_s & \text{if } j = i \\ e_s & \text{if } j = s+i \\ 0 & \text{otherwise} \end{cases},$$

and similarly for A'. So $E_{s,s}(A - A')(e_i) = (\lambda_s(A) - \lambda_s(A'))e_i$, and $E_{s,s}(A - A')$ sends all the other e_j to 0. As $E_{s,s}(A - A') \in I$, $E_{s,s}(A - A')e_i \in E_{s,s}(A - A')W$ by our assumption. So $\lambda_s(A) = \lambda_s(A')$. This means that we can find $\lambda_1, \dots, \lambda_{n-i} \in K$ (with $\lambda_j = 0$ for $j > r$) such that, for every $C \in I$, $Cv_0 = \sum_{j=i+1}^n \lambda_{j-i} Ce_j$. But then $v_0 - (\sum_{j=i+1}^n \lambda_{j-i} e_j)$ is in the kernel of every element of I, hence in V, which is absurd. This finishes the proof of $(*)$.

(2). Only 0 and $M_n(K)$ are ideals in $M_n(K)$. To prove this, we use the notation I_V from the previous question. It follows immediately from the definition of I_V that, for every subspace V of K^n and every invertible $A \in M_n(K)$, $I_V A = I_{AV}$. So if I_V is a left ideal, then $V = AV$ for every invertible $A \in M_n(K)$. This is only possible if $V = 0$ (then $I_V = M_n(K)$) or $V = K^n$ (then $I_V = 0$).

(3). We say that $a \in R$ is a left (resp. right) zero divisor if there exists $b \in R - \{0\}$ such that $ab = 0$ (resp. $ba = 0$). Show that an element of R is a left zero divisor if and only if it's a right zero divisor.

Let $a \in M_n(K)$ Then a is a left zero divisor if and only if $\mathrm{Ker}(a) \neq 0$, and a right zero divisor if and only if the rank of a is $< n$. But that these two conditions are equivalent, and that they are also equivalent to the fact that a is not invertible.

(4). Let $a \in M_n(K)$. Then a is left invertible if and only if $\mathrm{Ker}(a) = 0$, and right invertible if and only its rank is n. We know that these two conditions are equivalent. The last equivalence is already proved in the answer of (3).

(5). If a is an element of $M_n(\mathbb{Q})$, then we can write $a = \lambda a'$, with $\lambda \in \mathbb{Q}^\times$ and $a' \in M_n(\mathbb{Z})$. From this, it follows easily that elements of $M_n(\mathbb{Z})$ are left (resp.) right zero divisors in $M_n(\mathbb{Z})$ of and only if they are left (resp. right) zero divisors in $M_n(\mathbb{Q})$. So the equivalence of (3) stays true.

Let $a \in M_n(\mathbb{Z})$. If there exists $b \in M_n(\mathbb{Z})$ such that $ab = 1$ (resp. $ba = 1$), then a is invertible in $M_n(\mathbb{Q})$, $b = a^{-1}$, and we also have $ba = 1$ (resp. $ab = 1$). So the first two conditions of (4) are still equivalent. They imply the last condition by the remark above about zero divisors, but the converse is not true. For exemple, 2 (ie twice the identity matrix) is not a zero divisor, but it is also not invertible in $M_n(\mathbb{Z})$.

□

VII.1.5 Commutative semisimple rings

We say that a ring R is *simple* if it is nonzero and its only ideals are 0 and R. We say that R is *semisimple* if for every R-module M and every R-submodule N of M, there exists another R-submodule N' of M such that $M = N \oplus N'$ (that is, such that the map $N \times N' \to M,\ (x, y) \longmapsto x + y$ is an isomorphism).

Now take R a *commutative* ring.

(1). If R is simple, show that R is a field.

(2). Assume R semisimple, nonzero and Noetherian.[3]

(a) Show that we have $R = R' \times K$ (as rings), with K a field.

(b) Show that R is a direct product of fields.

Solution.

(1). Let $a \in R - \{0\}$. Then Ra is a nonzero ideal of R, so $Ra = R$, so a is invertible.

[3]The last hypothesis is unnecessary, and is removed in theorem I.1.10.5 of chapter I.

(2). (a) Let I be an ideal of R. As R is semisimple, there exists another ideal J of R such that $R = I \oplus J$. If e_I (resp. e_J) is the image of $1 \in R$ by the obvious map $R \to R/J = I$ (resp. $R \to R/I = J$), then it is a unit element for the multiplication in I (resp. J), and we have $1 = e_I + e_J$. So I and J are commutative rings, and we have $R = I \times J$ as rings. Now if we take I to be a maximal ideal, then $J = R/I$ will be a field. Note that I is also a semisimple ring, as ideal of I are just ideals of R contained in I.

(b) By the previous question, we can construct a descending sequence of ideals $I_0 = R \supset I_1 \supset I_2 \supset \dots$ and an ascending chain of ideals $J_0 = 0 \subset J_1 \subset J_2 \subset \dots$ of R such that $R = J_i \times I_i$ for every $i \in \mathbb{N}$ and each I_i/I_{i+1} is a field or zero. As R is Noetherian, the sequence $(J_i)_{i\in\mathbb{N}}$ becomes constant, so R is the product of the nonzero I_i/I_{i+1}, which are all fields.

□

VII.1.6 The $\mathbb{R}$-algebra of quaternions

Let $\mathbb{H}$ be the $\mathbb{R}$-algebra $\mathbb{R} \oplus \mathbb{R}i \oplus \mathbb{R}j \oplus \mathbb{R}k$, with the multiplication given by :

- $i^2 = j^2 = k^2 = -1$;
- $ij = -ji = k$, $jk = -kj = i$, $ki = -ik = j$.

Note that we have an obvious embedding $\mathbb{C} = \mathbb{R} \oplus \mathbb{R}i \subset \mathbb{H}$, so $\mathbb{H}$ is a $\mathbb{C}$-vector space, but not a $\mathbb{C}$-algebra.

(1). Why is $\mathbb{H}$ not a $\mathbb{C}$-algebra ?

(2). Show that $\mathbb{H}$ is a division ring. (Hint : If $x = a + bi + cj + dk \in \mathbb{H}$ with $a, b, c, d \in R$, its conjugate is defined to be $\overline{x} = a - bi - cj - dk$. What is $x\overline{x}$?)

(3). Show that $\mathbb{H} \otimes_{\mathbb{R}} \mathbb{C} \simeq M_2(\mathbb{C})$ as $\mathbb{C}$-algebras. (You can use an embedding $\mathbb{H} \subset M_2(\mathbb{C})$ given by choosing a $\mathbb{C}$-basis of $\mathbb{H}$.)

(4). Show that $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C} \simeq \mathbb{C} \times \mathbb{C}$ as $\mathbb{C}$-algebras. (If you're getting mixed up, try giving different names to the i in the two factors $\mathbb{C}$ of the tensor product.)

Solution.

(1). If $\mathbb{H}$ was a $\mathbb{C}$-algebra, then we would have $ab = ba$ for every $a \in \mathbb{C}$ and $b \in \mathbb{H}$. This is not true (just take $a = i$ and $b = j$).

(2). If $x = a + bi + cj + dk$, then $x\overline{x} = \overline{x}x = a^2 + b^2 + c^2 + d^2$. So $x\overline{x} \in \mathbb{R}_{\geq 0}$, and it is zero if and only if $x = 0$. Now if $x \neq 0$, then $\frac{1}{x\overline{x}}\overline{x}$ is an inverse of x (on both sides).

(3). As a $\mathbb{C}$-vector space, $\mathbb{H} = \mathbb{C} \oplus \mathbb{C}j \simeq \mathbb{C}^2$. Making $\mathbb{H}$ act on itself by left multiplication, we get a $\mathbb{C}$-linear map $u : \mathbb{H} \to \mathrm{Hom}_{\mathbb{C}}(\mathbb{H}, \mathbb{H}) \simeq M_2(\mathbb{C})$, and this map

is a map of algebras because the multiplication of $\mathbb{H}$ is associative (if $a, b, x \in \mathbb{H}$, $u(ab)(x) = (ab)x = a(bx) = u(a)(u(b)(x))$).

Consider the $\mathbb{C}$-linear map $v : \mathbb{H} \otimes_R \mathbb{C} \to M_2(\mathbb{C})$ that sends $a \otimes \lambda$ to λa. This a $\mathbb{C}$-algebra map, because

$$v((a \otimes \lambda)(b \otimes \mu)) = v((ab) \otimes (\lambda\mu)) = (\lambda\mu)(ab) = (\lambda a)(\mu b)$$

(we use the fact that $M_2(\mathbb{C})$ is a $\mathbb{C}$-algebra, ie that scalar matrices commute with every other matrix).

Let's calculate the images of $1, i, j, k \in \mathbb{H}$ by u. We have

$$u(1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, u(i) = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}, \; u(j) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \text{ and } u(k) = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

These four matrices generate $M_2(\mathbb{C})$ as a $\mathbb{C}$-vector space, so v is surjective. As $\mathbb{H} \otimes_{\mathbb{R}} \mathbb{C}$ and $M_2(\mathbb{C})$ are both $\mathbb{C}$-vector spaces of dimension 4, v is an isomorphism.

(4). Let $R = \mathbb{C} \otimes_R \mathbb{C}$ and $S = \mathbb{C} \times \mathbb{C}$. We want to construct $\mathbb{C}$-algebra maps $\varphi : R \to S$ and $\psi : S \to R$.

There are two ways to see φ. First we can consider the $\mathbb{R}$-basis $(1, i)$ of the second $\mathbb{C}$ in the tensor product. Then $R = \mathbb{C} \otimes 1 \oplus \mathbb{C} \otimes i$, and we set $\varphi(a \otimes 1 + b \otimes i) = (a + ib, a - ib)$. Or we just set $\varphi(x \otimes y) = (xy, x\overline{y})$ and extend this by linearity. These clearly give the same $\mathbb{C}$-linear map, and the fact that it respects multiplication is obvious on the second description.

Define $\psi : S \to R$ by $\psi(x, y) = \frac{x+y}{2} \otimes 1 + \frac{x-y}{2i} \otimes i$. Then ψ is $\mathbb{C}$-linear, and it is clearly the inverse of φ (use the first description of φ).

□

VII.1.7 Simple modules over some commutative rings

(1). Write a list of all the simple modules over $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{C}[x]$, $\mathbb{Q}[x]$ (up to isomorphism).

(2). Let $\Gamma = \mathbb{Z}/p\mathbb{Z}$ and R be the group algebra $k[\Gamma]$, where k is a field. Write a list of all the simple modules over R (up to isomorphism).

(3). Identify the group Γ of the previous question with the subgroup $\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$ of $GL_2(\mathbb{F}_p) := M_2(\mathbb{F}_p)^\times$, and use this injection $\Gamma \subset M_2(\mathbb{F}_p)$ to make Γ act on $\mathbb{F}_p^2$. This gives $M := \mathbb{F}_p^2$ the structure of a module on $R := \mathbb{F}_p[\Gamma]$. Find a Jordan-Hölder series for M. Is M a semisimple R-module ?

Solution.

(1). Let R be a ring and M be a simple R-module. Choose $x \in M - \{0\}$. Then the map ${}_RR \to M$, $a \longmapsto ax$ is surjective (its image is a nonzero submodule of M), so $M \simeq R/I$ with I a left ideal of R. As M is simple, I is maximal. Conversely, for every maximal left ideal I of R, R/I is a simple R-module. So to list all the simple R-module up to isomorphism, we just have to find all the maximal left ideals of R.

If $R = \mathbb{Z}$, this shows that the simple R-modules are the $\mathbb{Z}/p\mathbb{Z}$ with p a prime number. If $R = \mathbb{Q}$, the only simple R-module is $\mathbb{Q}$. If $R = \mathbb{C}[x]$, the simple R-modules are all (up to isomorphism) of the form $\mathbb{C}[x]/(x-a)$, with $a \in \mathbb{C}$. Note that $\mathbb{C}[x]/(x-a)$ is the $\mathbb{C}[x]$-module $\mathbb{C}$, where x acts by multiplication by a. Finally, if $R = \mathbb{Q}[x]$, the simple R-modules are all of the form $\mathbb{Q}[x]/(f)$, with f a monic irreducible polynomial in $\mathbb{Q}[x]$.

(2). Note that $R \simeq k[x]/(x^p - 1)$. So any R-module M is also a $k[x]$-module, and the R-submodules of M are its $k[x]$-submodules; in particular, M is simple as a R-module if and only if it is simple as a $k[x]$-module. So to find the simple R-modules, we just have to find the simple $k[x]$-modules on which x^p acts as 1. By the beginning of (a) (and the fact that $k[x]$ is a PID), every simple $k[x]$-module is isomorphic to a $k[x]/(f)$ with $f \in k[x]$ monic irreducible. Note that x^p acts as 1 on $k[x]/(f)$ if and only if f divides $x^p - 1$. Finally, the simple R-modules are the $k[x]/(f)$, with f an irreducible factor of $x^p - 1$. (In particular, there are only finitely many isomorphism classes of simple R-modules.)

(3). The identification is given by $x \longmapsto \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$. Note that by (b) and the fact that $x^p - 1 = (x-1)^p$ in $\mathbb{F}_p[x]$, we know that the only simple R-module is $\mathbb{F}_p$ with the trivial action of Γ.

Let (e_1, e_2) be the canonical basis of $\mathbb{F}_p^2$. Then $M_1 := \mathbb{F}_p e_1$ is a R-submodule of M, and both M_1 and M/M_1 are simple, so we've found our Jordan-Hölder series. The R-module M is not semisimple as we cannot write $M = M_1 \oplus M_2$ with M_2 another submodule. (Otherwise, M_2 would be isomorphic to M/M_1, so M be isomorphic to $\mathbb{F}_p^2$ with the trivial action of Γ, but this is not the case.)

□

VII.1.8 Group algebra of the quaternion group

Remember the $\mathbb{R}$-division algebra $\mathbb{H}$ of problem VII.1.6. Let Q be the subgroup $\{\pm 1, \pm i, \pm j, \pm k\}$ of $\mathbb{H}^\times$, and let $R = \mathbb{R}[Q]$ and $R_{\mathbb{C}} = \mathbb{C}[Q] (= R \otimes_R \mathbb{C})$.

(1). Show that there exists a $\mathbb{R}$-algebra R' such that $R \simeq R' \times \mathbb{H}$ (as $\mathbb{R}$-algebras).

(2). Find the isotypic components of ${}_RR$, and the multiplicities of the simple constituents (= Jordan-Hölder constituents) of M.

(3). Write R as a product of simple $\mathbb{R}$-algebras.

(4). Write a list of all the simple R-modules (up to isomorphism).

(5). Write $R_{\mathbb{C}}$ as a product of simple $\mathbb{C}$-algebras and find all the simple $R_{\mathbb{C}}$-modules (up to isomorphism).

Solution.

(1). Note that Q is isomorphic to the group Γ generated by the elements c_{-1}, c_i, c_j, satisfying the relations : $c_{-1}^2 = 1$, $c_i^2 = c_j^2 = c_{-1}$, $c_{-1}c_i = c_i c_{-1}$, $c_{-1}c_j = c_j c_{-1}$, $c_i c_j = c_{-1}c_j c_i$. We get the isomorphism $\iota : Q \to \Gamma$ by setting $\iota(-1) = c_{-1}$, $\iota(i) = c_i$, $\iota(j) = c_j$, $\iota(k) = c_i c_j$ and, for $\alpha \in \{i, j, k\}$, $\iota(c_{-\alpha}) = c_{-1}\iota(c_\alpha)$. So R is isomorphic to the quotient of $\mathbb{R}\langle x_{-1}, x_i, x_j\rangle$ by the ideal generated by $x_{-1}^2 - 1$, $x_i^2 - x_{-1}$, $x_j^2 - x_{-1}$, $x_i x_{-1} - x_{-1}x_i$, $x_j x_{-1} - x_{-1}x_j$ and $x_i x_j - x_{-1}x_j x_i$. A basis of the $\mathbb{R}$-vector space R is $(1, x_{-1}, x_i, x_j, x_{-1}x_i, x_{-1}x_j, x_i x_j, x_{-1}x_i x_j)$ (simply because $R = \mathbb{R}[Q]$).

We construct a $\mathbb{R}$-algebra map $\varphi : R \to \mathbb{H}$ by sending x_{-1} to -1, x_i to i and x_j to j. This φ is obviously surjective, so its kernel is dimension 4. As $1 + x_{-1}, x_i + x_{-1}x_i, x_j + x_{-1}x_j, x_i x_j + x_{-1}x_i x_j$ are all in $\operatorname{Ker}\varphi$ and linearly independent (by the description of the basis of R above), they form a basis of $\operatorname{Ker}\varphi$ as $\mathbb{R}$-vector space, and we see also that $\operatorname{Ker}\varphi$ is the ideal of R generated by $1 + x_{-1}$ (as $1 + x_{-1}$ is central, the left (or right) ideal is generates is an ideal). Let I be the $\mathbb{R}$-subspace of R generated by $1 - x_{-1}, x_i - x_{-1}x_i, x_j - x_{-1}x_j, x_i x_j - x_{-1}x_i x_j$, then we have $R = \operatorname{Ker}\varphi \oplus I$ and I is also the ideal of R generated by $1 - x_{-1}$ (again, the left, right and two-sided ideals generated by $1 - x_{-1}$ are equal because $1 - x_{-1}$ is central). So we have written $R = \operatorname{Ker}\varphi \oplus I$ with $\operatorname{Ker}\varphi$ and I ideals of R, which implies that $R' := \operatorname{Ker}\varphi$ and I are rings and that $R = R' \times I$ as rings, by remark I.1.3.16 in chapter I. (Also, these are obviously $\mathbb{R}$-subalgebras of R, because they are $\mathbb{R}$-subspaces.) It remains to notice that $I \simeq R/\operatorname{Ker}\varphi \simeq \mathbb{H}$.

(2). We already know that $\mathbb{H}$ is a simple R-algbera, because it's a division algebra. So it remains to decompose R'. Note that $R' = R/(x_{-1} - 1)$, so $R' \simeq \mathbb{R}[t_i, t_j]/(t_i^2 - 1, t_j^2 - 1) \simeq \mathbb{R}[t_i]/(t_i^2 - 1) \otimes_{\mathbb{R}} \mathbb{R}[t_j]/(t_j^2 - 1)$. By the Chinese remainder theorem, $\mathbb{R}[t]/(t^2 - 1) \simeq \mathbb{R}[t]/(t + 1) \times \mathbb{R}[t]/(t - 1) \simeq \mathbb{R} \times \mathbb{R}$. So we finally get

$$R' \simeq (\mathbb{R} \times \mathbb{R}) \otimes_{\mathbb{R}} (\mathbb{R} \times \mathbb{R}) \simeq \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$$

(using $\mathbb{R} \otimes_{\mathbb{R}} \mathbb{R} \simeq \mathbb{R}$).

(3). We have see that $R \simeq \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{H}$ as $\mathbb{R}$-algebras, and all these factors are simple $\mathbb{R}$-algebras, so they are the simple submodules of ${}_R R$. Now we have to calculate the action of R on them. We already now the action on $\mathbb{H}$. The four other factors are $\mathbb{R}[t_i]/(t_i \pm 1) \otimes_{\mathbb{R}} \mathbb{R}[t_j]/(t_j \pm 1)$ (with the four possibilities for the signs), with R acting by sending x_{-1} to 1, x_i to t_i and x_j to t_j. As $\mathbb{R}$-algebras, these are all isomorphic to $\mathbb{R}$. As R-modules, we get the R-modules corresponding to the following four $\mathbb{R}$-linear actions (=

representations) of Q on $\mathbb{R}$: the trivial one (sending every element of Q to 1),

$$\rho_{12} : -1 \longmapsto 1,\ i \longmapsto -1,\ j \longmapsto -1\ (\ k \longmapsto 1)$$

$$\rho_1 : -1 \longmapsto 1,\ i \longmapsto -1,\ j \longmapsto 1\ (k \longmapsto -1)$$

$$\rho_2 : -1 \longmapsto 1,\ i \longmapsto 1,\ j \longmapsto -1\ (k \longmapsto -1).$$

These are pairwise nonisomorphic, so we finally see that ${}_RR$ is the direct sum of the five simple R-modules described above, with multiplicities 1. In particular, R is semisimple.

(4). We have seen in problem VII.1.7(1) that every simple R-module is of the form R/I for I a maximal left ideal of R. As R is semisimple, this implies that every simple R-module is isomorphic to a simple submodule of R. We already gave a list of those in question (2).

(5). We know from problem VII.1.6 that $\mathbb{H} \otimes_R \mathbb{C} \simeq M_2(\mathbb{C})$. So question (3) implies that

$$R_{\mathbb{C}} = R \otimes_R \mathbb{C} \simeq \mathbb{C} \times \mathbb{C} \times \mathbb{C} \times \mathbb{C} \times M_2(\mathbb{C}).$$

All these factors are simple and semisimple $\mathbb{C}$-algebras. In particular, $R_{\mathbb{C}}$ is semisimple. As in (4), the simple $R_{\mathbb{C}}$-modules are all isomorphic to simple submodules of $R_{\mathbb{C}}$. The first four are the four 1-dimensional simple modules of R with scalars extended to $\mathbb{C}$ (they are simple because they're one-dimensional $\mathbb{C}$-vector spaces). The last simple R-module is $\mathbb{H}$, and $\mathbb{H} \otimes_R \mathbb{C} \simeq M_2(\mathbb{C})$ is not a simple $R_{\mathbb{C}}$-module, but is the direct sum of two simple submodules, both isomorphic to $\mathbb{C}^2$ with the usual action of $M_2(\mathbb{C})$ (and the action of $R_{\mathbb{C}}$ via the surjective map $R_{\mathbb{C}} \to M_2(\mathbb{C})$).

□

VII.1.9 A simple ring that is not semisimple

The goal of this problem is to construct simple rings that are not matrix rings over division rings (and hence not semisimple).

Let R be a ring. A *derivation* of R is an additive map $\delta : R \to R$ such that, for every $a, b \in R$, $\delta(ab) = a\delta(b) + \delta(a)b$.

(1). If $c \in R$, show that the map $\delta_c : R \to R$, $a \longmapsto ca - ac$ is a derivation. Such a derivation is called *inner*.

Let δ be a derivation of R. A *δ-ideal* of R is an ideal I of R such that $\delta(I) \subset I$. We say that R is *δ-simple* if $R \neq 0$ and its only δ-ideals are 0 and R.

The *differential polynomial ring* $R[x; \delta]$ is the R-module $R[x]$ with the multiplication given by $x^n x^m = x^{n+m}$ and $xa = ax + \delta(a)$, for $a \in R$.

(2). Show that $R[x; \delta]$ is indeed a ring.

(3). If $\delta = \delta_c$ with $c \in R$, show that the map $R[t] \to R[x;\delta]$, $t \longmapsto x - c$, is an isomorphism of rings.

(4). If δ is inner or R is not δ-simple, show that $R[x;\delta]$ is not simple.

(5). Conversely, we want to show that if δ is not inner and R is a δ-simple $\mathbb{Q}$-algebra, then $R[x;\delta]$ is simple. So assume that R is a δ-simple $\mathbb{Q}$-algebra, and that $R[x;\delta]$ contains a nonzero ideal $J \neq R[x;\delta]$.

(a) Let n be the minimum degree for the nonzero elements of J. (If $f \in R[x;\delta]$, write it as $\sum_{k\geq 0} a_k x^k$, and define the degree of f to be the biggest non-negative integer r such that $a_r \neq 0$.)

Show that $n > 0$ and that J contains an element g of the form $x^n + \sum_{k=0}^{n-1} a_k x^k$. (Hint : Use J to cook up a δ-ideal of R.)

(b) Show that δ is inner. (Hint : calculate $ga - ag$, for $a \in R$.)

(6). If $R \neq 0$, show that R is not left Artinian.

(7). Find a $\mathbb{Q}$-algebra R and a non-inner derivation δ on R such that R is δ-simple.

Solution.

(1). The map δ_c is obviously additive. Let $a, b \in R$. Then

$$\delta_c(ab) = c(ab) - (ab)c = (ca)b - (ac)b + (ac)b - (ab)c = \delta_c(a)b + a\delta_c(b).$$

(2). We have to check that, for every $a, b \in R$, $x1 = x$, $x(ab) = (xa)b$ and $x(a+b) = xa + xb$.

As $x1 = x + \delta(1)$, we want to show that $\delta(1) = 0$. But $\delta(1) = \delta(1 \cdot 1) = 1 \cdot \delta(1) + \delta(1) \cdot 1 = 2\delta(1)$, do indeed $\delta(1) = 0$.

We have

$$x(ab) = (ab)x + \delta(ab) = (ab)x + a\delta(b) + \delta(a)b$$
$$(xa)b = (ax + \delta(a))b = a(bx + \delta(b)) + \delta(a)b.$$

These are equal because $\delta(ab) = a\delta(b) + \delta(a)b$.

Finally,

$$x(a+b) = (a+b)x\delta(a+b) = ax + \delta(a) + bx + \delta(b) = xa + xb.$$

(3). To see that this map, that we'll call φ, is well-defined, we have to show that $x - c$ commutes with every element of $R[x;\delta]$. It suffices to show that it commutes with elements of R and with x. So let $a \in R$. We have :

$$(x-c)a = xa - ca = ax + \delta_c(a) - ca = ax + ca - ac - ca = ax - ac = a(x-c)$$

and

$$x(x-c) = x^2 - xc = x^2 - (cx + \delta_c(c)) = x^2 - cx = x(x-c).$$

Let's define a R-module map $\psi : R[x;\delta] \to R[t]$ by $\psi(x^n) = (t+c)^n$, for every $n \geq 0$. This is clearly the inverse of φ as a R-module map, so φ is an isomorphism (and ψ is a map of rings).

(4). If δ is inner, then $R[x;\delta] \simeq R[t]$, which is never simple (if $R = 0$, then $R[t] = 0$; if $R \neq 0$, then (t) is a nonzero proper ideal of $R[t]$).

If R is not δ-simple, let $I \neq 0, R$ be a δ-ideal of R, and let J be the R-submodule of $R[x;\delta]$ whose elements are the $\sum_{n\geq 0} a_n x^n$ with $a_n \in I$ for every n. Obviously, $J \neq 0, R[x;\delta]$. We want to show that J is an ideal of $R[x;\delta]$. For this, it suffices to show that $xJ \subset J$. Let $f = \sum_{n\geq 0} a_n x^n \in J$. Then

$$xf = \sum_{n\geq 0} (xa_n)x^n = \sum_{n\geq 0} (a_n x^{n+1} + \delta(a_n)x^n).$$

This is in J because $\delta(I) \subset I$.

(5). (a) First we show that, for every $g \in R[x;\delta]$, $\deg(xg - gx) \leq \deg(g)$. It is enough to show it for g of the form ax^n with $a \in R$, and then we have

$$xg - gx = (xa)x^n - ax^{n+1} = \delta(a)x^n.$$

Next we show that, for every $n \geq 0$ and $b \in R$, $\deg(x^n b - bx^n) \leq n-1$. We reason by induction on n. The result is obvious for $n = 0$, so let's assume that $n \geq 1$ and that know the result for $n-1$. Then

$$x^n b - bx^n = x^{n-1}(xb) - bx^n = x^{n-1}(bx + \delta(b)) - bx^n = (x^{n-1}b - bx^{n-1})x + x^{n-1}\delta(b)$$

is of degree $\leq \max(n-1, 1 + \deg(x^{n-1}b - bx^{n-1})$, and this is $\leq n-1$ by the induction hypothesis.

This implies in particular that, for every $g \in R[x;\delta]$ and $b \in R$, $\deg(gb - bg) < \deg(g)$.

Let $I \subset R$ be the union of $\{0\}$ and of the set of all leading coefficients of elements of J of degree n. It is clearly a left ideal of R. Let $a \in I$, choose $f, g \in R[x;\delta]$ such that $f = ax^n + g$, $\deg(g) \leq n-1$ and $f \in J$. Then for every $b \in R$, $fb \in J$, and we have

$$fb = ax^n b + gb = (ab)x^n + a(x^n b - bx^n) + gb$$

with $\deg(a(x^n b - bx^n) + gb) \leq n-1$, so $ab \in I$. So I is a right ideal of R. Moreover, $xf - fx \in J$, and we have

$$xf - fx = (xa)x^n + xg - ax^{n+1} - gx = \delta(a)x^n + xg - gx.$$

As $\deg(xg - gx) \leq n-1$, we have $\delta(a) \in I$. So I is a δ-ideal of R. As R is δ-simple and $I \neq 0$, we have $1 \in I$, so J contains an element g of the form $x^n + \sum_{k<n} a_k x^k$. As $J \neq r$, $n > 0$.

(b) Let's first show, by induction on n, that, for every $n \geq 1$ and every $a \in R$, $x^n a = ax^n + n\delta(a)x^{n-1} + h$, with $\deg(h) \leq n-2$. This is clear for $n = 1$, so assume that $n \geq 2$ and that the result is known for $n-1$. Then

$$x^n a = x(x^{n-1}a) = x(ax^{n-1} + (n-1)\delta(a)x^{n-2} + h),$$

with $\deg(h) \leq n-3$ (by the induction hypothesis). So

$$x^n a = ax^n + n\delta(a)x^{n-1} + (n-1)\delta^2(a)x^{n-2} + xh,$$

and we have $\deg((n-1)\delta^2(a)x^{n-2} + xh) \leq n-2$.

We have seen above that $\deg(ga - ag) < \deg(g)$ for every $a \in R$. By definition of n, this implies that $ga - ag = 0$ for every $a \in R$. Write $g = x^n + bx^{n-1} + h$, with $\deg(h) \leq n-2$. Then

$$\begin{aligned} ga - ag &= x^n a - ax^n + bx^{n-1}a - abx^{n-1} + ha - ah \\ &\quad n\delta(a)x^{n-1} + h_1 + (ba - ab)x^{n-1} + h_2 + ha - ah, \end{aligned}$$

with $\deg(h_1) \leq n-2$, $\deg(h_2) \leq n-2$ and $\deg(ha-ah) \leq n-2$. So $n\delta(a) = ba-ab$. As R is a $\mathbb{Q}$-algebra, this implies that $\delta = \delta_{n^{-1}b}$, so δ is inner.

(6). The sequence of ideals $(x) \supset (x^2) \supset (x^3) \supset \ldots$ does not stabilize.

(7). Let's take $R = \mathbb{Q}(t)$ and δ the derivation with respect to t. Then R is δ-simple because it is simple, because it is a field.

□

VII.1.10 Central simple algebras and the Brauer group

If R is a ring and $S \subset R$, the *centralizer of S in R* is

$$Z_R(S) = \{a \in R | \forall x \in S,\ ax = xa\}.$$

If $S = R$, we write $Z_S(R) = Z(R)$ and we call it the *center* of R.

In this problem, k will always be a field. We say that a k-algebra A is *central* if $Z(A) = k$, and we say that A is *finite* is $\dim_k(A) < \infty$.

(1). Let G be a finite group.

(a) When is $k[G]$ central (over k) ?

(b) When is $k[G]$ simple ?

(2). If R is a left Artinian simple ring and M is a finitely generated R-module, show that $\mathrm{End}_R(M)$ is a simple ring. If moreover R is a finite k-algebra, show that $\mathrm{End}_R(M)$ is also a finite k-algebra.

(3). If A and B are k-algebras, we make the tensor product $A \otimes_k B$ a ring by setting $(a \otimes b)(a' \otimes b') = (aa') \otimes (bb')$ and extending this by distributivity. Then $A \otimes_k B$ is also a k-algebra. (Remark : It is NOT true that every element of $A \otimes_k B$ is of the form $a \otimes b$. You have been warned.)

(a) Let A and A' be k-algbras, and let $B \subset A$ and $B' \subset A'$ be subalgebras. Show that

$$Z_{A \otimes_k A'}(B \otimes_k B') = Z_A(B) \otimes_k Z_{A'}(B').$$

(b) If A and B are k-algebras and n is a positive integer, construct an isomorphism

$$M_n(A \otimes_k B) \xrightarrow{\sim} M_n(A) \otimes_k B.$$

(c) Let A and B be k-algebras.

(i). If $\mathbb{D}$ is a central division k-algebra, show that we have a bijection between the set of ideals of $\mathbb{D} \otimes_k B$ and the set of ideals of B given by sending an ideal I of $\mathbb{D} \otimes_k B$ to $\{b \in B | 1 \otimes b \in I\}$ and an ideal J of B to $\mathbb{D} \otimes_k J$.

(ii). If A is a finite central simple k-algebra and B is simple, show that $A \otimes_k B$ is simple.

(iii). Given an example where A is simple but not central, B is simple and $A \otimes_k B$ is not simple.

(4). Let A be a finite central simple k-algebra.

(a) If K/k is an extension of fields, show that $A \otimes_k K$ is a finite central simple K-algebra.

(b) If k is algebraically closed, show that A is isomorphic to some $M_n(k)$.

(c) Show that $\dim_k(A)$ is the square of an integer. (Without using question (5).)

(5). Let A be a finite central simple k-algebra, let B be a simple k-subalgebra of A, and write $C = Z_A(B)$.

(a) Show that C is a simple k-algebra. (Hint : Let M be the unique simple A-module, identify C to the ring of endomorphisms of M that are linear for the action of some ring to be determined.)

(b) Show that $\dim_k(A) = \dim_k(B) \dim_k(C)$.

(c) If B is central, show that the multiplication map $B \otimes_k C \to A$ (that sends $b \otimes c$ to bc) is a k-algebra isomorphism.

(d) Let $k \subset K \subset A$ be a maximal commutative subfield of A, and suppose that A is a division algebra. Show that $Z_A(K) = K$ and $\dim_k(A) = \dim_k(K)^2$.

(6). (a) Let A be a finite central simple k-algebra. Show that the map

$$\begin{cases} A \otimes_k A^{op} & \to & \mathrm{End}_k(A) \\ a \otimes a' & \longmapsto & (x \longmapsto axa') \end{cases}$$

(exended by distributivity) is an isomorphism of k-algebras.

(b) Let A be a finite central division k-algebra. If $k \subset K \subset A$ is a maximal commutative subfield, show that $A \otimes_k K \simeq M_n(K)$, where $n = \dim_k(K)$. (Use the previous question, (5)(d) and (3)(a).)

(c) Let A and A' be finite central simple k-algebra. Show that the following are equivalent :

(i). There exists integers $m, m' \geq 1$ such that $M_m(A) \simeq M_{m'}(A')$ (as k-algebras).

(ii). There exists a k-division algebra $\mathbb{D}$ and integers $n, n' \geq 1$ such that $A \simeq M_n(\mathbb{D})$ and $A' \simeq M_{n'}(\mathbb{D})$.

If those conditions are satisfied, we say that A and A' are *similar* and write $A \sim A'$. This is obviously an equivalence relation on the set of isomorphism classes of finite central simple k-algebras, and we write $\mathrm{Br}(A)$ for the quotient of this set by $\sim$.

(d) If A, A', B, B' are finite central simple k-algebras such that $A \sim A'$ and $B \sim B'$, show that $A \otimes_k B$ and $A' \otimes_k B'$ are similar finite central simple k-algebras.

(e) Put the operation on $\mathrm{Br}(k)$ induced by the tensor product over k (this makes sense by the preceding question). Show that this makes $\mathrm{Br}(k)$ into a commutative group (the *Brauer group* of k).

(f) Calculate $\mathrm{Br}(k)$ for k algebraically closed.

(g) If I tell you that $\mathrm{Br}(\mathbb{R}) = \mathbb{Z}/2\mathbb{Z}$, can you give me a list of all finite $\mathbb{R}$-division algebras ?

(7). Reduced trace :

Let k be a field and A be a finite central simple k-algebra.

(a) Let $\varphi : A \to A$ be a an automorphism of k-algebras. Show that there exists $x \in A^\times$ such that $\varphi(a) = xax^{-1}$ for every $a \in A$.

Let $\sigma : k \to \Omega$ be a morphism of k into an algebraically closed field Ω. By (6)(b), there exists a positive integer n such that $A \otimes_k \Omega \simeq M_n(\Omega)$.

(b) Fix an isomorphism $A \otimes_k \Omega \simeq M_n(\Omega)$, and consider the morphism $T : A \to \Omega$ obtained by composing $A \to A \otimes_k \Omega$, $a \longmapsto a \otimes 1$, and $A \otimes_k \Omega \simeq M_n(\Omega) \stackrel{\mathrm{Tr}}{\to} \Omega$,

where Tr is the trace map. Show that this morphism T is independent of the choice of the isomorphism $A \otimes_k \Omega \simeq M_n(\Omega)$.

(c) Assume that $\mathrm{char}(k) = 0$. Show that we have $T = \sigma \circ \mathrm{Tr}_A$, where $\mathrm{Tr}_A : A \to k$ is a k-linear central morphism (ie such that $\mathrm{Tr}_A(ab) = \mathrm{Tr}_A(ba)$, for every $a, b \in A$). This morphism Tr_A is called the *reduced trace* of A. *(Hint : If $a \in A$, compare $T(a)$ and the trace of the k-linear endomorphism of A given by left multiplication by a.)*

Solution.

(1). (a) Only if $G = \{1\}$! We have seen in class that the element $\sum_{g \in G} g$ is central in $k[G]$, and this is in k if and only if G is trivial.

(b) Again, only if $G = \{1\}$. Indeed, the augmentation ideal of $k[G]$ is a proper two-sided ideal, and is only 0 when $G = \{1\}$.

(2). We have $R \simeq M_n(\mathbb{D})$, with $n \geq 1$ and $\mathbb{D}$ a division ring, and $V := \mathbb{D}^n$ is the only simple R-module up to isomorphism. As R is semisimple, M is a direct sum of copies of V, and as M is finitely generated, this direct sum if finite. So we may assume that $M = V^{\oplus m}$ for some $m \geq 0$. Then

$$\mathrm{End}_R(M) \simeq M_m(\mathrm{End}_R(V)) = M_m(\mathbb{D}^{op}).$$

If R is finite over k, so is $\mathbb{D}$, so M and $\mathrm{End}_R(M)$ are finite-dimensional k-vector spaces.

(3). (a) Remember that, if V and V' are k-vector spaces and the families $(e_i)_{i \in I}$ and $(e'_j)_{j \in I'}$ are bases of V and V', then $(e_i \otimes e'_j)_{(i,j) \in I \times I'}$ is a basis of $V \otimes_k V'$. In particular, if $x \in V \otimes_k V'$, there is unique family $(x_i)_{i \in I}$ of elements of V' such that $x = \sum_{i \in I} e_i \otimes x_i$.

Let's write $C = Z_A(B)$, $C' = Z_{A'}(B')$ and $C'' = Z_{A \otimes_k A'}(B \otimes_k B')$. We obviously have $C \otimes_k C' \subset C''$. So we have to show that $C'' \subset C \otimes_k C'$. For this, choose a basis $(e_i)_{i \in I}$ of A as a k-vector space. Let $x \in C''$, and write $x = \sum_{i \in I} e_i \otimes x_i$ with $x_i \in V$ (uniquely determined by x by the remark above). For every $b \in B'$, we have $x(1 \otimes b) = (1 \otimes b)x$, hence

$$\sum_{i \in I} e_i \otimes (bx_i - x_i b) = 0,$$

hence $bx_i - x_i b = 0$ for every $i \in I$, so all the x_i are in C' and $x \in A \otimes_k C'$. Similarly, choosing a basis $(e'_i)_{i \in I'}$ of the k-vector space C' and writing $x = \sum_{i \in I'} y_i \otimes e'_i$ with $y_i \in A$, we can show that $x \in C \otimes_k C'$.

(b) We define $\varphi : M_n(A) \otimes_k B \to M_n(A \otimes_k B)$ by sending $(a_{ij}) \otimes b$ to $(a_{ij} \otimes b)$ and extending this by linearity. As a map of abelian groups (or left A-modules), φ is simply the obvious isomorphism

$$A^{\oplus n^2} \otimes_k B \xrightarrow{\sim} (A \otimes_k B)^{\oplus n^2}.$$

But we should not forget to check that φ is a morphism of algebras. This is actually a straightforward check that follows directly from the definitions.

(c) (i). We prove that the two maps defined above are inverses of each other.

Let I be an ideal of $\mathbb{D} \otimes_k B$, and let $J = \{b \in B | 1 \otimes b \in J\}$. This is clearly and ideal of B, and we want to prove that $I = \mathbb{D} \otimes_k J$. The inclusion $I \supset \mathbb{D} \otimes_k J$ is obvious. Suppose that $I \neq \mathbb{D} \otimes_k J$. Choose a basis $(b_i)_{i \in A}$ of B as a k-vector space and a subset $A' \subset A$ such that $(b_i)_{i \in A - A'}$ is a basis of J. If $z \in \mathbb{D} \otimes_k B$, write $z = \sum_{i \in A} a_i \otimes b_i$ with the a_i in $\mathbb{D}$, and set $n(z) := |\{i \in A | a_i \neq 0\}|$. Choose $z \in I - \mathbb{D} \otimes_k J$ such that $n(z)$ is minimal (for z varying among elements of $I - \mathbb{D} \otimes_k J$). For every $i \in A - A'$, $a_i \otimes b_i \in \mathbb{D} \otimes_k J \subset I$, so $z - a_i \otimes b_i \in J - \mathbb{D} \otimes_k J$. By minimality of $n(z)$, $a_i = 0$. So $z = \sum_{i \in A'} a_i \otimes b_i$. Let $i_1, \ldots, i_r$ be the elements i of A' such that $a_i \neq 0$. Multiplying z by $a_{i_1}^{-1}$ on the left, we may assume that $a_{i_1} = 1$. If $a_{i_s} \in k$ for every s, then $z = 1 \otimes \left(\sum_{s=1}^r a_{i_s}^{-1} b_{i_s}\right)$, but then $\sum_{s=1}^r a_{i_s}^{-1} b_{i_s} \in J$ (by definition of J), contradiction. So there exists s such that $a_{i_s} \notin k$, and without loss of generality we may assume that $s = 2$. As the center of $\mathbb{D}$ is k, we can choose $a \in \mathbb{D}$ such that $a a_{i_2} - a_{i_2} a \neq 0$. Then

$$z' := (a \otimes 1) z - z (a \otimes 1) = \sum_{s=2}^{r} (a a_{i_s} - a_{i_s} a) \otimes b_{i_s} \in I$$

and $n(z') < n(z)$, so, by the minimality of $n(z)$, $z' \in \mathbb{D} \otimes_k J$. This is impossible because $a a_{i_2} - a_{i_2} a \neq 0$, so we have reached a contradiction.

For the other direction, let J be an ideal of B, let $I = \mathbb{D} \otimes_k J$, and let $J' = \{b \in B | 1 \otimes b \in I\}$. Obviously $J \subset J'$, so we have to show that $J' \subset J$. Let $V \subset J'$ be a k-subspace such that $J' = J \oplus V$. By the remark at the beginning of the solution of (a), we have $\mathbb{D} \otimes_k J' = (\mathbb{D} \otimes_k J) \oplus (\mathbb{D} \otimes_k V)$. So, if $b \in V$, $1 \otimes b$ cannot be in $I = \mathbb{D} \otimes_k J$ unless $b = 0$. So $V = 0$.

(ii). We have $A \simeq M_n(\mathbb{D})$, with $\mathbb{D}$ a division k-algebra and $n \geq 1$. If we embed $\mathbb{D}$ into $M_n(\mathbb{D})$ using $x \longmapsto x I_n$ (where I_n is the identity matrix), then this is a k-algebra map and it sends $Z(\mathbb{D})$ to $Z(M_n(\mathbb{D}))$. As A is central, so is $\mathbb{D}$. By (i), the set of ideals of $\mathbb{D} \otimes_k B$ is in bijection with the set of ideals of B; as B is simple, $\mathbb{D} \otimes_k B$ is also simple. We write $\mathbb{D} \otimes_k B \simeq M_m(\mathbb{D}')$, with $m \geq 1$ and $\mathbb{D}'$ a division k-algebra. By (b),

$$A \otimes_k B \simeq M_n(\mathbb{D} \otimes_k B) \simeq M_n(M_m(\mathbb{D}')) = M_{nm}(\mathbb{D}'),$$

so $A \otimes_k B$ is simple.

(iii). $k = \mathbb{R}$, $A = B = \mathbb{C}$. By PS1 5(d), $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C} \simeq \mathbb{C} \times \mathbb{C}$ is not a simple $\mathbb{R}$-algebra.

(4). (a) The K-algebra $A \otimes_k K$ is obviously finite. By (3)(a), we have

$$Z(A \otimes_k K) = Z(A) \otimes_k Z(K) = k \otimes_k K = K,$$

so $A \otimes_k K$ is a central K-algebra. By 2(c)(ii) (or 2(c)(i)), $A \otimes_k K$ is simple.

(b) We know that $A \simeq M_n(\mathbb{D})$, with $n \geq 1$ and $\mathbb{D}$ a finite division k-algebra. If k is algebraically closed, we have $\mathbb{D} = k$ by problem VII.1.3.

(c) Let K be an algebraically closed field containing k. By (a), $A \otimes_k K$ is a finite central simple K-algebra, so $\dim_K(A \otimes_k K)$ is a square by (b). But we have $\dim_k(A) = \dim_K(A \otimes_k K)$.

(5). (a) Write $A = M_n(\mathbb{D})$, with $\mathbb{D}$ a division k-algebra. Then $M := \mathbb{D}^n$, with the usual action of $M_n(\mathbb{D})$, is (up to isomorphism) the only simple A-module. We have seen in class that M also has a structure of right $\mathbb{D}$-module, and that $A = \mathrm{End}_{\mathbb{D}}(M)$. Now C is the subset of elements of $u \in \mathrm{End}_{\mathbb{D}}(M)$ such that, for every $b \in B$ and $x \in M$, $u(bx) = bu(x)$ (this follows from the definition of the isomorphism $A \xrightarrow{\sim} \mathrm{End}_{\mathbb{D}}(M)$).

We see the right $\mathbb{D}$-module structure on M as a left $\mathbb{D}^{op}$-module structure commuting with the action of A. This makes M into a left $A \otimes_k \mathbb{D}^{op}$-module, and C becomes the set of $B \otimes_k \mathbb{D}^{op}$ linear endomorphisms of M. We have seen in the proof of (3)(c)(ii) that $\mathbb{D}$ is a central k-algebra (because A is central). So $\mathbb{D}^{op}$ is also central. By (3)(c)(ii) again, this implies that $B \otimes_k \mathbb{D}^{op}$ is simple. As M is obviously finitely generated over $B \otimes_k \mathbb{D}^{op}$ (because $\dim_k M < \infty$), we can use question (2) to conclude that $C = \mathrm{End}_{B \otimes_k \mathbb{D}^{op}}(M)$ is a simple k-algebra.

(b) We use the notation of the proof of (a), and we write $a = \dim_k A$, $b = \dim_k B$ and $c = \dim_k C$. Let $B' = B \otimes_k \mathbb{D}^{op}$. We have seen that B' is simple, so $B' = M_m(\mathbb{D}')$, for some division k-algebra $\mathbb{D}'$. Let $M' = (\mathbb{D}')^m$ be its unique simple module. Then, as a B'-module, M is isomorphic to some $(M')^r$, and we have $C \simeq M_r(\mathrm{End}_{B'}(M')) \simeq M_r(\mathbb{D}'^{op})$. Note also that $r = \dim_k M / \dim_k M' = n \dim_k(\mathbb{D})/(m \dim_k(\mathbb{D}'))$ and $\dim_k B' = (\dim_k B)(\dim_k \mathbb{D}) = m^2 \dim_k \mathbb{D}'$. So

$$\begin{aligned} bc &= br^2 \dim_k(\mathbb{D}')^2 \\ &= b \dim_k(\mathbb{D}') \frac{n^2(\dim_k \mathbb{D})^2}{m^2(\dim_k \mathbb{D}')^2} \\ &= b \frac{n^2(\dim_k \mathbb{D})^2}{\dim_k B'} \\ &= n^2 \dim_k \mathbb{D} = a. \end{aligned}$$

(c) Let $u : B \otimes_k C \to A$ be the multiplication map. This is clearly a map of k-algebras. By (3)(c)(ii), $B \otimes_k C$ is simple, so u is injective (otherwise it would be 0, and this is not true). By (b), $\dim_k(B \otimes_k C) = \dim_k(A)$ and this is finite, so u is surjective.

(d) Let $A' = Z_K(A)$. It's a simple k-algebra by (a). As K is commutative, $K \subset A'$, and in fact $K \subset Z(A')$ by definition of A', so A' is simple K-algebra. Note that, if $x \in A' - \{0\}$, then $x^{-1} \in A'$: indeed, for every $y \in K$, we have $xy = yx$, hence $yx^{-1} = x^{-1}y$. So A' is a K-division algebra. Let $x \in A' - \{0\}$. Then $K[x]$ (the K-subalgebra of A' generated by x) is commutative, and it's also a field by the usual proof. (If $y \in A' - \{0\}$, then the elements $1, y, y^2, \ldots$ are linearly dependent

over K because $\dim_K A' < \infty$, so we have a relation $a_n y^n + \cdots + a_1 y + a_0 = 0$ with $a_n \neq 0$. As A' is a division algebra, after factoring out a power of y, we may assume that $a_0 \neq 0$, and then after divising by a_0 we may assume that $a_0 = 1$. Then $y^{-1} = -(a_n y^{n-1} + \cdots + a_2 y + a_1) \in K[x]$.) By the maximality of K, we have $K[x] = K$ for every $x \in A' - \{0\}$, which means that $Z_K(A) = A' = K$.

Now the fact that $\dim_k A = (\dim_k K)^2$ follows from (b).

(6). (a) Let's call this map φ. First we check that it's a morphism of rings (it's clear that it's k-linear) : If $a_1, a_2 \in A$ and $a_1', a_2' \in A^{op}$, then for every $x \in A$,

$$\varphi((a_1 \otimes a_1')(a_2 \otimes a_2'))(x) = \varphi((a_1 a_2) \otimes (a_2' a_1'))(x) = a_1 a_2 x a_2' a_2'$$

and

$$\varphi(a_1 \otimes a_1') \circ \varphi(a_2 \otimes a_2')(x) = \varphi(a_1 \otimes a_1')(a_2 x a_2') = a_1 a_2 x a_2' a_1',$$

so

$$\varphi((a_1 \otimes a_1')(a_2 \otimes a_2')) = \varphi(a_1 \otimes a_1') \circ \varphi(a_2 \otimes a_2').$$

By (3)(b)(ii), $A \otimes_k A^{op}$ is simple, so φ is injective. Moreover, $\dim_k(\mathrm{End}_k(A)) = (\dim_k(A))^2 = \dim_k(A \otimes_k A^{op})$, so φ is bijective.

(b) By the previous question, we have an isomorphism of k-algebras $\varphi : A \otimes_k A^{op} \xrightarrow{\sim} \mathrm{End}_k(A)$ sending $a \otimes a'$ to the map $x \longmapsto axa'$. As K is commutative, $K = K^{op}$, so we also see K as a subfield of A^{op}, and we have $Z_{A^{op}}(K) = K$ by (5)(d). By (3)(a), $A \otimes_k K = Z_A(k) \otimes_k Z_{A^{op}}(K) = Z_{A \otimes_k A^{op}}(k \otimes_k K)$. Using φ, this identifies $A \otimes_k K$ with the subalgebra A' of k-linear endomorphisms u of A such that $u(xb) = u(x)b$ for every $x \in A$ and $b \in K$. This is a K-algebra, where $b \in K$ acts by sending $u \in A'$ to the morphism $x \longmapsto u(xb) = u(x)b$. It follows from the definition of φ that it induces an isomorphism of K-algebras $A \otimes_k K \simeq A'$. Now seeing A as a K-vector space by maknig K act by left multiplication, we see that A' is isomorphic to $M_d(K)$, with $d = \dim_K A = \dim_k A / \dim_k K = \dim_k K$ (by (5)(d) again).

(c) As $M_n(M_m(R)) \simeq M_{nm}(R)$ for any ring R and any integers n, m, it's clear that (ii) implies (i). Let's show that (i) implies (ii). Let $m, m' \geq 1$ be integers such that $M_m(A) \simeq M_{m'}(A')$. Write $A = M_n(\mathbb{D})$ and $A' = M_{n'}(\mathbb{D}')$, with $\mathbb{D}, \mathbb{D}'$ division k-algebras and $n, n' \geq 1$. Then

$$M_{nm}(\mathbb{D}) \simeq M_m(A) \simeq M_{m'}(A) \simeq M_{n'm'}(\mathbb{D}').$$

We have seem in class that this implies that $\mathbb{D} \simeq \mathbb{D}'$ as k-algebras.

(d) Let $m, m' \geq 1$ be integers such that $M_m(B) \simeq M_{m'}(B')$. Then, by (3)(b),

$$M_m(A \otimes_k B)) \simeq A \otimes_k M_m(B) \simeq A \otimes_k M_{m'}(B') \simeq M_{m'}(A \otimes_k B'),$$

so

$$A \otimes_k B \sim A \otimes_k B'.$$

We see similarly that $A \otimes_k B' \sim A' \otimes_k B'$.

(e) As the tensor product is associative and commutative (up to isomorphism), the operation we put on $\mathrm{Br}(k)$ is associative and commutative. The class of the k-algebra k is clearly an identity element for this operation. By (a), for every finite central simple k-algebra,

$$A \otimes_k A^{op} \simeq \mathrm{End}_k(A) \simeq M_{\dim_k A}(k) \sim k,$$

which means that the class of A in $\mathrm{Br}_k(A)$ has an inverse, given by the class of A^{op}. So $\mathrm{Br}(k)$ is a commutative group.

(f) If k is algebraically closed, the only finite k-division algebra is k by problem VII.1.3, so every finite central simple k-algebra is similar to k, and $\mathrm{Br}(k) = \{1\}$.

(g) We already know two nonisomorphic finite central division $\mathbb{R}$-algebras, $\mathbb{R}$ and $\mathbb{H}$ (see problem VII.1.6 for $\mathbb{H}$). If $\mathrm{Br}(\mathbb{R}) = \mathbb{Z}/2\mathbb{Z}$, this means that they are the only finite central division $\mathbb{R}$-algebra. We still have to find the finite non-central division $\mathbb{R}$-algebras. Let $\mathbb{D}$ be a division $\mathbb{R}$-algebra such that $Z(\mathbb{D})$ contains $\mathbb{R}$ strictly. Then $Z(\mathbb{D})$ is a field, so it's a finite extension of $\mathbb{R}$ not equal to $\mathbb{R}$, so $Z(\mathbb{D}) = \mathbb{C}$ and $\mathbb{D}$ is a finite division $\mathbb{C}$-algebra. As $\mathbb{C}$ is algebraically closed, $\mathbb{D} = \mathbb{C}$.

(7). (a) The first $A \otimes_k \mathbb{D}$-module structure is given by taking $(a \otimes u) \cdot_1 m = au(m) = u(am)$, if $a \in A$, $u \in \mathbb{D}$ and $m \in M$. The second $A \otimes_k \mathbb{D}$-module structure is given by taking $(a \otimes u) \cdot_2 m = \varphi(a)u(m) = u(\varphi(a)m)$, if $a \in A$, $u \in \mathbb{D}$ and $m \in M$. By (3)(c)(ii), $A \otimes_k \mathbb{D}$ is a simple k-algebra, so it has a unique simple module up to isomorphism. As the two $A \otimes_k \mathbb{D}$-module structures on M make it a simple module, there exists an automorphism $\psi : M \to M$ such that $\psi((a \otimes u) \cdot_1 m) = (a \otimes u) \cdot_2 \psi(m)$, for every $a \in A$, $u \in \mathbb{D}$ and $m \in M$. As the factor $\mathbb{D}$ acts in the same way for both structures, ψ is in particular a $\mathbb{D}$-linear automorphism of M. But we know that $\mathrm{End}_{\mathbb{D}}(M) = A$, so there exists $x \in A^\times$ such that $\psi(m) = xm$ for every $m \in M$. We see that $xam = \varphi(a)xm$, for every $a \in A$ and every $m \in M$. As M is a faithful A-module, this implies that $\varphi(a)x = xa$ for every $a \in A$, hence $\varphi(a) = xax^{-1}$.

(b) By (a), if we have two isomorphisms $u_1 : A \otimes_k \Omega \xrightarrow{\sim} M_n(\Omega)$ and $u_2 : A \otimes_k \Omega \xrightarrow{\sim} M_n(\Omega)$, then there exists $g \in GL_n(\Omega)$ such that $u_2(x) = gu_1(x)g^{-1}$ for every $x \in A \otimes_k \Omega$. In particular, for every $a \in A$, $\mathrm{Tr}(u_2(a \otimes 1)) = \mathrm{Tr}(gu_1(a \otimes 1)g^{-1}) = \mathrm{Tr}(u_1(a \otimes 1))$.

(c) We identify k to its image by σ and stop writing σ. Then we just have to show that the image of Tr_A is contained in k.

Let $a \in A$, and let $m_a : A \to A$ be left multiplication by a. This is a k-linear endomorphism of A, so its $\mathrm{Tr}(m_a)$ is an element of k. This trace is also equal to the trace of the Ω-linear endomorphism of $A \otimes_k \Omega \simeq M_n(\Omega)$ given by left multiplication

by $a \otimes 1$. We know that, as a left $M_n(\Omega)$-module, $M_n(\Omega)$ is isomorphic to $(\Omega^n)^{\oplus n}$ (with the usual action of $M_n(\Omega)$ on Ω^n), so we finally get $\mathrm{Tr}(m_a) = nT(a)$. As $\mathrm{char}(k) = 0$, this implies that $T(a) \in k$.

□

VII.1.11 Irreducible representations of p-groups in characteristic p

Let G be a finite group and p a prime number. We say that G is a *p-group* if $|G|$ is a power of p. If G is a nontrivial p-group, then its center $Z(G)$ is nontrivial. [4]

(1). Suppose that the only irreducible representation of G on a finite-dimensional $\overline{\mathbb{F}}_p$-vector space is the trivial representation (i.e. $\overline{\mathbb{F}}_p$ with every $g \in G$ acting as identity). Show that G is a p-group. (Use the left regular $\overline{\mathbb{F}}[G]$-module $\overline{\mathbb{F}}[G]$.) [5]

(2). Conversely, if k is an algebraically closed field of characteristic p, G is a p-group and V is an irreducible representation of G on a finite-dimensional k-vector space, show that V is the trivial representation. *(Hint : Look at the subspace of vectors that are invariant by every element of $Z(G)$.)*

Solution.

(1). Let $V = \overline{\mathbb{F}}_p[G]$, seen as a left $\overline{\mathbb{F}}_p[G]$-module. This gives a group morphism $\rho : G \to GL(V)$, which is obviously injective. Choose a Jordan-Hölder series $V = V_0 \supset V_1 \supset \cdots \supset V_n = 0$. Then every V_i/V_{i+1} is a simple $\overline{\mathbb{F}}_p[G]$-module, hence equal to the trivial representation of G, and so $n = \dim_{\overline{\mathbb{F}}_p}(\overline{\mathbb{F}}_p[G]) = |G|$. Choose a basis $(e_1, \ldots, e_n)$ of V as a $\overline{\mathbb{F}}_p$-vector space such that, for every $i \in \{1, \ldots, n\}$, $(e_1, \ldots, e_i)$ is a basis of V_{n-i}. This gives an isomorphism $GL(V) \simeq GL_n(\overline{\mathbb{F}}_p)$, and the composition of this with ρ sends G injectively to the group $U(\overline{\mathbb{F}}_p)$ of upper triangular matrices in $GL_n(\overline{\mathbb{F}}_p)$ with ones on the diagonal. For every integer $r \geq 1$, let $U(\mathbb{F}_{p^r}) = U(\overline{\mathbb{F}}_p) \cap GL_n(\mathbb{F}_{p^r})$. Then $U(\overline{\mathbb{F}}_p) = \bigcup_{r \geq 1} U(\mathbb{F}_{p^r})$. As G is finite, its image by ρ is contained in $U(\mathbb{F}_{p^r})$ for r big enough. So there exists r such that G is isomorphic to a subgroup of $U(\mathbb{F}_{p^r})$. But $|U(\mathbb{F}_{p^r})| = (p^r)^{n(n-1)/2}$ is a power of p, so the order of G is also a power of p.

(2). We show the result by induction on the order of G. Suppose that G is abelian (this is the case, for example, if $|G| = p$). Let V be an irreducible representation of V. By Schur's lemma (and because k is algebraically closed), $\mathrm{End}_{k[G]}(V) = k$. But, as G is abelian, $k[G]$ is commutative, so the action of any element of $k[G]$ on V is in $\mathrm{End}_{k[G]}(V)$, so there exists a k-algebra map $u : k[G] \to k$ such that, for every $x \in k[G]$ and $v \in V$, $xv = u(x)v$. In particular, every k-subspace of V is a subrepresentation; as V is irreductible, $\dim V = 1$.

[4] See your favorite algebra textbook. Mine is Perrin's book [23].

[5] We could replace $\overline{\mathbb{F}}_p$ by a an arbitrary field of characteristic p in this question, and the result would stay true.

Giving u is the same as giving a morphism of groups $\rho : G \to k^\times$. But the only element of $k^\times$ that has order a power of p is 1 (because $t^{p^r} - 1 = (t-1)^{p^r}$ in $k[t]$), so ρ is trivial, so G acts trivially on V.

Now assume that G is not abelian, and that we know the result for every p-group of order $< |G|$. Let Z be the center of G, it's a nontrivial abelian p-group, so the induction hypothesis applies to it. Let V be an irreducible representation of G. We denote by V^Z the k-subspace of $v \in V$ such that $gv = v$ for every $g \in Z$. If $g \in G$ and $v \in V^Z$, then for every $h \in Z$,

$$h(gv) = (hg)v = (gh)v = g(hv) = gv,$$

so $gv \in V^Z$. This show that V^Z is actually a subrepresentation of V.

Let $V = V_0 \supset V_1 \supset \cdots \supset V_n = 0$ be a Jordan-Hölder series for V seen as a $k[Z]$-module. Then V_{n-1} is a simple $k[Z]$-module, so it's the trivial representation of Z by the induction hypothesis. This means that $V_{n-1} \subset V^Z$, and so $V^Z \neq 0$. As V is irreducible, $V^Z = V$. So the action of G on V factors G/Z, and we can see V as an irreducible representation of G/Z. By the induction hypothesis, this is the trivial representation of G/Z, and so V is the trivial representation of G.

□

VII.1.12 Another description of induction

Let R be a ring, G be a finite group and H be a subgroup of G. Choose a system of representatives $g_1, \ldots, g_r$ of G/H. Let M be a $R[H]$-module, write

$$I = \{f : G \to M | \forall h \in H, \forall g \in G,\ f(hg) = hf(g)\}.$$

We make G act on I by $(g \cdot f)(x) = f(xg)$ if $f \in I$ and $x, g \in G$.

Show that the map $I \to \mathrm{Ind}_H^G M$, $f \longmapsto \sum_{i=1}^r g_i \otimes f(g_i^{-1})$, is an isomorphism of $R[G]$-modules, and that it is independent of the choice of $g_1, \ldots, g_s$.

Solution. Let's call this map u. We first check that u is $R[G]$-linear. As u is obviously R-linear, we just need to show that it is compatible with the actions of G on its source and target. So let $f \in I$ and $g \in G$. Then we have

$$g \cdot u(f) = g \sum_{i=1}^r g_i \otimes f(g_i^{-1})$$

and

$$u(g \cdot f) = \sum_{i=1}^r g_i \otimes (g \cdot f)(g_i^{-1}) = \sum_{i \in r} g_i \otimes f(g_i^{-1} g).$$

Note that $g_1^{-1}, \dots, g_r^{-1}$ is a system of representatives of the quotient $H\backslash G$, and so is $g_1^{-1}g, \dots, g_r^{-1}g$. So we have a (unique) permutation $\sigma \in \mathfrak{S}_r$ and (unique) elements $h_1, \dots, h_r \in H$ such that $g_i^{-1}g = h_i g_{\sigma(i)}^{-1}$ for every i. This gives

$$u(g \cdot f) = \sum_{i=1}^{r} g_i \otimes f(h_i g_{\sigma(i)}^{-1}) = \sum_{i=1}^{r} (g_i h_i) \otimes f(g_{\sigma(i)}^{-1})$$

(using the properties of f and the fact that the tensor product is over $R[H]$). As $g_i h_i = g g_{\sigma(i)}$ for every i, we finally get

$$u(g \cdot f) = g \sum_{i=1}^{r} g_{\sigma(i)} f(g_{\sigma(i)}^{-1}) = g \cdot u(f).$$

Now we check that u is an isomorphism, by defining an inverse $v : \mathrm{Ind}_H^G M \to I$. Remember from class that, as a right $R[H]$-module, $R[G]$ is free with basis $(g_1, \dots, g_r)$. So, as an abelian group,

$$\mathrm{Ind}_H^G M = R[G] \otimes_{R[H]} M \simeq \bigoplus_{i=1}^{r} g_i R[H] \otimes_{R[H]} M \simeq \bigoplus_{i=1}^{r} M.$$

Using this isomorphism, we'll define v as a morphism $M^r \to I$. If $(m_1, \dots, m_r) \in M^r$, we send it to the sum $f_1 + \cdots + f_r \in I$, where, for every i, $f_i : G \to M$ is the function defined by

$$f_i(x) = \begin{cases} hm_i & \text{if } x = hg_i^{-1} \text{ with } h \in H \\ 0 & \text{otherwise} \end{cases}.$$

As the definition of f_i is R-linear in m_i, we see easily that v is indeed R-linear. We have to show that it is the inverse of u.

Let $f \in I$. Then $u(f) = \sum_{i=1}^{r} g_i \otimes f(g_i^{-1})$, which corresponds to the element $(f(g_1^{-1}), \dots, f(g_r^{-1}))$ of M^r. So $vu(f)$ is the sum $f_1 + \cdots + f_r$, where, for every i,

$$f_i(x) = \begin{cases} hf(g_i^{-1}) = f(hg_i^{-1}) & \text{if } x = hg_i \text{ with } h \in H \\ 0 & \text{otherwise} \end{cases}.$$

As $G = \coprod_{i=1^r} H g_i^{-1}$, we have indeed $f = f_1 + \cdots + f_r$.

Now let $x \in \mathrm{Ind}_H^G M$, write $x = \sum_{i=1}^{r} g_i \otimes m_i$ with $(m_1, \dots, m_r) \in M^r$, and write $f = v(x)$. Then

$$u(f) = \sum_{i=1}^{r} g_i \otimes f(g_i^{-1}) = \sum_{i=1}^{r} g_i \otimes m_i = x.$$

To finish, we have to show that the morphism u is independent of the choice of $g_1, \dots, g_r$. So let $g_1', \dots, g_r'$ be another system of representatives of G/H, and $u' : I \to \mathrm{Ind}_H^G M$ be the map that we get by using the g_i'. Up to changing the order of the g_i' (which obviously does not affect

u'), we may assume that there are $h_1, \dots, h_r \in H$ such that $g'_i = g_i h_i$ for every i. Then, for every $f \in I$,

$$u'(f) = \sum_{i=1}^{r} g'_i \otimes f(g_i'^{-1}) = \sum_{i=1}^{r} g_i h_i \otimes f(h_i^{-1} g_i^{-1}) = \sum_{i=1}^{r} g_i h_i \otimes h_i^{-1} f(g_i^{-1}) = u(f)$$

(as the tensor product is over $R[H]$).

□

VII.1.13 Representation ring of $\mathbb{Z}/p^r\mathbb{Z}$ in characteristic p

Take $G = \mathbb{Z}/p^r\mathbb{Z}$ and k of characteristic p. Show that $\mathrm{P}_k(G)$ is the free abelian group generated by $[k[G]]$ and calculate the map $\mathrm{P}_k(G) \to \mathrm{R}_k(G)$. [6] (Remember that $\mathrm{P}_k(G)$ and $\mathrm{R}_k(G)$ were introduced in definition I.4.6 of chapter I.)

Solution. We first show that a finitely generated $k[G]$-module is projective if and only if it is free. This will obviously imply that $\mathrm{P}_k(G)$ is the free abelian group generated by $[k[G]]$.

We already know that a free $k[G]$-module is projective. Conversely, let M be a projective $k[G]$-module of finite type. Note that $k[G] \simeq R := k[T]/(T^{p^r} - 1)$, so we can see M as a finitely generated $k[T]$-module. Using the structure theorem for finitely generated modules over PIDs, we see that, as a $k[T]$-module, M is a direct sum of $k[T]^s$ and of modules of the type $k[T]/(f^m)$, where the f are irreducible polynomials. As M is actually a $k[T]/(T^{p^r} - 1)$-module, we must have $s = 0$, and all the f^m that appear divide $T^{p^r} - 1$. As $char(k) = p$, $T^{p^r} - 1 = (T-1)^{p^r}$ in $k[T]$, so the only f that can appear in the decomposition above is $T - 1$, and we see that M is a direct sum of $k[T]/(T^{p^r} - 1)$-modules isomorphic to $k[T]/((T-1)^m)$, with $1 \leq m \leq p^r$. If $m = p^r$, $k[T]/((T-1)^m) = R \simeq k[G]$, so we just have to show that $k[T]/((T-1)^m)$ is not a projective R-module if $1 \leq m < p^r$. Fix m such that $1 \leq m < p^r$ and let $M = k[T]/((T-1)^m)$, then we have an obvious surjective R-module map $v : R \to M$ (sending T to T), and its kernel $M' := (T-1)^{p^r - m} R$ is isomorphic to $k[T]/(T^{p^r - m} - 1)$. If M were projective, we would have $R = M \oplus M'$ as R-modules. But the element $(T-1)^{\max(p^r - m, m)}$ acts as 0 on M and M' and not on R, so this is not possible and M cannot be projective.

Now we have to calculate the map $\mathrm{P}_k(G) \to \mathrm{R}_k(G)$. As $\mathrm{P}_k(G)$ is the free group on $[k[G]]$, $\mathrm{P}_k(G) \simeq \mathbb{Z}$ and we just have to calculate the image of this generator in $\mathrm{R}_k(G)$. We have seen in problem 7 of PS3 that the only simple $k[G]$-module is $\mathbb{1}$, so $\mathrm{R}_k(G) \simeq \mathbb{Z}$. Also, all the Jordan-Hölder factors of $k[G]$ are isomorphic to $\mathbb{1}$, and there are $\dim_k k[G] = |G| = p^r$ of them. Finally, the map $\mathrm{P}_k(G) \to \mathrm{R}_k(G)$ is isomorphic to the map $\mathbb{Z} \to Z$, $a \longmapsto p^r a$.

□

[6] This will be generalized to all p-groups in proposition III.8.1 of chapter III.

VII.1.14 Basic properties of induction

Let R be a ring and G a finite group.

(1). If $K \subset H \subset G$ are subgroups and V is a $R[K]$-module, show that

$$\mathrm{Ind}_H^G(\mathrm{Ind}_K^H V) \simeq \mathrm{Ind}_K^G V.$$

(2). If $H' \subset H \subset G$ are subgroups such that H' is normal in G, and if W is a $R[H]$-module on which H' acts trivially (so that we can also view W as a $R[H/H']$-module), show that

$$\mathrm{Ind}_H^G W \simeq \mathrm{Ind}_{H/H'}^{G/H'} W.$$

(3). Let G_1 and G_2 be two finite groups, and let $G = G_1 \times G_2$. If R is a commutative ring, $H_1 \subset G_1$ and $H_2 \subset G_2$ are subgroups, and V_1 (resp. V_2) is a $R[H_1]$-module (resp. a $R[H_2]$-module), show that

$$\mathrm{Ind}_{H_1 \times H_2}^{G_1 \times G_2}(V_1 \otimes_R V_2) \simeq (\mathrm{Ind}_{H_1}^{G_1} V_1) \otimes_R (\mathrm{Ind}_{H_2}^{G_2} V_2).$$

Solution.

(1). We have

$$\mathrm{Ind}_H^G(\mathrm{Ind}_K^H V) = R[G] \otimes_{R[H]} (R[H] \otimes_{R[K]} V) \simeq R[G] \otimes_{R[K]} V \simeq \mathrm{Ind}_K^G V.$$

(If you're unfamiliar with that property of tensor products : the isomorphism in the middle is given by $a \otimes (b \otimes v) \longmapsto (ab) \otimes v$, its inverse by $c \otimes v \longmapsto c \otimes (1 \otimes v)$. Note that, in the left hand side, $a \otimes (b \otimes v) = (ab) \otimes (1 \otimes v)$.)

(2). We have a $R[G]$-linear map

$$u : \mathrm{Ind}_H^G W = R[G] \otimes_{R[H]} W \to R[G/H'] \otimes_{R[H/H']} W = \mathrm{Ind}_{H/H'}^{G/H'} W$$

given by the obvious map $R[G] \to R[G/H']$. Let $g_1, \ldots, g_r$ be a system of representatives of G/H. Then $g_1' := g_1 H', \ldots, g_r' := g_r H'$ is a system of representatives of $(G/H')/(H/H')$. So we have isomorphisms of R-modules $W^r \to \mathrm{Ind}_H^G W$, $(w_1, \ldots, w_r) \longmapsto \sum_{i=1}^r g_i \otimes w_i$, and $W^r \to \mathrm{Ind}_{H/H'}^{G/H'} W$, $(w_1, \ldots, w_r) \longmapsto \sum_{i=1}^r g_i' \otimes w_i$. By these isomorphisms, u corresponds to the identity of W^r. So u is an isomorphism.

(3). First we note that the map $R[G_1] \otimes_R R[G_2] \to R[G_1 \times G_2]$, $a \otimes b \longmapsto ab$, is a ring isomorphism. (Indeed, it is obviously a map of rings and it sends the basis $(g_1 \otimes g_2)_{(g_1,g_2) \in G_1 \times G_2}$ of $R[G_1] \otimes_R R[G_2]$ to the basis $((g_1, g_2))_{(g_1,g_2) \in G_1 \times G_2}$ of $R[G_1 \times G_2]$. We have a similar isomorphism $R[H_1] \otimes_R R[H_2] \simeq R[H_1 \times H_2]$.

So we get :

$$\mathrm{Ind}_{H_1\times H_2}^{G_1\times G_2}(V_1\otimes_R V_2)\simeq (R[G_1]\otimes_R R[G_2])\otimes_{(R[H_1]\otimes_R R[H_2])}(V_1\otimes_R V_2).$$

We get a map of this into $(\mathrm{Ind}_{H_1}^{G_1} V_1)\otimes_R(\mathrm{Ind}_{H_2}^{G_2} V_2)$ by sending $(a_1\otimes a_2)\otimes(v_1\otimes v_2)$ to $(a_1\otimes v_1)\otimes(a_2\otimes v_2)$. (This is well-defined because it's additive in every variable, and if $b_i\in R[H_i]$ for $i=1,2$, then $((a_1\otimes a_2)(b_1\otimes b_2))\otimes(v_1\otimes v_2)$ and $(a_1\otimes a_2)\otimes((b_1\otimes b_2)(v_1\otimes v_2))$ are sent to the same element, ie $((a_1b_1)\otimes v_1)\otimes((a_2b_2)\otimes v_2)=(a_1\otimes(b_1v_1))\otimes(a_2\otimes(b_2v_2))$.) This is an isomorphism, because it has an inverse, given by sending $(a_1\otimes v_1)\otimes(a_2\otimes v_2)$ to $(a_1\otimes a_2)\otimes(v_1\otimes v_2)$. (Again, this is well-defined, and the verification is similar.)

□

VII.2 Chapter II exercises

VII.2.1 Representation rings and field extensions

In this problem, whenever k is a field and G is a group, we assume that all $k[G]$-modules are finite-dimensional over k.

Let G be a group, and let K/k be an extension of fields.

(1). If V and W are $k[G]$-modules, show that the obvious map

$$\mathrm{Hom}_G(V,W)\otimes_k K\to\mathrm{Hom}_G(V\otimes_k K, W\otimes_k K)$$

(sending $u\otimes x$ to xu if $x\in K$ and $u\in\mathrm{Hom}_k(V,W)$) is an isomorphism.

(2). If G is finite and $\mathrm{char}(k)$ does not divide $|G|$, show that the map $\mathrm{R}_k(G)\to\mathrm{R}_K(G)$ is injective. [7]

(3). If G is finite and k is algebraically closed of characteristic prime to $|G|$, show that the map $\mathrm{R}_k(G)\to\mathrm{R}_K(G)$ is bijective. [8]

Solution.

(1). We know that $\mathrm{Hom}_G(V,W)=\mathrm{Hom}_k(V,W)^G$ and $\mathrm{Hom}_G(V\otimes_k K, W\otimes_k K)=\mathrm{Hom}_K(V\otimes_k K, W\otimes_k K)^G$. (See remark II.1.1.9 of chapter II.) Notice that the map $\mathrm{Hom}_k(V,W)\otimes_k K\to\mathrm{Hom}_K(V\otimes_k K, W\otimes_k K)$, $u\longmapsto u\otimes 1$, is an isomorphism. (Choosing a basis $(e_i)_{i\in I}$ of V over k identifies

[7] The result is still true in any characteristic for a separable algebraic extension, for example by theorem 5.17 of [20]. But what happens in general ?

[8] The result is still true without the assumption on the characteristic of k, as we can see by using the characteristic zero case and corollary III.7.1 of chapter III.

$\mathrm{Hom}_k(V,W)$ with $\prod_{i\in I} W$. But $(e_i \otimes 1)_{i\in I}$ is a basis of $V \otimes_k K$ over K and gives an identification $\mathrm{Hom}_K(V\otimes_k, W\otimes_k K) \simeq \prod_{i\in I}(W\otimes_k K)$, and then the map above becomes the product over I of the identity maps $W \otimes_k K = W \otimes_k K$.) So we just have to prove the following fact :

If V is a $k[G]$-module, then the map $V^G \otimes_k K \to (V \otimes_k K)^G$, $x \longmapsto x \otimes 1$, is an isomorphism. This map is obviously injective, so we have to show that it is surjective. Let $(\alpha_j)_{j\in J}$ be a basis of K as a k-vector space. Let $v \in (V \otimes_k K)^G$, and write $v = \sum_{j\in J} v_j \otimes \alpha_j$ with $v_j \in V$ for every j. Then, for every $g \in G$, $gv = \sum_{j\in J}(gv_j) \otimes \alpha_j = v \sum_{j\in J} v_j \otimes \alpha_j$, which implies that $gv_j = v_j$ for every $j \in J$. So all the v_j are in V^G, and $v \in V^G \otimes_k K$. (This is very similar to the proof of (3)(a) in problem VII.1.10.)

(2). Note that the hypothesis ensures that all the modules over $k[G]$ and $K[G]$ are semisimple.

We know that $\mathrm{R}_k(G)$ is the free abelian group on the $[W]$, $W \in S_k(G)$. So proving that the map $u : \mathrm{R}_k(G) \to \mathrm{R}_K(G)$ is injective is equivalent to proving that the family $(u([W]))_{W\in S_k(G)}$ is linearly independent in $\mathrm{R}_K(G)$. For every $W \in S_k(G)$, let $S(W)$ be the set of $V \in S_K(G)$ such that V is isomorphic to a $K[G]$-submodule of $W \otimes_k K$.

Now we have to prove that, if W and W' are non-isomorphic simple $k[G]$-modules, then $S(W) \cap S(W') = \varnothing$, that is, $W \otimes_k K$ and $W' \otimes_k K$ have no simple factors in common. As $W \otimes_k K$ and $W' \otimes_k K$ are semisimple, this is the same as saying that $\mathrm{Hom}_G(W \otimes_k K, W' \otimes_k K) = 0$. But we know from (a) that $\mathrm{Hom}_G(W \otimes_k K, W' \otimes_k K) = \mathrm{Hom}_G(W,W') \otimes_k K = 0$.

(3). By theorem II.1.3.1 of chapter II, we know that the free abelian groups $\mathrm{R}_k(G)$ and $\mathrm{R}_K(G)$ have the same rank. (Indeed, the rank $\mathrm{R}_k(G)$ is equal to $\dim_k \mathscr{C}(G,k)$, and that of $\mathrm{R}_K(G)$ is equal to $\dim_K \mathscr{C}(G,K)$. As $\mathscr{C}(G,k) \otimes_k K = \mathscr{C}(G,K)$, these two dimensions are equal.) As the $[W]$ for $W \in S_k(G)$ (resp. $W \in S_K(G)$) form a basis of $\mathrm{R}_k(G)$ (resp. $\mathrm{R}_K(G)$), this means that $|S_k(G)| = |S_K(G)|$. So if we can show that $W\otimes_k K$ is irreducible for every $W \in S_k(G)$, this will imply that the map $\mathrm{R}_k(G) \to \mathrm{R}_K(G)$ sends a basis of $\mathrm{R}_k(G)$ to a basis of $\mathrm{R}_K(G)$, hence is an isomorphism.

So let W be a simple $k[G]$-module. By (1) and Schur's lemma, $\mathrm{End}_G(W \otimes_k K) = \mathrm{End}_G(W) \otimes_k K = k \otimes_k K = K$. If $W \otimes_k K$ were not irreducible, we could write $W \otimes_k K = V_1 \oplus V_2$ with $V_1, V_2 \neq 0$, and then the $K[G]$-linear endomorphisms $\begin{pmatrix} \mathrm{id}_{V_1} & 0 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 0 \\ 0 & \mathrm{id}_{V_2} \end{pmatrix}$ would generate a dimension 2 K-subspace of $\mathrm{End}_G(W\otimes_k K)$, which would contradict the calculation above. So $W\otimes_k K$ is irreducible.

□

VII.2.2 Some character tables

Let G be a finite group, and let C be a set of representatives of the set of conjugacy classes in G. Then giving a central function in $G \to k$ is the same as giving a function $C \to k$. The character table of G is a table showing the values of the characters of the irreducible representations of G (over k) at every element of C. For example, the character table of $\mathfrak{S}_2 \simeq \{\pm 1\}$ (over $\mathbb{Q}$) is

	$\mathbb{1}$	sgn
1	1	1
-1	1	-1

Find representatives for the sets of conjugacy classes and write the character tables in the following situations :

(1). $G = \mathfrak{S}_3$, $k = \mathbb{Q}$;

(2). $G = \{\pm 1, \pm i, \pm j, \pm k\}$, $k = \mathbb{R}$;

(3). $G = \{\pm 1, \pm i, \pm j, \pm k\}$, $k = \mathbb{C}$.

Solution.

(1). $G = \mathfrak{S}_3$, $k = \mathbb{Q}$: We use the decomposition in cycles to determine the conjugacy classes in $\mathfrak{S}_n$: Every element of $\mathfrak{S}_n$ is a unique way a product $c_1 \dots c_r$ where the c_i are cycles with pairwise disjoint supports (note that then the c_i commute), and two elements $c_1 \dots c_r$ and $c'_1 \dots c'_s$ written in this way are conjugate if and only (i) $r = s$ and (ii) up to reordering the c'_i, the cycles c_i and c'_i have the same length for every $i \in \{1, \dots, r\}$.

In particular, we have three conjugacy classes in G : the conjugacy class of 1, the conjugacy class of transpositions (a representative is (12)) and the conjugacy class of 3-cycles (a representative is (123)).

We have seen in class that the irreducible representations of $\mathfrak{S}_3$ over $\mathbb{Q}$ are $\mathbb{1}$, sgn and the space $V = \{(x_1, x_2, x_3) \in \mathbb{Q}^3 | x_1 + x_2 + x_3 = 0\}$ (with $\mathfrak{S}_3$ acting by permuting the coordinates).

So the character table is

	$\mathbb{1}$	sgn	V
1	1	1	2
(12)	1	-1	0
(123)	1	1	-1

(2). $G = \{\pm 1, \pm i, \pm j, \pm k\}$, $k = \mathbb{R}$: We see easily that the conjugacy classes in G are the following : $\{1\}$, $\{-1\}$, $\{\pm i\}$, $\{\pm j\}$, $\{\pm k\}$. We have seen that the irreducible representations of G over $\mathbb{R}$ are $\mathbb{1}$, ε_1, ε_2, ε_3 and $\mathbb{H}$ (where $\mathbb{H}$ has the obvious action of G by left multiplication, and the others are the 1-dimensional representations defined in class (see

also the character table for their values)). So the character table is :

	$\mathbb{1}$	ε_1	ε_2	ε_3	$\mathbb{H}$
1	1	1	1	1	4
-1	1	1	1	1	-4
i	1	1	-1	-1	0
j	1	-1	1	-1	0
k	1	-1	1	-1	0

(3). $G = \{\pm 1, \pm i, \pm j, \pm k\}$, $k = \mathbb{C}$: The only change is that the irreducible representation $\mathbb{H}$ of G over $\mathbb{R}$ splits over $\mathbb{C}$ as $V \oplus V$, with V a simple $\mathbb{C}[G]$-module (ie $\mathbb{H} \otimes_{\mathbb{R}} \mathbb{C} \simeq V \oplus V$ as $\mathbb{C}[G]$-module). So we get the following character table :

	$\mathbb{1}$	ε_1	ε_2	ε_3	V
1	1	1	1	1	2
-1	1	1	1	1	-2
i	1	1	-1	-1	0
j	1	-1	1	-1	0
k	1	-1	1	-1	0

□

VII.2.3 Calculating representation rings

(1). Let n be a positive integer, and let μ_n be the group of nth roots of 1 in k. Assume that k is algebraically closed. Show that, as a $\mathbb{Z}$-algebra, $\mathrm{R}_k(\mathbb{Z}/n\mathbb{Z})$ is isomorphic to $\mathbb{Z}[\mu_n]$.

(2). Let $G = \{\pm 1, \pm i, \pm j, \pm k\}$. The theorem quoted at the beginning says that $\mathrm{R}_{\mathbb{C}}(G)$ is a subring of $\mathscr{C}(G, \mathbb{C}) \simeq \mathbb{C}^5$. But do we have $\mathrm{R}_{\mathbb{C}}(G) \simeq \mathbb{Z}^5$ as a ring ? *(Hint : Look for idempotents.)*

Solution.

(1). Write $n = mp^r$, where $p = char(k)$ and $p \not| m$ (if $char(k) = 9$, we take $n = m$). Then μ_n is a cyclic group of order m. Let ζ_n be a generator of μ_n. The irreducible representations of $\mathbb{Z}/n\mathbb{Z}$ over k are the $\varepsilon_0, \ldots, \varepsilon_{m-1}$, where $\varepsilon_i : \mathbb{Z}/n\mathbb{Z} \to k^\times$ is the morphism of groups that sends 1 to ζ_n^i. For every $i \in \{0, \ldots, m-1\}$, let $c_i = [\varepsilon_i] \in \mathrm{R}_k(\mathbb{Z}/n\mathbb{Z})$. Then $(c_0, \ldots, c_{n-1})$ is a basis of $\mathrm{R}_k(\mathbb{Z}/n\mathbb{Z})$ as a $\mathbb{Z}$-module. Also, for every i, j, $c_i c_j = [\varepsilon_i \otimes \varepsilon_j]$, and $\varepsilon_i \otimes \varepsilon_j : \mathbb{Z}/n\mathbb{Z} \to k^\times$ sends 1 to ζ_n^{i+j}. So the $\mathbb{Z}$-linear isomorphism $\mathrm{R}_k(\mathbb{Z}/n\mathbb{Z}) \to \mathbb{Z}[\mu_n]$ that sends c_i to ζ_n^i pour tout $i \in \{0, \ldots, n-1\}$ is a map of rings.

(2). With the notation of problem 4, we write $c_r = [\varepsilon_r] \in \mathrm{R}_{\mathbb{C}}(G)$ for $r \in \{1, 2, 3\}$, and $d = [V]$. Then $(1, c_1, c_2, c_3, d)$ is a basis of the $\mathbb{Z}$-module $\mathrm{R}_{\mathbb{C}}(G)$. Using the character table of problem 4 (remember that the character of a tensor product is the product of the characters),

we see that $c_1^2 = c_2^2 = c_3^2 = 1$, $c_1c_2 = c_3$, $c_2c_3 = c_1$, $c_3c_1 = c_2$, $c_1d = c_2d = c_3d = d$ and $d^2 = 1 + c_1 + c_2 + c_3$.

Now let x be an element of $\mathrm{R}_{\mathbb{C}}(G)$, write $x = \alpha + \beta_1 c_1 + \beta_2 c_2 + \beta_3 c_3 + \gamma d$, with $\alpha, \beta_1, \beta_2, \beta_3, \gamma \in \mathbb{Z}$, and suppose that $x^2 = x$. By the multplication table we established above, the coefficient of 1 in x^2 is $\alpha^2 + \beta_1^2 + \beta_2^2 + \beta_3^2 + \gamma^2$. As $x = x^2$, this is equal to α, which is only possible if $\alpha = 1$ and $\beta_1 = \beta_2 = \beta_3 = \gamma = 0$. So 1 is the only idempotent of $\mathrm{R}_{\mathbb{C}}(G)$, and we cannot have $\mathrm{R}_{\mathbb{C}}(G) \simeq \mathbb{Z}^5$ as a ring.

□

VII.2.4 Representations of products

Let G_1 and G_2 be two finite groups, and let $G = G_1 \times G_2$. Let k be an algebraically closed field of characteristic 0.

(1). Let V_1 (resp. V_2) be an irreducible representation of G_1 (resp. G_2) (over k). Show that $V_1 \otimes_k V_2$ is an irreducible representation of G.

(2). If V_1, V_1' (resp. V_2, V_2') are irreducible representations of G_1 (resp. G_2), show that the following are equivalent :

  (a) the representations $V_1 \otimes_k V_2$ and $V_1' \otimes_k V_2'$ of G are isomorphic;

  (b) V_1 and V_1' are isomorphic and V_2 and V_2' are isomorphic.

(3). Show that every irreducible representation of G is of the form $V_1 \otimes_k V_2$, with V_1 (resp. V_2) an irreducible representation of G_1 (resp. G_2). (There are several ways to do this.)

(4). Is (1) still true if k is not algebraically closed ? *(Hint : problem VII.1.6(4).)*

Solution.

(1). We have seen in class that the character of $V := V_1 \otimes_k V_2$ is the map $(g_1, g_2) \longmapsto \chi_{V_1}(g_1)\chi_{V_2}(g_2)$. So we have :

$$\langle V, V\rangle_G = \frac{1}{|G_1 \times G_2|} \sum_{(g_1,g_2)\in G_1\times G_2} \chi_{V_1}(g_1)\chi_{V_2}(g_2) = \langle V_1, V_1\rangle_{G_1}\langle V_2, V_2\rangle_{G_2}.$$

As these three brackets are non-negative integers, we deduce that $\langle V, V\rangle_G = 1$ (i.e. V is irreducible) if and only if $\langle V_1, V_1\rangle_{G_1} = \langle V_2, V_2\rangle_{G_2} = 1$ (i.e. both V_1 and V_2 are irreducible).

(2). It's obvious that (b) implies (a), so let's show that (a) implies (b). Assuming (a), let's show for example that V_1 and V_1' are isomorphic (the case of V_2 and V_2' is similar). Let $d = \dim_k V_2$ and $d' = \dim_k V_2'$. Then, for every $g \in G_1$,

$$d\chi_{V_1}(g) = \chi_{V_1}(g)\chi_{V_2}(1) = \chi_{V_1\otimes_k V_2}(g, 1) = \chi_{V_1'\otimes_k V_2'}(g, 1) = \chi_{V_1'}(g)\chi_{V_2'}(1) = d'\chi_{V_1'}(g).$$

So χ_{V_1} and $\chi_{V_1'}$ are proportional. As V_1 and V_1' are irreducible, this is only possible if $V_1 \simeq V_1'$.

(3). By the two previous questions, we have a map $S_k(G_1) \times S_k(G_2) \to S_k(G_1 \times G_2)$, and it's injective. We want to show that this map is bijective, and for this we can for example show that its source and target have the same cardinality.

Let $X(G_1)$ (resp. $X(G_2)$, resp. $X(G_1 \times G_2)$) be the set of conjugacy classes in G_1 (resp. G_2, resp. $G_1 \times G_2$). We have an obvious surjective map $u : G_1 \times G_2 \to X(G_1 \times G_2)$, and two elements $(g_1, g_2), (g_1', g_2') \in G_1 \times G_2$ are conjugate in $G_1 \times G_2$ if and only g_i and g_i' are conjugate in G_i for $i = 1, 2$. So u induces a bijection $X(G_1) \times X(G_2) \xrightarrow{\sim} X(G_1 \times G_2)$. As $|S_k(G)| = X(G)$ for $G \in \{G_1, G_2, G_1 \times G_2\}$, this proves that $|S_k(G_1) \times S_k(G_2)| = |S_k(G_1 \times G_2)|$.

Here is another way to prove the conclusion of (3). We identify G_1 and G_2 with the subgroups $G_1 \times \{1\}$ and $\{1\} \times G_2$ of G. Let V an irreducible representation of G. Let $V = W_1 \oplus \cdots \oplus W_m$ be the decomposition of V in isotypic components as a representation of G_1. Then every element of G_2 stabilizes all the W_i. (A useful trick : if $g \in G_2$, then g centralizes G_1, so the endomorphism of V given the action of g is a G_1-equivariant isomorphism, and it is clear from the definition of isotypic components that such an isomorphism must respect them.) So the W_i are $k[G]$-representations of V. As V is irreducible, we have $V = W_1$, that is, there is a simple $k[G_1]$-module V_1 and $r_1 \geq 1$ such that $V \simeq V_1^{\oplus r_1}$ as representations of G_1. Similarly, as a representation of G_2, V is isomorphic to a $V_2^{\oplus r_2}$, with V_2 a simple $k[G_2]$-module. Let $i = 1, 2$. In the decomposition of $k[G_i]$ as a product of simple k-algebras, let $M_{n_i}(k)$ be the factor corresponding to V_i; then the action of $k[G_i]$ on V factors through its quotient $M_{n_i}(k)$. So the action of $k[G]$ on V factors through $A := M_{n_1}(k) \otimes M_{n_2}(k) \simeq M_{n_1 n_2}(k)$. But this k-algebra is simple, so it has a unique simple module up to isomorphism. As $V_1 \otimes_k V_2$ and V are simple A-modules, we deduce that they must isomorphic as A-modules, and hence as $k[G]$-modules.

(4). No. Take $k = \mathbb{R}$, $G_1 = G_2 = \{\pm 1, \pm i\}$, and $V_1 = V_2 = \mathbb{C}$ with the obvious action. We have seen in problem VII.1.6(4) that there is a $\mathbb{R}$-linear isomorphism $V_1 \otimes_{\mathbb{R}} V_2 \xrightarrow{\sim} \mathbb{C} \oplus \mathbb{C}$, $(x, y) \longmapsto xy \oplus x\overline{y}$. If we use this to transport the action of G to $\mathbb{C} \oplus \mathbb{C}$, we see that each summand is stable by G, so the representation $V_1 \otimes_k V_2$ is not irreducible.

□

VII.2.5 Character of small symmetric and exterior powers

Let G be a group, k be a field of characteristic not diving 6 and V be a representation of G on a finite-dimensional k-vector space.

For every $n \geq 0$, $S^n V$ and $\wedge^n V$ be the representations of G on the nth symmetric and exterior

powers of V.[9] Show that, for every $g \in G$,

$$\chi_{S^2V}(g) = \frac{\chi(g)^2 + \chi(g^2)}{2}$$

$$\chi_{\wedge^2 V}(g) = \frac{\chi(g)^2 - \chi(g^2))}{2}$$

$$\chi_{S^3V}(g) = \frac{\chi(g)^3 + 3\chi(g^2)\chi(g) + 2\chi(g^3)}{6}$$

and

$$\chi_{\wedge^3 V}(g) = \frac{\chi(g)^3 - 3\chi(g^2)\chi(g) + 2\chi(g^3)}{6}.$$

Solution. We may assume that k is algebraically closed. Denote the representation by $\rho : G \to \mathrm{GL}(V)$. Let $g \in G$, and choose a basis $(e_1, \dots, e_n)$ of V in which the matrix of $\rho(g)$ is upper triangular with diagonal entries $\lambda_1, \dots, \lambda_n$. Then the image in S^2V (resp. $\wedge^2 V$) of the family $(e_i \otimes e_j + e_j \otimes e_i)_{1 \le i \le j \le n}$ (resp. $(e_i \otimes e_j - e_j \otimes e_i)_{1 \le i < j \le n}$) of T^2V is a basis. In this basis and using the lexicographic order on $\{1, \dots, n\}^2$, $S^2\rho(g)$ (resp. $\wedge^2\rho(g)$) is upper triangular with diagonal entries $(\lambda_i\lambda_j)_{1 \le i \le j \le n}$ (resp. $(\lambda_i\lambda_j)_{1 \le i < j \le n}$). So

$$\chi_{S^2V}(g) = \sum_{1 \le i \le j \le n} \lambda_i\lambda_j = \frac{1}{2}\left(\left(\sum_{i=1}^n \lambda_i\right)^2 + \sum_{i=1}^n \lambda_i^2\right) = \frac{1}{2}(\chi_V(g)^2 + \chi_V(g^2)),$$

and

$$\chi_{\wedge^2 V}(g) = \sum_{1 \le i < j \le n} \lambda_i\lambda_j = \frac{1}{2}\left(\left(\sum_{i=1}^n \lambda_i\right)^2 - \sum_{i=1}^n \lambda_i^2\right) = \frac{1}{2}(\chi_V(g)^2 - \chi_V(g^2)).$$

Similarly, the image in S^3V (resp. $\wedge^3 V$) of the family $(\sum_{\sigma \in S_3} e_{i_{\sigma(1)}} \otimes e_{i_{\sigma(2)}} \otimes e_{i_{\sigma(3)}})_{1 \le i_1 \le i_2 \le i_3 \le n}$ (resp. $(\sum_{\sigma \in S_3} \mathrm{sgn}(\sigma) e_{i_{\sigma(1)}} \otimes e_{i_{\sigma(2)}} \otimes e_{i_{\sigma(3)}})_{1 \le i_1 < i_2 < i_3 \le n}$) of T^3V is a basis. In this basis and using the lexicographic order on $\{1, \dots, n\}^3$, $S^3\rho(g)$ (resp. $\wedge^3\rho(g)$) is upper triangular with diagonal entries $(\lambda_{i_1}\lambda_{i_2}\lambda_{i_3})_{1 \le i_1 \le i_2 \le i_3 \le n}$ (resp. $(\lambda_{i_1}\lambda_{i_2}\lambda_{i_3})_{1 \le i_1 < i_2 < i_3 \le n}$).

Note that

$$\sum_{1 \le i_1 \le i_2 \le i_3 \le n} \lambda_{i_1}\lambda_{i_2}\lambda_{i_3} = \frac{1}{6}\left(\left(\sum_{i=1}^n \lambda_i\right)^3 + 3\sum_{1 \le i,j \le n} \lambda_i^2\lambda_j + 2\sum_{i=1}^n \lambda_i^3\right)$$

and

$$\sum_{1 \le i_1 < i_2 < i_3 \le n} \lambda_{i_1}\lambda_{i_2}\lambda_{i_3} = \frac{1}{6}\left(\left(\sum_{i=1}^n \lambda_i\right)^3 - 3\sum_{1 \le i,j \le n} \lambda_i^2\lambda_j + 2\sum_{i=1}^n \lambda_i^3\right),$$

[9] Symmetric powers are defined in problem VII.6.7, and exterior powers in section VI.9.1 of chapter VI.

so

$$\chi_{S^3V}(g) = \frac{1}{6}(\chi(g)^3 + 3\chi(g^2)\chi(g) + 2\chi(g^3))$$

and

$$\chi_{\wedge^3V}(g) = \frac{1}{6}(\chi(g)^3 - 3\chi(g^2)\chi(g) + 2\chi(g^3)).$$

□

VII.2.6 Using Mackey's irreducibility criterion (representations of $\mathrm{GL}_2(\mathbb{F}_q)$, part 1)

Let $\mathbb{F}$ be a finite field, let $G = GL_2(\mathbb{F})$, and consider the subgroup B of upper triangular matrices in G. Let k be an algebraically closed field of characteristic 0 and $\omega_1, \omega_2 : \mathbb{F}^\times \to k^\times$ two morphisms of groups. We consider the representation $\rho : B \to k^\times$ given by

$$\rho \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = \omega_1(a)\omega_2(d).$$

(1). When is $\mathrm{Ind}_B^G \rho$ irreducible ? (Give a condition on ω_1 and ω_2.)

(2). Calculate the character of $\mathrm{Ind}_B^G \rho$.

Solution.

(1). First we need a system of representatives for G/B. I claim that the matrix $g_c := \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}$ for $c \in \mathbb{F}$ and $w := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ form such a system. Indeed, an easy calculation show that

$$g_c B = \left\{ \begin{pmatrix} x & y \\ z & t \end{pmatrix} \in GL_2(\mathbb{F}) | x \neq 0 \text{ and } z = cx \right\}$$

for $c \in \mathbb{F}$, and

$$wB = \left\{ \begin{pmatrix} 0 & y \\ z & t \end{pmatrix} \in GL_2(\mathbb{F}) \right\}.$$

Obviously, $GL_2(\mathbb{F})$ is the disjoint union of these subsets.

To see whether $\mathrm{Ind}_B^G \rho$ is irreducible, we use Mackey's irreducibility criterion. First note that ρ is irreducible.

Let $c \in \mathbb{F} - \{0\}$ (if $c = 0$, $g_c \in B$). Then, for every $x = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in B$, we have

$$g_c x g_c^{-1} = \begin{pmatrix} a - bc & b \\ c(a - (bc + d)) & bc + d \end{pmatrix}.$$

So

$$B \cap g_c B g_c^{-1} = \left\{ \begin{pmatrix} a & b \\ 0 & a-bc \end{pmatrix}, a \in \mathbb{F}^\times,\ b \in \mathbb{F} \right\},$$

and for an element x of $B \cap g_c B g_c^{-1}$ written as above, we have $\rho(x) = \omega_1(a)\omega_2(a-bc)$ and $\rho(g_c^{-1} x g_c) = \omega_1(a-bc)\omega_2(a)$. These two representations of $B \cap g_c B \cap g_c^{-1}$ are 1-dimensional, so they have a simple factor in common if and only if they are isomorphic, and this is equivalent to $\omega_1 = \omega_2$.

Also, wBw^{-1} is the set of lower triangular matrices, so $B \cap wBw^{-1}$ is the set of diagonal matrices, and, if $x = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, then $\rho(x) = \omega_1(a)\omega_2(b)$ and $\rho(w^{-1}xw) = \omega_2(a)\omega_1(b)$. Again, these two representations of $B \cap wBw^{-1}$ have a simple factor in common if and only if they are isomorphic, and this is equivalent to $\omega_1 = \omega_2$.

Finally, Mackey's irreducibility criterion shows that $\mathrm{Ind}_B^G \rho$ is irreducible if and only $\omega_1 \neq \omega_2$.

(2). Let χ be the character of $\mathrm{Ind}_B^G \rho$, and let $X = \{g_c, c \in \mathbb{F}\} \cup \{w\}$. According to the formula for the character of an induced representation, for every $x \in G$,

$$\chi(x) = \sum_{g \in X, g^{-1}xg \in B} \rho(g^{-1}xg).$$

Also, we know that $\chi(x)$ depends only on the conjugacy class of x. We have three types of conjugacy classes in $GL_2(\mathbb{F})$:

(i) If x is diagonalizable over $\mathbb{F}$, then it's conjugate to a diagonal matrix $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, with $a, b \in \mathbb{F}^\times$ (duh).

(ii) If x is not diagonalizable over $\mathbb{F}$ but is diagonalizable over an algebraic closure of $\overline{\mathbb{F}}$, then the eigenvalues of x are in the unique degree 2 extension $\mathbb{F}'$ of $\mathbb{F}$. This $\mathbb{F}'$ is generated by the square root of an element $u \in \mathbb{F}^\times - (\mathbb{F}^\times)^2$, and x is conjugated to a matrix of the form $\begin{pmatrix} a & b \\ ub & a \end{pmatrix}$, with $a \in \mathbb{F}$ and $b \in \mathbb{F}^\times$ (if $b = 0$, we are in case (i)).

(iii) If x is not diagonalizable over any extension of $\mathbb{F}$, then it is conjugated to a matrix of the form $\begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}$, with $a \in \mathbb{F}^\times$.

In case (i), we have $w^{-1}xw \in B$ and $\rho(w^{-1}xw) = \omega_2(a)\omega_1(b)$, and $g_c^{-1}xg_c = \begin{pmatrix} a & 0 \\ c(a-b) & b \end{pmatrix} \in B$ if and only $c = 0$ or $a = b$, with $\rho(g_c^{-1}xg_c) = \omega_1(a)\omega_2(b)$ in both cases. So

$$\chi(x) = \begin{cases} \omega_1(a)\omega_2(b) + \omega_2(a)\omega_1(b) & \text{if } a \neq b \\ (1 + |\mathbb{F}|)\omega_1(a)\omega_2(a) & \text{if } a = b \end{cases}.$$

In case (ii), $w^{-1}xw = \begin{pmatrix} a & ub \\ b & a \end{pmatrix} \notin B$, and $g_c^{-1}xg_c = \begin{pmatrix} a+bc & b \\ b(u-c^2) & a-bc \end{pmatrix} \notin B$, so

$$\chi(x) = 0.$$

In case (iii), $w^{-1}xw \notin B$, and $g_c^{-1}xg_c = \begin{pmatrix} a+c & 1 \\ -c^2 & a-c \end{pmatrix}$ is in B if and only $c = 0$, and in that case $\rho(g_c^{-1}xg_c) = \omega_1(a)\omega_2(a)$. So

$$\chi(x) = \omega_1(a)\omega_2(a).$$

□

VII.2.7 Rationality problems

In this problem, k is a field of characteristic 0 and G is a finite group. Let $k \subset \Omega$ be an extension of k. Remember that we have a commutative square of injective maps

$$\begin{array}{ccc} \mathrm{R}_k(G) & \longrightarrow & \mathrm{R}_\Omega(G) \\ \downarrow & & \downarrow \\ \mathscr{C}(G,k) & \longrightarrow & \mathscr{C}(G,\Omega) \end{array}$$

(By corollary II.1.2.9 of chapter 2 and problem VII.2.1.)

We will use this to identify $\mathrm{R}_k(G)$ (resp. $\mathrm{R}_\Omega(G)$) with its image in $\mathscr{C}(G,k)$ (resp. $\mathscr{C}(G,\Omega)$).

We say that a representation V of G over Ω is *realizable over* k if there exists a $k[G]$-module W and a $\Omega[G]$-linear isomorphism $W \otimes_k \Omega \simeq V$.

(1). If V is a $\Omega[G]$-module, show that V is realizable over k if and only if $\chi_V \in \mathrm{R}_k(G)$, and that in that case, any two $k[G]$-modules W, W' such that $W \otimes_k \Omega \simeq V$ and $W' \otimes_k \Omega \simeq V$ are isomorphic (over $k[G]$).

Now suppose that Ω is algebraically closed, and write $\mathrm{R}(G) = \mathrm{R}_\Omega(G)$. (We have seen in class that $\mathrm{R}_\Omega(G)$ is independent of the choice of the algebraically closed extension Ω of k, hence the notation.) Let $\mathrm{R}'_k(G)$ be the space of elements $\chi \in R(G)$ that take their values in k (when seen as central functions on G). We obviously have $\mathrm{R}_k(G) \subset \mathrm{R}'_k(G)$, and we want to investigate the difference between the two.

Let $k[G] = A_1 \times \cdots \times A_r$ be the decomposition of $k[G]$ into simple k-algebras. For every i, write $A_i = M_{n_i}(\mathbb{D}_i)$ with $\mathbb{D}_i$ a division algebra and $n_i \geq 1$, let $V_i = \mathbb{D}_i^{n_i}$ be the unique simple A_i-module, $K_i = Z(\mathbb{D}_i)$ (a finite extension of k) and $m_i = \sqrt{\dim_{K_i}(\mathbb{D}_i)}$ (this is an integer by problem VII.1.10(4), and it's called the *Schur index* of the simple $k[G]$-module V_i). We also set

$\chi_i = \chi_{V_i} \in \mathscr{C}(G,k)$ and denote by $\psi_i \in \mathscr{C}(G,K_i)$ the function that takes $g \in G$ to the reduced trace of its image in A_i (see problem VII.1.10(7)). Finally, let Σ_i be the set of k-linear field morphisms $K_i \to \Omega$.

(2). Show that every $|\Sigma_i| = [K_i : k]$.

(3). Show that $(\chi_1, \dots, \chi_r)$ is a basis of $\mathrm{R}_k(G)$.

(4). If $i \in \{1, \dots, r\}$ and $\sigma \in \Sigma_i$, we write $A_{i,\sigma} = A_i \otimes_{K_i} \Omega$, where we use the morphism $\sigma : K_i \to \Omega$ to form the tensor product. (Note that A_i is naturally a K_i-algebra.) Show that this is a simple Ω-algebra of dimension $n_i^2 m_i^2$, where $m_i = \sqrt{\dim_{K_i} \mathbb{D}_i}$ (an integer by (4) of problem VII.1.10), and that we have

$$\Omega[G] \simeq \prod_{i=1}^{r} \prod_{\sigma \in \Sigma_i} A_{i,\sigma}.$$

(5). For every $i \in \{1, \dots, r\}$ and $\sigma \in \Sigma_i$, let $W_{i,\sigma}$ be the unique simple $A_{i,\sigma}$-module and let $\psi_{i,\sigma}$ be its character.

Show that $\psi_{i,\sigma} = \sigma \circ \psi_i$, that every irreducible representation of G over Ω is isomorphic to a unique $W_{i,\sigma}$, and that $\chi_i = m_i \sum_{\sigma \in \Sigma_i} \psi_{i,\sigma}$ for every $i \in \{1, \dots, r\}$.

(6). Let $\chi \in R(G)$. By (d), we can write $\chi = \sum_{i=1}^{r} \sum_{\sigma \in \Sigma_i} a_{i,\sigma} \psi_{i,\sigma}$, with $a_{i,\sigma} \in \mathbb{Z}$. Show that $\chi \in \mathrm{R}'_k(G)$ if and only if, for every $i \in \{1, \dots, r\}$ and every $\sigma, \tau \in \Sigma_i$, $a_{i,\sigma} = a_{i,\tau}$.

(7). Show that $(m_1^{-1}\chi_1, \dots, m_r^{-1}\chi_r)$ is a basis of $\mathrm{R}'_k(G)$. (In particular, the quotient $\mathrm{R}'_K(G)/\mathrm{R}_k(G)$ is finite of order $m_1 \dots m_r$, and $\mathrm{R}_k(G) = \mathrm{R}'_k(G)$ if and only if all the $\mathbb{D}_i$ are commutative.)

Solution.

(1). If $\chi_V \in \mathrm{R}_k(G)$, then we have $\chi_V = \sum_{W \in S_k(G)} n_W \chi_W$, with the n_W in $\mathbb{Z}$. If W and W' are in $S_k(G)$, then $\mathrm{Hom}_{k[G]}(W, W') \otimes_k \Omega = \mathrm{Hom}_{\Omega[G]}(W \otimes_k \Omega, W' \otimes_k \Omega)$ (by (1) of problem VII.2.1), which is zero unless $W \simeq W'$. So if $W \not\simeq W'$, the $\Omega[G]$-modules $W \otimes_k \Omega$ and $W' \otimes_k \Omega$ have no simple factor in common. For every $W \in S_k(G)$, let S_W be the set of $M \in S_\Omega(G)$ that are simple factors of $W \otimes_k \Omega$ and write $[\chi_W] = \sum_{M \in S_W} n_{W,M} \chi_M$ in $R_\Omega(G)$, with the $n_{W,M}$ non-negative integers (and at least one of them nonzero).

As the S_W are pairwise disjoint in $S_\Omega(G)$, there are no cancellations in the expression

$$\chi_V = \sum_{W \in S_k(G)} n_W \sum_{M \in S_W} n_{W,M} \chi_M,$$

so all the n_W are nonnegative integers. Hence we can form the $k[G]$-module $V' = \bigoplus_{W \in S_k(G)} W^{\oplus n_W}$. We have $\chi_V = \chi_{V'} = \chi_{V' \otimes_k \Omega}$, so $V \simeq V' \otimes_k \Omega$, so V is defined over k.

Let W, W' be two $k[G]$-modules such that $W \otimes_k \Omega \simeq W' \otimes_k \Omega \simeq V$. Then $\chi_W = \chi_{W\otimes_k\Omega} = \chi_{W'\otimes_k\Omega} = \chi_{W'}$, so $W \simeq W'$.

(2). By the primitive element theorem, there exists $x \in K_i$ such that $K_i = k[x]$. Let $f \in k[T]$ be the minimal polynomial of x over k. Then $K_i \simeq k[T]/(f)$, so f is irreducible, $\deg(f) = [K_i : k]$, and $\sigma \longmapsto \sigma(x)$ is a bijection between Σ_i and the set of roots of f in Ω.

(3). We have seen in class that the classes of the simple $k[G]$-modules form a basis of the $\mathbb{Z}$-module $\mathrm{R}_k(G)$. So there is nothing to do.

(4). The Ω-algebra $A_{i,\sigma}$ is simple by problem VII.1.10(3)(c)(ii). (We can apply this because K_i is the center of A_i, as $K_i = Z(\mathbb{D}_i)$ and $A_i = M_{n_i}(\mathbb{D}_i)$, and

$$\dim_\Omega(A_{i,\sigma}) = \dim_{K_i}(A_i) = n_i^2 \dim_{K_i}(\mathbb{D}_i).$$

Fix i. By the primitive element theorem, there is a $x \in K_i$ such that $K_i = k[x]$. Let $f \in k[T]$ be the minimal polynomial of x, then we have an isomorphism $k[T]/(f) \xrightarrow{\sim} K_i$ given by $T \longmapsto x$. The set of roots of f in Ω is $\{\sigma(x), \sigma \in \Sigma_i\}$, and we get an isomorphism

$$K_i \otimes_k \Omega \simeq k[T]/(f) \otimes_f \Omega \simeq \Omega[T]/(f) \simeq \Omega^{\Sigma_i}$$

by sending $a \otimes b \in K_i \otimes \Omega$ to $(\sigma(a)b)_{\sigma\in\Sigma_i}$. This gives an isomorphism

$$A_i \otimes_k \Omega \simeq A_i \otimes_{K_i} (K_i \otimes_k \Omega) \simeq \prod_{\sigma\in\Sigma_i} A_{i,\sigma},$$

and we finally get

$$\Omega[G] \simeq (\prod_{i=1}^{r} A_i) \otimes_k \Omega \simeq \prod_{i=1}\prod_{\sigma\in\Sigma_i} A_{i,\sigma}.$$

(5). By the previous question, the irreducible representations of G over Ω are exactly the simple $A_{i,\sigma}$-modules, for $i \in \{1, \ldots, r\}$ and $\sigma \in \Sigma_i$. There is exactly one for each i and σ as above, an it's $W_{i,\sigma}$.

For every i and $\sigma \in \Sigma_i$, let $V_{i,\sigma} = V_i \otimes_{K_i} \Omega$, where we use $\sigma : K_i \to \Omega$ to form the tensor product. Then as above, we get an isomorphism

$$V_i \otimes_k \Omega \simeq \prod_{\sigma\in\Sigma_i} V_{i,\sigma},$$

and the factor $A_{i,\sigma}$ of $A_i \otimes_k \Omega$ acts on $V_i \otimes_k \Omega$ through $V_{i,\sigma}$. The unique simple $A_{i,\sigma}$-module if of dimension $n_i m_i$ over Ω, and $\dim_\Omega(V_{i,\sigma}) = \dim_{K_i} V_i = n_i m_i^2$ (because we know that $V_i \simeq \mathbb{D}_i^{\oplus n_i}$), so $V_{i,\sigma} \simeq W_{i,\sigma}^{\oplus m_i}$. Finally, we get

$$\chi_{V_i} = \sum_{\sigma\in\Sigma_i} \chi_{V_{i,\sigma}} = m_i \sum_{\sigma\in\Sigma_i} \psi_{i,\sigma}.$$

It remains to calculate the $\psi_{i,\sigma}$. Fix $i \in \{1, \dots, r\}$ and $\sigma \in \Sigma_i$. Let $a \in A_i$. Then by definition of the reduced trace, $\sigma(\psi_i(a))$ is the trace of $a \otimes 1 \in A_{i,\sigma}$, where the trace is defined by identifying $A_{i,\sigma}$ with the matrix algebra $M_{n_i m_i}(\Omega)$. In other words, it's the trace of $a \otimes 1$ on the unique simple $A_{i,\sigma}$-module $W_{i,\sigma}$, that is, $\psi_{i,\sigma}(a)$.

(6). For every $g \in G$, we have

$$\chi(g) = \sum_{i=1}^{r} \sum_{\sigma \in \Sigma_i} a_{i,\sigma} \sigma(\psi_i(g)).$$

We know that $\chi(g) \in k$ if and only if, for every $\tau \in \mathrm{Gal}(\Omega/k)$, $\tau(\chi(k)) = \chi(k)$. For such a τ, we have

$$\tau(\chi(g)) = \sum_{i=1}^{r} \sum_{\sigma \in \Sigma_i} a_{i,\sigma} \tau(\sigma(\psi_i(g))) = \sum_{i=1}^{r} \sum_{\sigma \in \Sigma_i} a_{i,\tau^{-1}\sigma} \sigma(\psi_i(g)),$$

because $a_{i,\sigma} \in \mathbb{Z}$ (so $\tau(a_{i,\sigma}) = a_{i,\sigma}$) and right multiplication by τ (and τ^{-1}) preserves Σ_i. To finish the proof, we just have to note that, for every $\sigma, \sigma' \in \Sigma_i$, there exists $\tau \in \mathrm{Gal}(\Omega/k)$ such that $\tau\sigma = \sigma'$.

(7). By (3) and (6), a basis of $\mathrm{R}'_k(G)$ is given by the elements $\sum_{\sigma \in \Sigma_i} \psi_{i,\sigma}$, for $i \in \{1, \dots, r\}$. By (5), those are equal to the $m_i^{-1}\chi_i$.

□

VII.2.8 Hecke algebra

Let k be an algebraically closed field of characteristic 0.

Let $H \subset G$ be finite groups, and let $\chi : H \to k^\times$ be a morphism of groups (i.e. a 1-dimensional representation of H over k).

Let $\mathscr{H}$ be the space of functions $f : G \to k$ such that, for every $h, h' \in H$ and $g \in G$, $f(hgh') = \chi(hh')f(g)$. The *convolution* of two functions f_1 and f_2 of $\mathscr{H}$ is the function $f_1 * f_2$ defined by :

$$(f_1 * f_2)(g) = \frac{1}{|H|} \sum_{x \in G} f_1(x) f_2(x^{-1}g).$$

Using Frobenius reciprocity (theorem I.5.4.3 of chapter I) and problem VII.1.12, construct an isomorphism (of k-vector spaces) $\mathrm{End}_{k[G]}(\mathrm{Ind}_H^G \chi) \xrightarrow{\sim} \mathscr{H}$, and show that it sends the multiplication of $\mathrm{End}_{k[G]}(\mathrm{Ind}_H^G \chi)$ to the convolution on $\mathscr{H}$. [10]

[10] In particular, the convolution of two functions of $\mathscr{H}$ is still in $\mathscr{H}$, and the convolution makes $\mathscr{H}$ into a k-algebra. Both these facts can also easily be checked directly.

Solution. Let $\mathscr{E} = \mathrm{End}_{k[G]}(\mathrm{Ind}_H^G k_\chi)$. Frobenius reciprocity gives an isomorphism of k-vector spaces $\varphi : \mathscr{E} \xrightarrow{\sim} \mathrm{Hom}_{k[H]}(k_\chi, \mathrm{Res}_H^G \mathrm{Ind}_H^G k_\chi)$, and problem VII.1.12 gives an isomorphism of $k[G]$-modules $\mathrm{Ind}_H^G \xrightarrow{\sim} \mathscr{F}$, where

$$\mathscr{F} = \{f : G \to k | \forall h \in H,\ g \in G,\ f(hg) = \chi(h)f(g)\},$$

with the action of G given by $(gf)(x) = f(xg)$ for any $g, x \in G$ and $f \in \mathscr{F}$. Putting these two results together, we get an injective k-linear map $\psi : \mathscr{E} \to \mathscr{F}$, $u \longmapsto \varphi(u)(1)$. (This map is injective because $\varphi(u) \in \mathrm{Hom}_{k[H]}(k_\chi, \mathrm{Res}_H^G \mathrm{Ind}_H^G k_\chi)$ is k-linear, hence uniquely determined by the image of 1.) Note that $\mathscr{H} \subset \mathscr{F}$ (as k-vector spaces).

I claim that the image of ψ is $\mathscr{H}$. Indeed, let $u \in \mathscr{E}$. Then $\varphi(u)$ is H-linear, so, if $f = \psi(u)$, then for every $h \in H$,

$$hf = h(\varphi(u)(1)) = \varphi(u)(h \cdot 1) = \varphi(u)(\chi(h)) = \chi(h)\varphi(u)(1) = \chi(h)f,$$

and hence for every $g \in G$,

$$(hf)(g) = f(gh) = \chi(h)f(g).$$

So $f = \psi(u)$ is indeed in $\mathscr{H}$. Conversely, let $f \in \mathscr{H}$, and define a k-linear map $u : k_\chi \to Res_H^G \mathrm{Ind}_H^G k_\chi$ by setting $u(1) = f$. Then for every $h \in H$, for every $g \in G$,

$$u(h \cdot 1)(g) = (\chi(h)u(1))(g) = \chi(h)f(g) = f(gh) = (hf)(g) = (hu(1))(g),$$

so that u is actually $k[H]$-linear. We obviously have $\psi(u) = f$.

Finally, we have to check that ψ sends the composition on $\mathscr{E}$ to the convolution on $\mathscr{H}$. Let $u_1, u_2 \in \mathscr{E}$, and let $f_1 = \psi(u_1)$, $f_2 = \psi(u_2)$. First we identify $\varphi(u_1u_2)$. Remember that we have an injective map of $k[H]$-modules

$$k_\chi = k[H] \otimes_{k[H]} k_\chi \to k[G] \otimes_{k[H]} k_\chi = \mathrm{Res}_H^G \mathrm{Ind}_H^G k_\chi,$$

and that φ is given by restriction to $k_\chi \subset \mathrm{Res}_H^G \mathrm{Ind}_H^G k_\chi$. So $\varphi(u_1u_2)$ is the composition

$$k_\chi \xrightarrow{\varphi(u_2)} \mathrm{Res}_H^G \mathrm{Ind}_H^G k_\chi \xrightarrow{Res_H^G u_1} \mathrm{Res}_H^G \mathrm{Ind}_H^G k_\chi.$$

Next we identify the endomorphism of $\mathscr{F}$ corresponding to u_1. Remember that the morphism $a : \mathrm{Ind}_H^G k_\chi \to \mathscr{F}$ sends $g \otimes v$ to the function $x \longmapsto \begin{cases} \chi(h)v & \text{if } x = hg,\ h \in H \\ 0 & \text{otherwise} \end{cases}$. In the other direction, if $(g_i)_{i \in I}$ is a system of representatives of G/H, then the morphism $\mathscr{F} \to \mathrm{Ind}_H^G k_\chi$ sends f to $\sum_{i \in I} g_i \otimes f(g_i^{-1})$. (This is from problem VII.1.12.) So the endomorphism of $\mathscr{F}$ corresponding to u_1 sends $f \in \mathscr{F}$ to

$$au_1\left(\sum_{i \in I} g_i \otimes f(g_i^{-1})\right) = a\left(\sum_{i \in I} g_i u_1(1 \otimes f(g_i^{-1}))\right) = a\left(\sum_{i \in I} g_i(1 \otimes \varphi(u_1)(f(g_i^{-1}))\right)$$

$$= \sum_{i \in I} g_i f_1(f(g_i^{-1})) = \sum_{i \in I} f(g_i^{-1})(g_i f_1).$$

We get $\psi(u_1 u_2)$ by applying this endomorphism to f_2. So

$$\psi(u_1 u_2) = \sum_{i \in I} f_2(g_i^{-1})(g_i f_1),$$

and, for every $x \in G$,

$$\psi(u_1 u_2)(x) = \sum_{i \in I} f_2(g_i^{-1}) f_1(x g_i) = \frac{1}{|H|} \sum_{g \in G} f_1(xg) f_2(g^{-1}).$$

On the other hand, for every $x \in G$,

$$(f_1 * f_2)(x) = \frac{1}{|H|} \sum_{g' \in G} f_1(g') f_2((g')^{-1} x).$$

Making the change of variables $g' = xg$, we see that these two sums are equal.

□

VII.2.9 Multiplicity-free modules

We say that a $k[G]$-module V is *multiplicity-free* if the multiplicity of every simple $k[G]$-module in V is at most 1 (i.e. $V = W_1 \oplus \cdots \oplus W_r$, where the W_i are pairwise non-isomorphic simple $k[G]$-modules).

Show that V is multiplicity-free if and only if $\mathrm{End}_{k[G]}(V)$ is commutative.

Solution. Write $V = \bigoplus_{W \in S_k(G)} W^{\oplus n(W)}$. Then

$$\mathrm{End}_{k[G]}(V) = \prod_{W \in S_k(G)} M_{n(W)}(k),$$

and this is commutative if and only $n(w) \leq 1$ for every $W \in S_k(G)$.

□

VII.2.10 Hecke algebra and multiplicities

[11]

We use the notation of problem VII.2.8 For every $g \in G$, we write $H_g = H \cap gHg^{-1}$ and $\chi^g : H_g \to k^\times$, $h \longmapsto \chi(g^{-1}hg)$. Let $\alpha : G \to G$ be a bijection such that :

- $\alpha(gh) = \alpha(h)\alpha(g), \forall h, g \in G$;
- $\alpha(H) = H$, and $\chi \circ \alpha_{|H} = \chi$;
- for every $g \in G$, if $\mathrm{Hom}_{H_g}(\chi_{|H_g}, \chi^g) \neq 0$, then there exists $g' \in HgH$ such that $\alpha(g') = g'$.

We define a k-linear automorphism $\widetilde{\alpha}$ of $\mathscr{H}$ by sending f to $\widetilde{\alpha}(f) = f \circ \alpha$.

(1). Show that $\widetilde{\alpha}$ does indeed any f in $\mathscr{H}$ to an element of $\mathscr{H}$, and that $\widetilde{\alpha}(f_1 * f_2) = \widetilde{\alpha}(f_2) * \widetilde{\alpha}(f_1), \forall f_1, f_2 \in \mathscr{H}$.

(2). Show that α is the identity map on $\mathscr{H}$. (Hint : Can you find a basis of $\mathscr{H}$?)

(3). Show that $\mathrm{Ind}_H^G \chi$ is multiplicity-free.

Solution. We first note that $\alpha(1) = \alpha(1^2) = \alpha(1)^2$, so $\alpha(1) = 1$. Consequently, for every $g \in G$,

$$1 = \alpha(1) = \alpha(gg^{-1}) = \alpha(g^{-1})\alpha(g),$$

so that $\alpha(g^{-1}) = \alpha(g)^{-1}$.

(1). The first part follows easily from the first two conditions on α.

Let $x \in G$. Then

$$\widetilde{\alpha}(f_1 * f_2)(x) = (f_1 * f_2)(\alpha(x)) = \frac{1}{|H|} \sum_{g \in G} f_1(g) f_2(g^{-1}\alpha(x)).$$

On the other hand,

$$\begin{aligned}(\widetilde{\alpha}(f_2) * \widetilde{\alpha}(f_1))(x) &= \frac{1}{|H|} \sum_{g \in G} f_2(\alpha(g)) f_1(\alpha(g^{-1}x)) \\ &= \frac{1}{|H|} \sum_{g \in G} f_2(\alpha(g)) f_1(\alpha(x)\alpha(g)^{-1}) \\ &= \frac{1}{|H|} \sum_{g' \in G} f_1(g') f_2((g')^{-1}\alpha(x)) = \widetilde{\alpha}(f_1 * f_2)(x)\end{aligned}$$

(using the change of variables $g' = \alpha(x)\alpha(g)^{-1}$).

[11] this due to somebody ?

(2). For every $g \in G$, let e_g be the characteristic function of HgH. Note that this function depends only on the class of g in $H\backslash G/H$, and that $\sum_{g\in H\backslash G/H} e_g = 1$ (the constant function 1 on G). If $f \in \mathscr{H}$, then $f = \sum_{g\in H\backslash G/H} fe_g$, and $fe_g \in \mathscr{H}$ for every g. To show that $\widetilde{\alpha}(f) = f$, it suffices to show that $\widetilde{\alpha}(fe_g) = fe_g$ for every $g \in G$. In other words, we may assume that f is supported on some double class HgH, with $g \in G$. By the first two conditions on α, $\alpha(HgH) = H\alpha(g)H$, so $\widetilde{\alpha}(f)$ is supported on the double class $H\alpha(g)H$. There are two possibilities :

(a) If there exist $h, h' \in H$ such that $hgh' = g$, this means that $H_g = H \cap gHg^{-1} \neq \{1\}$. If $f(g) \neq 0$, then $\chi(hh') = 1$ (because $f(g) = f(hgh') = \chi(hh')f(g)$), so $\chi^g(h) = \chi((h')^{-1}) = \chi(h)$, so $\mathrm{Hom}_{H_g}(\chi_{H_g}, \chi^g) \neq 0$. By the third assumption on α, there exists $g' \in HgH$ such that $\alpha(g') = g'$. We have $\widetilde{\alpha}(f)(g') = f(\alpha(g')) = f(g')$. As $\widetilde{\alpha}(f)$ and f are both in $\mathscr{H}$ and supported on $Hg'H$, this implies that $\widetilde{\alpha}(f) = f$.

(b) Otherwise, $H_g = \{1\}$, so $\mathrm{Hom}_{H_g}(\chi_{|H_g}, \chi^g) = k \neq 0$, so there exists $g' \in HgH$ such that $\alpha(g') = g'$. As in the first case, this implies that $\widetilde{\alpha}(f) = f$.

(3). By (1) and (2), $f_1 * f_2 = f_1 * f_1$ for every $f_1, f_2 \in \mathscr{H}$. By problem VII.2.8, $\mathrm{End}_{k[G]}(\mathrm{Ind}_H^G \chi)$ is commutative. By problem VII.2.9, this implies that $\mathrm{Ind}_H^G \chi$ is multiplicity-free.

□

VII.2.11 Characters of a finite field

Let $\mathbb{F}_q$ be a finite field. We write $\widehat{\mathbb{F}}_q$ for the set of group morphisms $\psi : (\mathbb{F}_q, +) \to k^\times$. Let $\psi \neq 1$ be an element of $\widehat{\mathbb{F}}_q$. Show that the map

$$\left\{ \begin{array}{ccc} \mathbb{F}_q & \to & \widehat{\mathbb{F}}_q \\ a & \longmapsto & (x \longmapsto \psi(ax)) \end{array} \right.$$

is a bijection.

Solution. Let's call this map u. First we show that u is injective. Let $a, b \in \mathbb{F}_q$ such that $u(a) = u(b)$. Then, for every $x \in \mathbb{F}_q$, $\psi((a-b)x) = \psi(ax)\psi(bx)^{-1} = 1$. As $\mathbb{F}_q$ is a field, if $a - b \neq 0$, we get that $\psi(y) = 1$ for every $y \in \mathbb{F}_q$, which contradicts the fact that $\psi \neq 1$. So $a - b = 0$, ie $a = b$.

Now note that the group $\mathbb{F}_q$ is commutative, so all its irreducible representations over k are of dimension 1, so $\widehat{\mathbb{F}}_q$ is actually $S_k(\mathbb{F}_q)$. Also, the conjugacy classes in $\mathbb{F}_q$ are singletons (again using the commutativity of the additive group $\mathbb{F}_q$), so there are $|\mathbb{F}_q|$ of them, and we get $|\widehat{\mathbb{F}}_q| = |\mathbb{F}_q|$. As the map $u : \mathbb{F}_q \to \widehat{\mathbb{F}}_q$ is injective, it is automatically bijective.

□

VII.2.12 Representations of $\mathrm{GL}_2(\mathbb{F}_q)$, part 2

Let $\mathbb{F}_q$ be a finite field, let $G = GL_2(\mathbb{F}_q)$, let $N = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \subset G$. We identify N with the additive group $\mathbb{F}_q$ by sending $x \in \mathbb{F}_q$ to $\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$. For every $\psi \in \widehat{\mathbb{F}}_q - \{1\}$ (see problem VII.2.11), we write $V_\psi = \mathrm{Ind}_N^G \psi$.

(1). Show that, if $\psi, \psi' \in \widehat{\mathbb{F}}_q - \{1\}$, then $V_\psi \simeq V_{\psi'}$.

This representation V_ψ is called the *Gelfand-Graev representation* of $\mathrm{GL}_2(\mathbb{F}_q)$.

We fix $\psi \in \widehat{\mathbb{F}}_q - \{1\}$.

(2). Show that V_ψ is multiplicity-free. *(Hint :* $\alpha \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d & b \\ c & a \end{pmatrix}$.*)*

(3). Calculate $|G|$, find all the conjugacy classes in G and their cardinalities.

(4). Calculate the character of V_ψ.

(5). Show that $V_\psi \simeq (V_\psi)^*$.

(6). Calculate the number of simple components of V_ψ.

(7). Let W be a simple $k[G]$-module. Assume that $q > 2$. Show that the following are equivalent :

(a) $\langle W, V_\psi \rangle_G \neq 0$ (ie W is a simple component of V_ψ);

(b) $\langle \psi, \mathrm{Res}_N^G W \rangle_N \neq 0$;

(c) for every $\psi' \in \widehat{\mathbb{F}}_q - \{1\}$, $\langle \psi', \mathrm{Res}_N^G W \rangle_N \neq 0$;

(d) $\dim_k(W) \geq 2$.

(You can admit the following useful facts : (1) The commutator subgroup of G is $SL_2(\mathbb{F}_q)$. (2) The group $SL_2(\mathbb{F}_q)$ is generated by N and by $N' := \begin{pmatrix} 1 & 0 \\ * & 0 \end{pmatrix}$.)

(If $q = 2$, then $GL_2(\mathbb{F}_q) = SL_2(\mathbb{F}_q) \simeq \mathfrak{S}_3$, and the useful fact is not true anymore.)

Solution.

(1). By problem VII.2.11, there exists a unique $a \in \mathbb{F}_q$ such that $\psi'(x) = \psi(ax)$ for every $x \in \mathbb{F}_q$. As $\psi' \neq 1$, $a \neq 0$. Let $t = \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \in G$. Then, if $n = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \in N$, we have $tnt^{-1} = \begin{pmatrix} 1 & ax \\ 0 & 1 \end{pmatrix}$, so that $\psi'(n) = \psi(tnt^{-1})$. We define $u : V_\psi \to V_{\psi'}$ by $u(g \otimes x) = (gt) \otimes x$, for every $g \in G$ and $x \in k_\chi$. Then :

- u is well-defined : Let $n \in N$, $g \in G$, $x \in k_\psi$. We have to show that $u((gn) \otimes x) = u(g \otimes (nx))$. But

  $$u((gn)\otimes x) = (gnt)\otimes x = ((gt)(t^{-1}nt))\otimes x = (gt)\otimes(\psi'(t^{-1}nt)x) = \psi(n)((gt)\otimes x)$$

  (using that $t^{-1}nt \in N$), and

  $$u(g \otimes (nx)) = u(\psi(n)(g \otimes x)) = \psi(n)((gt) \otimes x).$$

- u is obviously $k[G]$-linear.
- u is an isomorphism, because it has an inverse v given by $v(g \otimes x) = (gt^{-1}) \otimes x$. (The proof that v is well-defined is similar to the proof that u is well-defined.)

(2). Let's use the hint and apply problem VII.2.10 with α given by $\alpha \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d & b \\ c & a \end{pmatrix}$. The second condition of problem VII.2.10 is clear. Let $w = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Then $w = w^{-1}$ and $\alpha(g) = wg^T w$ for every $g \in G$, so the first condition is also clear. Obviously the third condition for $g \in G$ depends only on the double class NgN, so we start by finding representatives for these double classes. Let $T = \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}$. If $G = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \in T$, then an easy calculation gives

$$NgN = \left\{ \begin{pmatrix} x & y \\ z & t \end{pmatrix} \in G | z = 0,\ x = a \text{ and } t = b \right\}$$

and

$$NwgN = \left\{ \begin{pmatrix} x & y \\ z & t \end{pmatrix} \in G | z = a \text{ and } y - t^{-1}xz = b \right\}.$$

So

$$G = \left(\coprod_{t \in T} NtN \right) \sqcup \left(\coprod_{t \in T} NwtN \right).$$

Let $t = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \in T$. Consider the double class NtN. We have $tNt^{-1} = N$, so $N_t = N$, and ψ^t is given by

$$\psi^t \left(\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \right) = \psi \left(t^{-1} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} t \right) = \psi \left(\begin{pmatrix} 1 & a^{-1}bx \\ 0 & 1 \end{pmatrix} \right),$$

so $\mathrm{Hom}_{N_t}(\psi, \psi^t) \neq 0$ if and only $\psi = \psi^t$, if and only if $a^{-1}b = 1$, ie $a = b$. In that case, we have $\alpha(t) = t$.

Now consider the double class $NwtN$. We have $wtN(wt)^{-1} = wNw^{-1} = \begin{pmatrix} 1 & 0 \\ * & 1 \end{pmatrix}$, so $N_{wt} = \{1\}$, so $\mathrm{Hom}_{N_{wt}}(\chi_{|N_{wt}}, \chi^t) = k \neq 0$. On the other hand, $wt = \begin{pmatrix} 0 & b \\ a & 0 \end{pmatrix}$, so $\alpha(wt) = wt$.

(3). The cardinality of G is the number of bases in $\mathbb{F}_q^2$, ie $(q^2-1)(q^2-q)$. We already found representatives of the conjugacy classes in the solution of (2) of problem VII.2.6, they are :

(a) $t = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, $a, b \in \mathbb{F}_q^\times$. The corresponding class has cardinality $|G/Z_G(t)|$. If $a \neq b$, then $Z_G(T) = T$, so $|G/Z_G(t)| = q(q+1)$. If $a = b$, then $Z_G(t) = G$, so $|G/Z_G(t)| = 1$.

(b) $g = \begin{pmatrix} a & b \\ ub & a \end{pmatrix}$, where u is a fixed element of $\mathbb{F}_q^\times - (\mathbb{F}_q^\times)^2$ and $a, b \in \mathbb{F}_q$ are such that $a^2 - ub^2 \neq 0$ and $b \neq 0$. Then we easily see that

$$Z_G(g) = \left\{ \begin{pmatrix} a' & b' \\ ub' & a' \end{pmatrix}, \text{ with } a', b' \in \mathbb{F}_q \text{ st } (a')^2 - u(b')^2 \neq 0 \right\},$$

so $|Z_G(g)| = q^2 - 1$ and $|G/Z_G(g)| = q^2 - q$.

(c) $n = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}$, $a \in \mathbb{F}_q^\times$. Then $Z_G(n) = \left\{ \begin{pmatrix} x & y \\ 0 & x \end{pmatrix}, x, y \in \mathbb{F}_q, x \neq 0 \right\}$, so $|Z_G(n)| = q(q-1)$ and $|G/Z_G(n)| = q^2 - 1$.

(4). We use the formula for the character of an induced representation. The only conjugacy classes that intersect N are the ones with representatives $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. So $\chi_{V_\psi}(g) = 0$ if g is not conjugate to one of these two matrices. Also,

$$\chi_{V_\psi}\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) = \dim_k(V_\psi) = |G/N| = (q^2-1)(q-1).$$

Finally, if $n = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, then $\chi_{V_\psi}(n) = \frac{1}{|N|} \sum_{g \in G,\ gng^{-1} \in N} \chi(gng^{-1})$. We see easily that $gng^{-1} \in N$ if and only if $g \in B$, where $B = \begin{pmatrix} * & * \\ 0 & * \end{pmatrix}$. If $g = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in B$, then $gng^{-1} = \begin{pmatrix} 1 & ad^{-1} \\ 0 & 1 \end{pmatrix}$. So

$$\chi_{V_\psi}(n) = \frac{1}{q} \sum_{a,d \in \mathbb{F}_q^\times, b \in \mathbb{F}_q} \psi(ad^{-1}) = (q-1) \sum_{x \in \mathbb{F}_q^\times} \psi(x) = 1 - q,$$

because

$$\sum_{x\in\mathbb{F}_q^\times}\psi(x) = -1 + \sum_{x\in\mathbb{F}_q}\psi(x) = -1 + \langle\psi, 1\rangle_{\mathbb{F}_q} = -1$$

(as $\psi \neq 1$).

Note that this calculation gives another proof of the result of (1).

(5). It is enough to show that $\chi_{V_\psi} = \chi_{V_\psi^*}$. But $\chi_{V_\psi^*}(g) = \chi_{V_\psi}(g^{-1})$, for every $g \in G$. Let n be as in the solution of (d). As $n^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ is in the conjugacy class of n, $\psi_{V_\psi}(n^{-1}) = \psi_{V_\psi}(n)$. By (4), this gives the desired result.

(6). If $V_\psi = \bigoplus_{W\in S_k(G)} W^{\oplus n(W)}$, then $\langle V_\psi, V_\psi\rangle_G = \sum_{W\in S_k(G)} n_W^2$. By (2), V_ψ is multiplicity-free, so $\sum_{W\in S_k(G)} n_W^2$ is the number of simple components of V_ψ. So we have to calculate $\langle V_\psi, V_\psi\rangle_G$. By definition and using (3), (4) and (5), $\langle V_\psi, V_\psi\rangle_G$ is equal to

$$\frac{1}{|G|}\sum_{g\in G}\chi_{V_\psi}(g)^2 = \frac{1}{(q^2-1)(q^2-q)}\left(((q^2-1)(q-1))^2 + (1-q)^2(q^2-1)\right) = q(q-1).$$

(7). By Frobenius reciprocity, $\langle W, V_\psi\rangle_G = \langle\psi, \mathrm{Res}_N^G W\rangle_N = \dim_k \mathrm{Hom}_{k[N]}(\psi, \mathrm{Res}_H^G W)$. This gives the equivalence of (a) and (b).

Using Frobenius reciprocity as above and (1), we get the fact that (a) implies (c). As (c) obviously implies (b), it also implies (a), so (a) and (c) are equivalent.

Suppose that $\langle W, V_\psi\rangle_G = 0$. Then, using (1) and the fact that (a) implies (b), we get $\langle\psi', \mathrm{Res}_N^G W\rangle_N = 0$ for every $\psi' \in \widehat{\mathbb{F}}_q - \{1\}$. So N acts trivially on W. Let N' be as in the useful admitted fact, the $N' = wNw^{-1}$ with w as in the solution of (2), so N' also acts trivially on W. As $SL_2(\mathbb{F}_q)$ is generated by N and N', $SL_2(\mathbb{F}_q)$ acts trivially on W, so the action of G on W factors through $G/SL_2(\mathbb{F}_q) \xrightarrow{\sim} \mathbb{F}_q^\times$ (the isomorphism being given by the determinant). As $\mathbb{F}_q^\times$ is abelian and W is irreducible, $\dim_k W = 1$. This show that (d) implies (a).

Suppose that $\dim_k(W) = 1$. Then the action of G on W is given by a morphism of groups $G \to k^\times$, so it factors through the abelianization of G, which is $G/SL_2(\mathbb{F}_q)$ by the useful fact. As $N \subset SL_2(\mathbb{F}_q)$, this show that N acts trivially on W, so (b) is false. This show that (b) implies (d).

□

VII.2.13 Representations of $\mathrm{GL}_2(\mathbb{F}_q)$, part 3

Let $\mathbb{F}_q$ be a finite field, $G = GL_2(\mathbb{F}_q)$, B be the subgroup of upper triangular matrices in G, N be the subgroup of unipotent upper triangular matrices in G (ie elements of B with both diagonal

entries equal to 1), T be the subgroup of diagonal matrices in G.

If ω_1, ω_2 are morphisms $\mathbb{F}_q^\times \to \mathbb{C}^\times$, we used them in problem VII.2.6 to define a representation $I(\omega_1, \omega_2)$ of G (that was denoted by $\mathrm{Ind}_B^G \rho$ there). We saw that $I(\omega_1, \omega_2)$ is irreducible if and only if $\omega_1 \neq \omega_2$.

(1). Show that $I(1,1) = \mathbb{1} \oplus \mathrm{St}$, with St an irreducible representation of G. (This representation St is called the *Steinberg representation*.)

(2). Calculate the character of St.

(3). For every morphism of groups $\omega : \mathbb{F}_q^\times \to \mathbb{C}^\times$, show that $I(\omega, \omega) = (\omega \circ \det) \oplus ((\omega \circ \det) \otimes \mathrm{St})$.

(4). We say that an irreducible representation V of G is *cuspidal* if $V^N = \{0\}$. Show that V is cuspidal if and only if it is not a simple constituent of one of the $I(\omega_1, \omega_2)$. (Simple constituents of the $I(\omega_1, \omega_2)$ are called *principal series representations*.)

(5). Find the number of isomorphism classes of cuspidal representations of G.

(6). Suppose that $q > 2$. Let V be a cuspidal representation of G. Show that $\dim_{\mathbb{C}}(V) = q - 1$. *(Hint : for every nontrivial morphism $\psi : N \to \mathbb{C}^\times$, show that $\langle \psi, \mathrm{Res}_N^G V \rangle_N = 1$.)*

Solution.

(1). Let $V = I(1,1)$. We have

$$\langle V, V \rangle_G = \langle \mathbb{1}, \mathrm{Res}_B^G V \rangle_B = \frac{1}{|B|} \sum_{g \in B} \chi_V(g).$$

We have calculated χ_V in problem VII.2.6. If $g = \begin{pmatrix} a & b \\ 0 & a \end{pmatrix}$ with $b \neq 0$, then $\chi_V(g) = 1$, and there $(q-1)^2$ such elements in B. If $g = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$, then $\chi_V(g) = 1 + q$, and there $q - 1$ such elements in B. If $g = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$ with $a \neq c$, then $\chi_V(g) = 2$, and there $(q-1)(q-2)q$ such elements in B. Also, $|B| = (q-1)^2 q$. So finally

$$\langle V, V \rangle_G = \frac{1}{(q-1)^2 q}((q-1)^2 + (q+1)(q-1) + 2(q-1)(q-2)q) = 2,$$

and this implies that V has two irreducible components. As

$$\mathrm{Hom}_G(V, \mathbb{1}) = \mathrm{Hom}_B(\mathbb{1}, \mathbb{1}) = \mathbb{C},$$

one of these components is $\mathbb{1}$. So $V = \mathbb{1} \oplus \mathrm{St}$, with St an irreducible representation of G.

(2). We have $\chi_{\mathrm{St}} = \chi_{I(1,1)} - 1$. So $\chi_{\mathrm{St}}(g) = 1$ if g has two distinct eigenvalues in $\mathbb{F}_q$, $\chi_{\mathrm{St}}(g) = q$ if $g = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ with $a \in \mathbb{F}_q^\times$, and $\chi_{\mathrm{St}}(g) = 0$ otherwise.

(3). Let $\chi = \omega \det : G \to \mathbb{C}$. Then

$$\chi \oplus (\chi \otimes \mathrm{St}) = \chi \otimes I(1,1) = \mathrm{Ind}_B^G(\chi_{|B} \otimes \mathbb{1}) = I(\omega, \omega).$$

(4). For $\omega_1, \omega_2 : \mathbb{F}_q^\times \to \mathbb{C}^\times$, we use the same notation $\omega_1 \otimes \omega_2$ to denote the map $B \to \mathbb{C}$, $\begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \longmapsto \omega_1(a)\omega_2(c)$, and its restriction to T. We also use the same notation for the representations they define.

I claim that, for every $\mathbb{C}[B]$-module W, the inclusion

$$\mathrm{Hom}_T(\omega_1 \otimes \omega_2, W^N) = \mathrm{Hom}_B(\omega_1 \otimes \omega_2, W^N) \subset \mathrm{Hom}_B(\omega_1 \otimes \omega_2, W)$$

is an equality. Indeed, N acts trivially on $\omega_1 \otimes \omega_2$, so every element of $\mathrm{Hom}_B(\omega_1 \otimes \omega_2, W)$ has image contained in W^N.

Now let's apply this to a representation V of G. For every $\omega_1, \omega_2 : \mathbb{F}_q^\times \to \mathbb{C}^\times$, we have

$$\mathrm{Hom}_G(I(\omega_1, \omega_2), V) = \mathrm{Hom}_B(\omega_1 \otimes \omega_2, \mathrm{Res}_B^G V) = \mathrm{Hom}_T(\omega_1 \otimes \omega_2, (\mathrm{Res}_B^G V)^N).$$

If V is cuspidal, then this is 0 for any ω_1 and ω_2, so V cannot be a simple constituent of a $I(\omega_1, \omega_2)$. If V is not cuspidal, then the representation $(\mathrm{Res}_B^G V)^N$ of T is not trivial, so it contains some $\omega_1 \otimes \omega_2$, and then we have a nonzero morphism $I(\omega_1, \omega_2) \to V$. As V is irreducible, this morphism is surjective, so V is a simple constituent of $I(\omega_1, \omega_2)$.

(5). First let's calculate the number of isomorphism classes of irreducible representations of G. This is the same as the number of conjugacy classes in G. We have $q-1$ elements in the center of G, $q-1$ conjugacy classes of non-diagonalizable elements, $(q-1)(q-2)/2$ conjugacy classes of elements that have distinct eigenvalues in $\mathbb{F}_q$, and $(q^2-q)/2$ conjugacy classes of elements that have eigenvalues in $\mathbb{F}_{q^2} - \mathbb{F}_q$. So this gives $(q-1)(q+1)$ conjugacy classes in total.

Now let's count isomorphism classes of non-cuspidal irreducible representations. If $\omega_1, \omega_2 : \mathbb{F}_1^\times \to \mathbb{C}^\times$ are distinct, then $I(\omega_1, \omega_2)$ is irreducible. Moreover, the calculation of the character of $I(\omega_1, \omega_2)$ in (2) of problem VII.2.6 shows that $I(\omega_1, \omega_2) \simeq I(\omega_1', \omega_2')$ if and only if $(\omega_1, \omega_2) = (\omega_1', \omega_2')$ or (ω_2', ω_1'). So we get $(q-1)(q-2)/2$ irreducible representations this way. For every $\omega : \mathbb{F}_q^\times \to \mathbb{C}^\times$, we also get the representations $\omega \circ \det$ and $(\omega \circ \det) \otimes \mathrm{St}$, hence $2(q-1)$ representations in total. So the number of isomorphism classes of cuspidal irreducible representations is

$$(q-1)(q+1) - (q-1)(q-2)/2 - 2(q-1) = q(q-1)/2.$$

(6). By (7) of problem VII.2.12, for every nontrivial morphism $\psi : N \to \mathbb{C}^\times$,

$$\langle \psi, \mathrm{Res}_N^G V \rangle_N = \langle \mathrm{Ind}_N^G \psi, V \rangle_G = 1.$$

As $V^N = 0$, $\mathrm{Res}_N^G V$ is a sum of nontrivial irreducible representations of N. As N is commutative, all these representations are of dimension 1. By the calculation above, every nontrivial dimension 1 representation of N appears in $\mathrm{Res}_N^G V$ with multiplicity 1. So $\dim_C V = |N| - 1 = q - 1$.

□

VII.2.14 Representations of $\mathrm{GL}_2(\mathbb{F}_q)$, part 4

We use the notation of problem VII.2.13. Fix an element $u \in \mathbb{F}_q^\times - (\mathbb{F}_q^\times)^2$, and let

$$E = \left\{ \begin{pmatrix} a & b \\ ub & a \end{pmatrix},\ a, b \in \mathbb{F}_q \right\} \subset M_2(\mathbb{F}_q).$$

If $g = \begin{pmatrix} a & b \\ ub & a \end{pmatrix}$ is an element of E, we set

$$\overline{g} = \begin{pmatrix} a & -b \\ -ub & a \end{pmatrix} \in E$$

and $N(g) = a^2 - ub^2 \in \mathbb{F}_q$. We also set $S = E \cap G$.

(1). Show that E is a commutative subring of $M_2(\mathbb{F}_q)$, that it is isomorphic to $\mathbb{F}_{q^2}$, and that this identifies S with $\mathbb{F}_{q^2}^\times$.

(2). Show that the map $E \to E$, $g \longmapsto \overline{g}$, is a morphism of rings and that $N : S \to \mathbb{F}_q^\times$ is a morphism of groups.

(3). Fix a nontrivial group morphism $\psi : \mathbb{F}_q \to \mathbb{C}^\times$. If $f : E \to \mathbb{C}$ is a function, we define another function $\widehat{f} : E \to \mathbb{C}$ by

$$\widehat{f}(x) = -q^{-1} \sum_{y \in E} f(y)\psi(\mathrm{Tr}(\overline{x}y)).$$

Show that $\widehat{\widehat{f}}(x) = f(-x)$.

(4). We say that an element u of G is *unipotent* if $u - 1 \in M_2(\mathbb{F}_q)$ is nilpotent.

Let A be a maximal commutative subgroup of G, and suppose 1 is the only unipotent element of A. Show that there exists $g \in G$ such that $A = gTg^{-1}$ or $A = gSg^{-1}$.

(We will use the group S to construct the cuspidal representations of G in the next problem.)

Solution.

(1). Note that, if q is even, then $\mathbb{F}_q^\times = (\mathbb{F}_q^\times)^2$. So the problem is empty in this case, and so we may assume that q is odd.

The set E is clearly stable by addition and substraction. If $g = \begin{pmatrix} a & b \\ ub & a \end{pmatrix}$ and $g' = \begin{pmatrix} a' & b' \\ ub' & a' \end{pmatrix}$, then $gg' = g'g = \begin{pmatrix} aa'+ubb' & ab'+a'b \\ u(ab'+a'b) & aa'+ubb' \end{pmatrix} \in E$ and, if $\det(g) = N(g) \neq 0$, then $g^{-1} = N(g)^{-1}\overline{g} \in E$.

So E is a commutative subring of $M_2(\mathbb{F}_q)$ and its set of invertible elements is S. Note also that E has q^2 elements.

Let $g = \begin{pmatrix} a & b \\ ub & a \end{pmatrix}$ be non-invertible. Then $a^2 = ub^2$. As u is not a square in $\mathbb{F}_q$, this is only possible if $a = b = 0$, ie $g = 0$. So E is a field with q^2 elements, and hence is isomorphic to $\mathbb{F}_{q^2}$. We have already seen that $S = E^\times$.

(2). The map $N : S \to \mathbb{F}_q^\times$ is a morphism of groups because it is the restriction to S of the determinant on $GL_2(\mathbb{F}_q)$. Showing that $g \longmapsto \overline{g}$ is a morphism of rings is an easy calculation. (We can also notice that, by (1), this morphism identifies to the action of the nontrivial element of $\mathrm{Gal}(\mathbb{F}_{q^2}/\mathbb{F}_q)$ on $\mathbb{F}_{q^2}$.)

(3). Let $F = \widehat{\widehat{f}}$. Let $x \in E$. Then

$$F(x) = -q^{-1}\sum_{y\in E} \widehat{f}(y)\psi(\mathrm{Tr}(\overline{x}y)) = q^{-2}\sum_{y,z\in E} f(z)\psi(\mathrm{Tr}(\overline{y}z))\psi(\mathrm{Tr}(\overline{x}y)).$$

As $\mathrm{Tr}(g) = \mathrm{Tr}(\overline{g})$ for every $g \in E$, $\mathrm{Tr}(\overline{x}y) = \mathrm{Tr}(x\overline{y}) = \mathrm{Tr}(\overline{y}z)$ for every $y \in E$. As ψ is a morphism of groups from E to $\mathbb{C}$,

$$F(x) = q^{-2}\sum_{z\in E} f(z)\sum_{y\in E}\psi(\mathrm{Tr}(\overline{y}(x+z))).$$

If z is fixed, then $y \longmapsto \psi(\mathrm{Tr}(\overline{y}(x+z)))$ if a morphism of groups $E \to \mathbb{C}$, trivial if and only if $x + z = 0$. So

$$\sum_{y\in E}\psi(\mathrm{Tr}(\overline{y}(x+z))) = \begin{cases} 0 & \text{if } x+z \neq 0 \\ q^2 & \text{if } x+z = 0. \end{cases}$$

Finally, we get $F(x) = f(-x)$.

(4). Suppose that A contains an element x that is not diagonalizable (over $\overline{\mathbb{F}}_q$). Then, after replacing A by a conjugate, we may assume that $x = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}$, with $a \in \mathbb{F}_q^\times$. Then, for every $r \in \mathbb{Z}$,

$$x^r = \begin{pmatrix} a^r & ra^{r-1} \\ 0 & a^r \end{pmatrix},$$

so $x^{q-1} = \begin{pmatrix} 1 & -a^{q-2} \\ 0 & 1 \end{pmatrix} \in A - \{1\}$ is idempotent, which contradicts the hypothesis. So every element of A is diagonalizable.

If every element of A has eigenvalues in $\mathbb{F}_q$, then there exists $g \in G$ such that $A \subset gTg^{-1}$. By maximality of A, $A = gTg^{-1}$.

Otherwise, there exists $h \in GL_2(\mathbb{F}_{q^2})$ such that, for every $x \in A$, $h^{-1}xh$ is diagonal in $GL_2(\mathbb{F}_{q^2})$. Fix a square root α of u in $\mathbb{F}_{q^2}$. Let $x \in A$. As $x \in GL_2(\mathbb{F}_q)$, the diagonal entries of $h^{-1}xh$ are of the form $a_x + \alpha b_x$ and $a_x - \alpha b_x$. Let S' be the set of elements of $GL_2(\mathbb{F}_{q^2})$ that are diagonal with diagonal entries equal to $a+\alpha b$ and $a-\alpha b$, with $a, b \in \mathbb{F}_q$; note that $|S'| \leq q^2 - 1$. Then $h^{-1}Ah \subset S'$. Similarly, there exists $h' \in GL_2(\mathbb{F}_{q^2})$ such that $h'Sh'^{-1} \subset S'$. As $|S| = q^2 - 1 \geq |S'|$, we have $S' = h'Sh'^{-1}$, so $A \subset (hh')S(hh')^{-1}$. We are not done because we don't know that $hh' \in GL_2(\mathbb{F}_q)$. If we could find $g \in G$ such that $A \subset gSg^{-1}$, then we could conclude by maximality of A that $A = gSg^{-1}$. Note that A is isomorphic to a subgroup of $S \simeq \mathbb{F}_{q^2}^\times$, so A is cyclic. Let x be a generator of A. We just need to find $g \in G$ such that $g^{-1}xg \in S$, and then we'll have $g^{-1}Ag \subset S$. But x has eigenvalues of the form $a + \alpha b$ and $a - \alpha b$, with $a, b \in \mathbb{F}_q$, so it cnojugate in $GL_2(\mathbb{F}_q)$ to the element $\begin{pmatrix} a & b \\ ub & a \end{pmatrix}$ of S, and we are done.

□

VII.2.15 Representations of $\mathrm{GL}_2(\mathbb{F}_q)$, part 5

This is a continuation of problems VII.2.13 and VII.2.14, and we use the notation of these problems. We also suppose that q is odd, and we fix a nontrivial group morphism $\psi : \mathbb{F}_q \to \mathbb{C}^\times$.

(1). Let $\chi : S \to \mathbb{C}$ be a morphism of groups. We let

$$\mathscr{W}(\chi) = \{f : E \to \mathbb{C} | f(yx) = \chi(y)^{-1} f(x)\ \forall x, y \in E \text{ satisfying } N(y) = 1\}.$$

Calculate $\dim_{\mathbb{C}} \mathscr{W}(\chi)$.

(2). For every $a \in \mathbb{F}_q^\times$ and $c \in \mathbb{F}_q$, let

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

and

$$n(c) = \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}.$$

Let $H = \mathrm{SL}_2(\mathbb{F}_q) \subset G$, and let H' be the group generated by elements t_a, $a \in \mathbb{F}_q^\times$, n_c, $c \in \mathbb{F}_q$, and s, subject to the following relations :

$$t_{a_1} t_{a_2} = t_{a_1 a_2}, \quad n_{c_1} n_{c_2} = n_{c_1 + c_2},$$

$$t_a n_c t_a^{-1} = n_{a^2 c},$$
$$s t_a s = t_{-a^{-1}},$$

and, if $c \neq 0$,

$$s n_c s = t_{-c^{-1}} n_{-c} s n_{-c^{-1}}.$$

Show that there is a morphism $\varphi : H' \to H$ that sends t_a to $t(a)$, n_c to $n(c)$ and s to w, and that it is an isomorphism.

(Hint : Write $B_H = B \cap H$, $N_H = N \cap H$, $T = T_H$. When you're trying to construct an inverse of φ, show first that $H = B_H \sqcup B_H w B_H$, and that $B_H = T_H N_H$ and $B_H w B_H = N_H T_H w N_H$.)

(3). Show that there exists a unique representation ρ of H on $\mathscr{W}(\chi)$ such that, for every $f \in \mathscr{W}(\chi)$ and $x \in E$:

- If $a \in \mathbb{F}_q^\times$,
$$\left(\rho \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} f\right)(x) = f(ax).$$
- If $c \in \mathbb{F}_q$,
$$\left(\rho \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix} f\right)(x) = \psi(N(x)c) f(x).$$
- $(\rho(w)f)(x) = \widehat{f}(x)$.

(Remember that $\widehat{f}$ is defined in problem VII.2.14(3).)

This is (a particular case of) the *Weil representation*.

About checking the last relation : No, you don't need to know how to calculate Gauss sums. It's easier than it seems. Look more closely at that sum, and remember the properties of $N : E \to \mathbb{F}_q$ that you (hopefully) proved in (1). (For example, that this map is surjective.)

(4). Show that there is a unique extension of ρ to a morphism of groups $\rho : G \to GL(\mathscr{W}(\chi))$ (ie a representation of G on $\mathscr{W}(\chi)$) such that, for every $a \in \mathbb{F}_q^\times$, every $f \in \mathscr{W}(\chi)$ and $x \in E$,

$$\left(\rho \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} f\right)(x) = \chi(b) f(bx),$$

where b is any element of S such that $N(b) = a$.

(5). Suppose that χ is not trivial on the subgroup $\mathrm{Ker}(N)$ of S. Show that $\mathscr{W}(\chi)$ is cuspidal and irreducible.

(6). Show that every cuspidal irreducible representation of G is of the form $\mathscr{W}(\chi)$, for χ satisfying the condition of (5).

(Hint : Calculate the character of $\mathscr{W}(\chi)$ on $\begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix} \in G$.)

Solution.

(1). Let $E^1 = \mathrm{Ker}(N : E^\times \to \mathbb{F}_q^\times)$ be the set of norm 1 elements of E. Let's calculate the cardinality of E^1 for later usage. Things will be easier if we write an isomorphism $\mathbb{F}_{q^2} \xrightarrow{\sim} E$ more explicitely (see problem VII.2.14(1) for the existence of such an isomorphism). Let $\alpha \in \mathbb{F}_{q^2}$ be a square root of u. Then $\mathbb{F}_{q^2} = \mathbb{F}_q[\alpha]$, so every element of E can be written in a unique way as $a + \alpha b$, with $a, b \in \mathbb{F}_q$. We define a map $\iota : \mathbb{F}_{q^2} \to E$ by $\iota(a + \alpha b) = \begin{pmatrix} a & b \\ ub & a \end{pmatrix}$. Then this is obviously a bijection that preserves 0, 1 and addition. It's also compatible with multiplication by the explicit calculation of the product of two elements of E in the solution of (1) of problem VII.2.14. Now note that $\alpha^{2q} = u^q = u$, so α^q is also a square root of u; as $\alpha^q \neq \alpha$ (otherwise α would be in $\mathbb{F}_q$), we have $\alpha^q = -\alpha$. So, for every $a + \alpha b \in \mathbb{F}_{q^2}$,

$$(a + \alpha b)^q = a^q + \alpha^q b^q = a - \alpha b.$$

In other words, for every $g \in E$, $\overline{g} = g^q$ and $N(g) = g^{q+1}$. As E is a field, this means that E^1, the set of zeros of the polynomial $x^{q+1} - 1$ in E, has at most $q + 1$ elements. If we know about separable polynomials, it's obvious that the polynomial $x^{q+1} - 1$ is separable and hence $|E^1| = q + 1$. Otherwise, we observe that $|E^1| = |E^\times||N(E^\times)|^{-1} \geq (q^2 - 1)/(q - 1) = q + 1$, and so $|E^1| = q + 1$. (Also, we proved that $N : E^\times \to \mathbb{F}_q^\times$ is surjective.)

An element $f \in \mathscr{W}(\chi)$ is totally determined by its values on a set of representatives of E/E^1. Let $(x_i)_{i\in I}$ be a set of representatives of $E^\times/E^1$; then $(0, x_i, i \in I)$ is a set of representatives of E/E^1. For every $i \in I$, there exists a unique function $f_i \in \mathscr{W}(\chi)$ such that $f_i(e_i) = 1$ and $f_i(e_j) = 0$ if $j \neq i$; it's given by

$$f_i(x) = \begin{cases} \chi(y)^{-1} & \text{if } x = ye_i,\ y \in E^1 \\ 0 & \text{otherwise.} \end{cases}$$

If $\chi_{|E^1} = 1$, then the function f_0 that sends 0 to 1 and every element of $E^\times$ to 0 is also an element of $\mathscr{W}(\chi)$. If $\chi_{|E^1} \neq 1$, then, for every $f \in \mathscr{W}(\chi)$, $f(0) = 0$ (indeed, choose $y \in E^1 - \{1\}$ such that $\chi(y) \neq 1$, and note that $f(0) = f(y0) = \chi(y)^{-1} f(0)$).

Finally, we see that if $\chi_{|E^1} = 1$, then the family $(f_0, f_i, i \in I)$ is a basis of $\mathscr{W}(\chi)$, so

$$\dim_{\mathbb{C}} \mathscr{W}(\chi) = 1 + |E^\times/E^1| = 1 + (q^2 - 1)/(q + 1) = q.$$

If $\chi_{|E^1} \neq 1$, then the family $(f_i)_{i\in I}$ is a basis of $\mathscr{W}(\chi)$, so

$$\dim_{\mathbb{C}} \mathscr{W}(\chi) = |E^\times/E^1| = q - 1.$$

(2). To show that φ exists, we just have that the images of the generators satisfy the relations in H, which is an easy calculation.

To show that φ is an isomorphism, we construct its inverse ψ. Define $\psi : H \to H'$ by

$$\psi\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{cases} n(ac^{-1})t(-c^{-1})sn(dc^{-1}) & \text{if } c \neq 0 \\ t(a)n(ba^{-1}) & \text{if } c = 0. \end{cases}$$

We have $\psi \circ \varphi = \mathrm{id}_{H'}$, because this is true on the generators on H'. And an easy calculation shows that $\varphi \circ \psi = \mathrm{id}_H$. So φ is bijective. As it is a morphism of groups by construction, it's an isomorphism of groups.

(3). By (2), we can instead check that there is a unique morphism of groups $\rho' : H' \to \mathrm{GL}(\mathscr{W}(\chi))$ (which will be $\rho \circ \varphi$) that is given by the formulas above on the genetaors t_a, n_c and s. That is, we just have to check that the images of these generators in $\mathscr{W}(\chi)$ satisfy the relations defining H'.

Let's write $T_a = \rho'(t_a)$, $N_c = \rho'(N_c)$, $S = \rho'(s)$. Let $f \in \mathscr{W}(\chi)$ and $x \in E$. The, if $a_1, a_2 \in \mathbb{F}_q^\times$ and $c_1, c_2 \in \mathbb{F}_q$,

$$(T_{a_1}T_{a_2}f)(x) = (T_{a_2}f)(a_1x) = f(a_1a_2x) = (T_{a_1a_2}f)(x)$$

and

$$(N_{c_1}N_{c_2}f)(x) = \psi(N(x)c_1)(N_{c_2}f)(x) = \psi(N(x)c_1)\psi(N(x)c_2)f(x) = (N_{c_1+c_2}f)(x),$$

which gives the first two relations.

For every $a \in \mathbb{F}_q^\times$ and $c \in \mathbb{F}_q$, noting by $T_a^{-1} = T_{a^{-1}}$ by the first relation and that $N(ax) = a^2x$ by definition of N, we have

$$(T_aN_cT_a^{-1}f)(x) = (N_cT_{a^{-1}}f)(ax) = \psi(N(ax)c)(T_{a^{-1}}f)(ax) = \psi(N(x)a^2c)f(x),$$

which is equal to $(N_{a^2c}f)(x)$. This gives the third relation.

For every $a \in \mathbb{F}_q^\times$, we have

$$\widehat{f}(ax) = -q^{-1}\sum_{y\in E} f(y)\psi(\mathrm{Tr}(\overline{ax}y)) = -q^{-1}\sum_{y'\in E} f(a^{-1}y')\psi(\mathrm{Tr}(\overline{x}y') = (\widehat{T_{a^{-1}}f})(x)$$

(using the fact that $\overline{ax} = a\overline{x}$ and the change of variables $y' = ay$). Also, problem VII.2.14(3) implies that $\widehat{\widehat{h}} = T_{-1}h$ for every$h \in \mathscr{W}(\chi)$. So

$$ST_aSf = ST_a\widehat{f} = S\widehat{T_{a^{-1}f}} = \widehat{\widehat{T_{a^{-1}}f}} = T_{-a^{-1}}f,$$

and this gives the fourth relation.

Finally, let $c \in \mathbb{F}_q^\times$. Then

$$(SN_{-c^{-1}}f)(x) = -q^{-1}\sum_{y\in E}\psi(-c^{-1}N(y))f(y)\psi(\mathrm{Tr}(\overline{x}y)),$$

so

$$\begin{aligned}(T_{-c^{-1}}N_{-c}SN_{-c^{-1}}f)(x) &= \psi(-cN(-c^{-1}x))(SN_{-c^{-1}}f)(-c^{-1}x)\\ &= -q^{-1}\psi(-c^{-1}N(x))\sum_{y\in E}\psi(-c^{-1}N(y)-c^{-1}\mathrm{Tr}(\overline{x}y))f(y)\\ &- q^{-1}\sum_{y\in E}\psi(c^{-1}(N(x)+N(y)+\mathrm{Tr}(\overline{x}y))f(y)\\ &= q^{-1}\sum_{y\in E}\psi(c^{-1}N(x+y))f(y).\end{aligned}$$

On the other hand,

$$(N_cSf)(x) = \psi(cN(x))\widehat{f}(x),$$

so

$$\begin{aligned}(SN_cSf)(x) &= -q^{-1}\sum_{y\in E}\psi(cN(y))\psi(\mathrm{Tr}(\overline{x}y))\widehat{f}(y)\\ &= q^{-2}\sum_{y,z\in E}\psi(cN(y)+\mathrm{Tr}(\overline{x}y)+\mathrm{Tr}(\overline{y}z))f(z).\end{aligned}$$

Noting that

$$cN(y)+\mathrm{Tr}(\overline{x}y)+\mathrm{Tr}(\overline{y}z) = cN(y)+\mathrm{Tr}(y(\overline{x+z})) = cN(y+c^{-1}(x+z))-c^{-1}N(x+z),$$

we get

$$(SN_cSf)(x) = q^{-2}\sum_{z\in E}\psi(-c^{-1}N(x+z))f(z)\sum_{y\in E}\psi(cN(y+c^{-1}(x+z))).$$

For every $t\in E$, let

$$\Sigma(c,t) = \sum_{y\in E}\psi(cN(y+c^{-1}t)).$$

Then doing the change of variables $y' = y+c^{-1}t$, we see that $\Sigma(c,t) = \Sigma(c,0)$. Also, we have shown in (1) that $N : E^\times \to \mathbb{F}_q^\times$ is surjective, so there exists $d\in E^\times$ such that $N(d) = c$, and we have

$$\Sigma(c,0) = \sum_{y\in E}\psi(cN(y)) = \sum_{y\in E}\psi(N(dy)) = \Sigma(1,0).$$

Finalement, let's calculate $\Sigma := \Sigma(1,0)$. Remember that we showed in (1) that $N : E^\times \to \mathbb{F}_q^\times$ is surjective and that its kernel E^1 has cardinality $q+1$. So we get

$$\Sigma = \sum_{y\in E}\psi(N(y)) = \psi(0)+|E^1|\sum_{a\in\mathbb{F}_q^\times}\psi(a) = 1+(q+1)(-1) = -q,$$

because, as $\psi : \mathbb{F}_q \to \mathbb{C}^\times$ is a nontrivial character,

$$\sum_{a\in\mathbb{F}_q^\times}\psi(a) = -1+\sum_{b\in\mathbb{F}_q}\psi(b) = -1.$$

Coming back to the calculation of SN_cSf, we get

$$(SN_cSf)(x) = q^{-2}\sum_{z\in E}\psi(-c^{-1}N(x+z))f(z)(-q) = q^{-1}\sum_{z\in E}\psi(c^{-1}N(x+z))f(z),$$

which is indeed equal to $(T_{-c^{-1}}N_{-c}SN_{-c^{-1}}f)(x)$, and so the fifth relation is also proved.

(4). First, note that if $f \in \mathscr{W}(\chi)$, $x \in E$ and $b, b' \in E$ are two elements such that $N(b) = N(b') \in \mathbb{F}_q^{-1}$, then $b'b^{-1} \in E^1$,

$$\chi(b')f(b'x) = \chi(b')f((b'b^{-1})bx) = \chi(b')\chi(b'b^{-1})^{-1}f(bx) = \chi(b)f(bx).$$

So the formula given in the statement of (4) does not depend on the choice of b.

For every $a \in \mathbb{F}_q^\times$, write $t'(a) = \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}$. Then, if $g \in G$, $t'(\det(g)^{-1})g \in H$, so G is generated by H and the $t'(a)$, $a \in \mathbb{F}_q^\times$, which gives the uniqueness of the extension of ρ to G.

If $a \in \mathbb{F}_q^\times$, let T'_a be the linear automorphism of $\mathscr{W}(\chi)$ defined in the statement of (4). We define $\rho : G \to \mathscr{W}(\chi)$ by taking $\rho(g) = T'_{\det(g)}\rho(t'(\det(g)^{-1})g)$. (This is the only possible choice.) We have to check that this is a morphism of groups. If $g_1, g_2 \in G$, and if we set $a_i = \det(g_i)$ and $h_i = t'(a_i)^{-1}g_i$, then

$$g_1g_2 = t'(a_1)h_1t'(a_2)h_2 = t'(a_1a_2)(t'(a_2)^{-1}h_1t'(a_2))h_2$$

and $t'(a_2)^{-1}h_1t'(a_2)$. As we know that $\rho_{|H}$ is a morphism of groups, we need to check two things :

$$T'_{a_1}T'_{a_2} = T'_{a_1a_2}$$

and

$$\rho(t'(a_2)^{-1}h_1t'(a_2)) = T'_{a_2^{-1}}\rho(h_1)T'_{a_2}.$$

For the first equality, choose $b_1, b_2 \in E$ such that $N(b_1) = a_1$, $N(b_2) = a_2$. Let $f \in \mathscr{W}(\chi)$ and $x \in E$. Then

$$(T'_{a_1}T'_{a_2}f)(x) = \chi(b_1)(T'_{a_2}f)(b_1x) = \chi(b_1b_2)f(b_1b_2x) = (T'_{a_1a_2}f)(x).$$

For the second equality, take $a \in \mathbb{F}_q^\times$ and $h \in H$. We want to show that $T'_{a^{-1}}\rho(h)T'_a = \rho(t'(a^{-1}))ht'(a))$. As $T'_{a^{-1}} = T'^{-1}_a$, it's enough to check this equality for h one of the generators of H given in (b). Fix $b \in E$ such that $N(b) = a$, and fix $f \in \mathscr{W}(\chi)$ and $x \in E$.

Suppose that $h = t(a_1)$, with $a_1 \in \mathbb{F}_q^\times$. Then $t'(a)^{-1}ht'(a) = h$, and

$$(T'_{a^{-1}}\rho(h)T'_af)(x) = \chi(b)^{-1}(T_{a_1}T'_af)(b^{-1}x) = \chi(b^{-1})(T'_af)(a_1b^{-1}x) = f(a_1x),$$

so $T'_{a^{-1}}T_{a_1}T'_a f = T_{a_1} f$.

Suppose that $h = n(c)$, with $c \in \mathbb{F}_q$. Then $t'(a^{-1})ht'(a) = n(a^{-1}c)$. On the other hand, $(T'_{a^{-1}}N_cT'_a f)(x)$ is equal to

$$\chi(b)^{-1}(N_cT'_a f)(b^{-1}x) = \chi(b)^{-1}\psi(cN(b^{-1}x))(T'_a f)(b^{-1}x) = \psi(ca^{-1}N(x))f(x),$$

so we do get $T'_{a^{-1}}N_cT'_a f = N_{a^{-1}c}f$.

Finally, suppose that $h = w$. Then

$$t'(a^{-1})ht'(a) = \begin{pmatrix} 0 & a^{-1} \\ -a & 0 \end{pmatrix} = t(a^{-1})h.$$

We need to calculate the Fourier transform of the function $x \longmapsto f(bx)$. It sends x to

$$-q^{-1}\sum_{y\in E} f(by)\psi(\mathrm{Tr}(\overline{x}y)) = -q^{-1}\sum_{y'\in E} f(y')\psi(\mathrm{Tr}(\overline{x}b^{-1}y')) = \widehat{f}(\overline{b}^{-1}x).$$

So $(ST'_a f)(x) = \chi(b)\widehat{f}(\overline{b}^{-1}x)$, and

$$(T'_{a^{-1}}ST'_a f)(x) = \widehat{f}(b^{-1}\overline{b}^{-1}x) = \widehat{f}(a^{-1}x) = (T_{a^{-1}}Sf)(x).$$

This finishes the proof.

(5). First we calculate $\mathscr{W}(\chi)^N$. Let $f \in \mathscr{W}(\chi)$ such that $N(c)f = f$ for every $c \in \mathbb{F}_q$. Then, if $x \in E$, $\psi(cN(x))f(x) = f(x)$ for every $c \in \mathbb{F}_q$. This implies that $f(x) = 0$ if $x \neq 0$. As $\psi_{|E^1} \neq 1$, we know that $f(0) = 0$ (see the proof of (a), but this is pretty easy). So $f = 0$. So we have shown that $\mathscr{W}(\chi)^N = 0$.

This implies that every irreducible component V of $\mathscr{W}(\chi)$ also satisfies V^N, ie is irreducible cuspidal. But we have seen in problem VII.2.13(6) that every irreducible cuspidal representation of G has dimension $q-1$, and we have seen in (a) that $\mathscr{W}(\chi)$ has dimension $q-1$ (because $\chi_{|E^1} \neq 1$), so $\mathscr{W}(\chi)$ cannot have more than one irreducible component, which means that it is irreducible cuspidal.

(6). We have seen in problem VII.2.13(5) that there are $q(q-1)/2$ classes of irreducible cuspidal representations of G. Let's count the number N of morphisms of groups $\chi : E^\times \to \mathbb{C}^\times$ that are trivial on E^1. As $E^\times/E^1 \xrightarrow{\sim} \mathbb{F}_q^\times$ is commutative of order $q-1$ (see (1)), the number of χ that are trivial on E^1 is equal to the number of group morphisms $\mathbb{F}_q^\times \to \mathbb{C}^\times$, which is $q-1$. Also, as $E^\times$ is commutative, the number of morphism of groups $E^\times \to \mathbb{C}^\times$ is $|E^\times| = q^2 - 1$. So, finally, $N = q^2 - 1 - (q-1) = q(q-1)$.

Obviously, some of the $\mathscr{W}(\chi)$ are going to be isomorphic. The easiest way to see whether two representations are isomorphic is to compare their characters. We don't want to do the full calculation, but we're basically forced to.

Let's first calculate the value of $\chi_{\mathscr{W}(\chi)}$ on the matrix $g := \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix} \in G$. This matrix is equal to $t'(a^2)t(a^{-1})n(a^{-1})$, so, if $f \in \mathscr{W}(\chi)$ and $x \in E$,

$$(\rho(g)f)(x) = \chi(a)\psi(a^{-1}N(x))f(x).$$

Now assume that $\chi_{|E^1} \neq 1$, and remember the basis $(f_i)_{i\in I}$ constructed in the proof of (a). As $N : E^\times \to \mathbb{F}_q^\times$ is constant on the orbits of E^1, the calculation above show that $\rho(g)f_i = \chi(a)\psi(a^{-1}N(x_i))$. So we get

$$\chi_{\mathscr{W}(\chi)}(g) = \sum_{i\in I} \chi(a)\psi(a^{-1}N(x_i)) = \chi(a) \sum_{x\in E/E^1} \psi(a^{-1}x) = \chi(a) \sum_{x\in\mathbb{F}_q^\times} \psi(a^{-1}x).$$

As χ is non-trivial,

$$\sum_{x\in\mathbb{F}_q^\times} \psi(a^{-1}x) = \sum_{x'\in\mathbb{F}_q^\times} \psi(x') = -1 + \sum_{x'\in\mathbb{F}_q} \psi(x') = -1.$$

So finally

$$\chi_{\mathscr{W}(\chi)}(g) = -\chi(a).$$

Let's also calculate $\chi_{\mathscr{W}(\chi)}$ on the matrix $g = \begin{pmatrix} a & b \\ ub & a \end{pmatrix}$, with $a \in \mathbb{F}_q$, $b \in \mathbb{F}_q^\times$ and $a^2 - ub^2 \neq 0$, using the proofs of (b) and (d) to decompose this as a product of matrices whose images by ρ we know. It's an awful calculation, and should give

$$\chi_{\mathscr{W}(\chi)}(g) = -(\chi(z) + \chi(\overline{z})) = -(\chi(z) + \chi(z^q)),$$

where $z = a + \alpha b \in E$.

As $E \simeq \mathbb{F}_{q^2}$, the group $E^\times$ is cyclic. Let's choose a generator ζ of $E^\times$. Now let χ, χ' be two morphisms of groups $E^\times \to \mathbb{C}^\times$ that are non-trivial on E^1 and such that $\mathscr{W}(\chi) \simeq \mathscr{W}(\chi')$. Then $\chi_{\mathscr{W}(\chi)} = \chi_{\mathscr{W}(\chi')}$, and by the calculation above $\chi(z) + \chi(z^{q+1}) = \chi'(z) + \chi'(z^q)$ for every $z \in E^\times$. Writing $\xi = \chi(\zeta)$ and $\xi' = \chi'(\zeta)$, this translates to

$$\xi^n + (\xi^q)^n = \xi'^n + (\xi'^q)^n$$

for every $n \in \mathbb{Z}$, and is equivalent to the condition $\{\xi, \xi^q\} = \{\xi', \xi'^q\}$. Also, the condition $\chi_{|E^1} \neq 1$ (resp. $\chi'_{|E^1} \neq 1$) is equivalent $\xi^{q-1} \neq 1$, ie $\xi \neq \xi^q$ (resp. $\xi' \neq \xi'^q$), because E^1 is the subgroup $\{x \in E^\times | x^{q+1} = 1\}$ of $E^\times$, hence is generated by ζ^{q-1}. Finally, we see that, if $\mathscr{W}(\chi) \simeq \mathscr{W}(\chi')$, then either $\xi = \xi'$, which means that $\chi = \chi'$, or $\xi' = \xi^q$, which means that $\chi' = \chi^q$. This implies that there are at least $N/2 = q(q-1)/2$ distinct $\mathscr{W}(\chi)$. As the total number of irreducible cuspidal representations of G is $q(q-1)/2$, there are exactly $q(q-1)/2$ distinct $\mathscr{W}(\chi)$, and every irreducible cuspidal representation is of that form.

□

VII.2.16 Induction and characters

Let k be a field of characteristic 0, and let $\alpha : H \to G$ be a morphism of finite groups. For every $f \in \mathbb{C}(H, k)$, define a function $\mathrm{Ind}_H^G f : G \to k$ by

$$f(g) = \frac{1}{|H|} \sum_{(s,h)\in G\times H | s^{-1}gs=\alpha(h)} f(h).$$

Show that, for every representation V on a finite-dimensional k-vector space, we have

$$\chi_{\mathrm{Ind}_H^G V} = \mathrm{Ind}_H^G \chi_V.$$

Solution. We may assume that k is algebraically closed.

Write $f = \chi_V$. First note that, if α is injective, the definition of Ind_H^G is the same as in definition II.3.1.1 of chapter II. Also, for every $g \in \alpha(H)$,

$$(\mathrm{Ind}_H^{\alpha(H)} f)(g) = \frac{1}{|H|} \sum_{(s,h)\in\alpha(H)\times H | s^{-1}gs=\alpha(h)} f(h) = \frac{|\alpha(H)|}{|H|} \sum_{h\in H | g=\alpha(h)} f(h),$$

so, for every $g \in G$,

$$(\mathrm{Ind}_{\alpha(H)}^G \mathrm{Ind}_H^{\alpha(H)} f)(g) = \frac{1}{|H|} \sum_{s\in G | s^{-1}gs\in\alpha(H)} \frac{|\alpha(H)|}{|H|} \sum_{h\in H | s^{-1}gs=\alpha(h)} f(h) = (\mathrm{Ind}_H^G f)(g).$$

Using this, the transitivity of induction and theorem II.3.1.2 of chapter II, we are reduced to the case where $G = \alpha(H)$.

In this case, by lemma I.5.2.5 and proposition I.5.5.3 of chapter I, $\mathrm{Ind}_H^G V = V^{\mathrm{Ker}(\alpha)}$. Now we have to generalize lemmas II.1.2.3 and II.1.2.4 of chapter II. Write $K = \mathrm{Ker}(\alpha)$. We want to show that, for every $g \in G$,

$$\chi_{V^K}(g) = \frac{1}{|K|} \sum_{h\in\alpha^{-1}(g)} \chi_V(h).$$

Let $V = \bigoplus_{W\in S_k(K)} W^{\oplus n_W}$ be the decomposition into irreducibles of $\mathrm{Res}_K^H V$. Fix $g \in G$, pick $h_0 \in H$ such that $g = \alpha(h_0)$ and write $c = \sum_{x\in K} h_0 x = \sum_{h\in\alpha^{-1}(g)} h \in k[H]$. Then, for every $x \in K$, $xc = cx = x$. In particular, c centralizes $k[K] \subset k[H]$, and so, by Schur's lemma, it stabilizes every summand $W^{\oplus n_W}$ in V and acts on $W^{\oplus n_W}$ by a $n_W \times n_W$ matrix A_W with coefficients in $\mathrm{End}_{k[K]}(W) = k$.

Fix $W \in S_k(K)$, write ρ_W for the map $K \to \mathrm{End}_k(W)$ and $n = n_W$. Let $\lambda_1, \ldots, \lambda_n$ be the diagonal coefficients of A_W. Then, using the fact that $cx = c$ for every $x \in K$, we see as in the proof of lemma I.1.2.3 of chapter I that, for every $i \in \{1, \ldots, n\}$ and every $x \in K$,

$$\lambda_i \rho_W(x) = \lambda_i \mathrm{id}_W.$$

So, if $W \not\simeq \mathbb{1}_K$, we must have $\lambda_1 = \cdots = \lambda_n = 0$.

Also, if $W = \mathbb{1}_K$, then the action of c on the summand $W^{\oplus n_W} = V^K$ is equal to that of $|K|g$.

Finally, this shows that

$$\chi_V(c) = \sum_{h \in \alpha^{-1}(g)} \chi_V(h) = |K| \chi_{V^K}(g),$$

which is the result we wanted to prove.

□

VII.3 Chapter III exercises

VII.3.1 Discrete valuation rings

A *discrete valuation ring* is a commutative principal ideal domain A such that A has a unique nonzero prime ideal $\wp$. Let π be a generator of $\wp$; we call π a *uniformizer* of A. The quotient $k = A/\wp$ is called the *residue field* of A. We denote by K the fraction field of A.

(1). Show that k is indeed a field.

(2). Show that, for every $x \in K^\times$, there exists $n \in \mathbb{Z}$ and $u \in A^\times$ uniquely determined such that $x = u\pi^n$, and that n does not depend on the choice of the uniformizer π.

If $x = u\pi^n$ as in (2), we write $n = v(x)$ and we say that n is the *valuation* of x. We also set $v(0) = \infty$.

(3). Show that :

(a) $v : K^\times \to \mathbb{Z}$ is surjective.

(b) For every $x, y \in K$, $v(xy) = v(x) + v(y)$ and $v(x + y) \geq \inf(v(x), v(y))$.

(c) $A = \{x \in K | v(x) \geq 0\}$ and $\wp = \{x \in K | v(x) \geq 1\}$.

(With the convention that, for every $n \in \mathbb{Z}$, $\infty n = \infty$ and $\infty > n$.)

(4). Let k be a field and $k[[T]]$ be the ring of formal series over k. Show that $k[[T]]$ is a discrete valuation ring.

(5). Show that $\mathbb{Z}_{(p)}$ is a discrete valuation ring. (Remember that $\mathbb{Z}_{(p)}$ is the localization of $\mathbb{Z}$ at the prime ideal (p), that is, the subring of $\mathbb{Q}$ generated by $\mathbb{Z}$ and the $1/n$, for every $n \in \mathbb{Z} - (p)$.)

Solution.

(1). As A is a domain, 0 is also a prime ideal of A. So A has exactly two prime ideals and cannot be a field. As maximal ideals are prime, the only maximal ideal of A is $\wp$, and so $k = A/\wp$ is a field.

(2). As A is a principal ideal domain, it is a unique factorization domain. The only irreducible element of A is π (up to multiplying by an invertible element), so every element a of $A - \{0\}$ can be written as $a = u\pi^n$, with $u \in A^\times$ and $n \geq 0$ uniquely determined by a. Also, n is the unique nonnegative integer such that $a \in \wp^n - \wp^{n+1}$ (in other words, the biggest nonnegative integer such that $a \in \wp^n$), so it does not depend on the choice of $\wp$. Write $n = v(a)$. Clearly, if $a, b \in A - \{0\}$, then $v(ab) = v(a) + v(b)$.

As every element of $K^\times$ is of the form ab^{-1} for $a, b \in A - \{0\}$, and as v sends multiplication in A to addition in $\mathbb{Z}$, these results extend to elements of $K^\times$.

(3). (a) For every $n \in \mathbb{Z}$, $v(\wp^n) = n$. So v is surjective.

(b) The first property is clear. For the second property, we may assume that $x, y \in A$ (using the first property), and that they both nonzero (otherwise the conclusion is obvious). Let $n = v(x)$ and $m = v(y)$. Then $x+y \in \wp^{\inf(n,m)}$, so $v(x+y) \geq \inf(n, m)$. (Here we use the fact that, for $a \in A - \{0\}$, $v(a)$ is the biggest nonnegative integer n such that $a \in \wp^n$.)

(c) Obviously, for every $a \in A$, $v(a) \geq 0$. Now let $x \in K^\times$ such that $v(x) \geq 0$, and write $x = ab^{-1}$, with $a, b \in A - \{0\}$, Let $n = v(a)$, $m = v(b)$. Then $a = u\pi^n$ and $b = v\pi^m$, with $u, v \in A^\times$. So $x = uv^{-1}\pi^{n-m} \in A$. The second equality follows from the characterization of v we gave in (b).

(4). Either you know that $k[[T]]$ is a principal ideal domain, and then it's easy because it's also local with maximal ideal (T), so its unique irreducible element (up to invertibles) is T. Or you don't, and then the easiest way is to use problem VII.3.2, with the valuation v on $\mathrm{Frac}(k[[T]]) = k((T))$ given by taking the order of $f \in k((T))$ at 0.

(5). We know that $\mathbb{Z}[p^{-1}]$ is a principal ideal domain, because it is a localization of the principal ideal domain $\mathbb{Z}$. Prime ideals of $\mathbb{Z}[p^{-1}]$ are prime ideals of $\mathbb{Z}$ that are contained in (p), so only 0 and (p) are left.

□

VII.3.2 Discrete valuation fields

Let K be a field and $v : K^\times \to \mathbb{Z}$ be a surjective group morphism such that $v(x + y) \geq \inf(v(x), v(y))$ for every $x, y \in K^\times$ such that $x + y \neq 0$. (We say that (K, v) (or simply K) is a *discrete valuation field*.) Show that $A := \{0\} \cup \{x \in K^\times | v(x) \geq 0\}$ is a discrete valuation ring (see problem VII.3.1 for the definition of a discrete valuation ring). (The ring A is called is the *valuation ring* of K.)

Solution. First, A is obviously a domain, because it's a subring of a field.

We claim that $A^\times = \{x \in K | v(x) = 0\}$. Indeed, if $a \in A^\times$, then a and a^{-1} are in A, so $v(a) \geq 0$ and $v(a^{-1}) = -v(a) \geq 0$, hence $v(a) = 0$. Conversely, let $a \in K$ such that $v(a) = 0$. Then $v(a) \geq 0$ and $v(a^{-1}) = -v(a) \geq 0$, so both a and a^{-1} are in A, so $a \in A^\times$.

Choose $\pi \in K$ such that $v(\pi) = 1$. We have $\pi \in A$ by definition of A. Let's show that, for every $a \in A - \{0\}$, there exists $u \in A^\times$ and $n \geq 0$, uniquely determined by a, such that $a = u\pi^n$. First, if $a = u\pi^n$, then $v(a) = v(u) + nv(\pi) = n$, and $u = a\pi^{-n}$, so u and n are determined by a. Also, for every $a \in A - \{0\}$, we have $n := v(a) \geq 0$ and $u := a\pi^{-n} \in A^\times$ (because $v(u) = 0$), which gives the existence of u and n.

Now we show that A is a principal ideal domain whose only nonzero prime ideal if (π). Let I be a nonzero ideal of A. If a is any nonzero element of I and $n = v(a)$, then $a = u\pi^n$ with $u \in A^\times$, so $\pi^n \; in I$. Let n_0 be the smallest nonnegative integer such that $\pi^n \in I$. (This exists because every nonempty subset of $\mathbb{Z}_{\geq 0}$ has a smallest element.) Let's show that $I = (\pi^{n_0})$. As $\pi^{n_0} \in I$, we obviously have $I \supset (\pi^{n_0})$. Conversely, let $a \in I$, let $n = v(a)$. Then we see as before that $\pi^n \in I$, so $n \geq n_0$ by definition of n_0, so $a\pi^{-n_0} \in A$ (because its valuation is ≥ 0), so $a \in (\pi^{n_0})$. Finally, we show that I is prime if and only if $n_0 = 1$. if $n_0 > 1$, then I is not prime because $\pi \in I$ and $\pi^{n_0} \in I$. Conversely, assume that $n_0 = 1$, and let $a, b \in A$ such that $ab \in I$. If $ab = 0$, then $a = 0$ or $b = 0$ because A is domain, so $a \in I$ or $b \in I$. If $ab \neq 0$, then $v(ab) = v(a) + v(b) \geq 1$, so $v(a) \geq 1$ or $v(b) \geq 1$m hence $a \in I$ or $b \in I$.

□

VII.3.3 Completion of a discrete valuation ring

Let A be a discrete valuation ring with maximal ideal m. (See problem VII.3.1.) The *completion* of A is

$$\widehat{A} = \varprojlim_n (A/m^n) := \{(x_n) \in \prod_{n \geq 0} A/m^n | \forall n \geq 0, \; x_n = x_{n+1} \mod m^n\}.$$

We define a map $A \to \widehat{A}$ by sending $x \in A$ to the family $(x + m^n)_{n \geq 0} \in \prod_{n \geq 0} A/m^n$.

(1). Show that the map $A \to \widehat{A}$ is injective. (We use it to identify A to a subring of $\widehat{A}$. If $A = \widehat{A}$, we say that A is *complete*.)

(2). Show that $\widehat{A}$ is a discrete valuation ring with maximal ideal $\widehat{m} := m\widehat{A}$, and that the obvious map $A/m^n \to \widehat{A}/\widehat{m}^n$ is an isomorphism for every n. (In particular, $\widehat{A}$ is complete.)

(3). We consider the following three topologies on $\widehat{A}$:

(a) The topology induced by the product topology on $\prod_{n \geq 0} A/m^n$, where we put the discrete topology on each A/m^n. (This is the topology on $\prod_{n \geq 0} A/m^n$ generated by

the open sets $\prod_{n\geq 0} X_n$, where $X_n \subset A/m^n$ for every n and $X_n = A/m^n$ for all but a finite number of n's.)

(b) The topology generated by the open sets $x + \widehat{m}^n$, for $x \in \widehat{A}$ and $n \geq 0$.

(c) The topology given by the distance function $d(x,y) = c^{v(x-y)}$, where c is a real number such that $0 < c < 1$ and $v : \widehat{A} \to \mathbb{Z}_{\geq 0} \cup \{\infty\}$ is the valuation of $\widehat{A}$. (With the convention that $c^\infty = 0$.)

Show that these three topologies are equal (in particular, the last one does not depend on c), that $\widehat{A}$ is a complete metric space and that, if k is finite, then $\widehat{A}$ is Hausdorff and compact.

(4). Let $k = A/m = \widehat{A}/\widehat{m}$ be the residual field of $\widehat{A}$. Let $f \in A[t]$ be a polynomial such that its image $\overline{f}$ in $k[t]$ is nonzero, and let $\overline{x}$ be a *simple* root of $\overline{f}$ in k. Show that there exists a root x of f in $\widehat{A}$ such that $x \mod m = \overline{x}$.

This statement is called *Hensel's lemma*.

(5). Give an example of a non-complete dicrete valuation ring where (4) fails.

Solution.

(1). The kernel of the map $A \to \widehat{A}$ is $\bigcap_{n\geq 0} m^n$, so we have to show that this intersection is 0. Let v be the valuation of A. If $x \in \bigcap_{n\geq 0} m^n$, then $v(x) \geq n$ for every $n \geq 0$, so $v(x) = \infty$, so $x = 0$.

(2). First we extend v to $\widehat{A}$. Let $x = (x_n) \in \widehat{A}$. We choose elements y_n of A lifting the $x_n \in A/m^n$. For every $n, k \geq 0$, we have $y_{n+k} = y_n$ modulo m^n, so $v(y_{n+k}) \geq \inf(v(y_n), n)$ and $v(y_n) \geq \inf(v(y_{n+k}), n)$. If $y_n \in m^n$ for every $n \geq 0$, then $x = 0$, and we take $v(x) = \infty$. Otherwise, let n_0 be the smallest nonnegative integer such that $y_{n_0} \notin m^{n_0}$. Then, for every $n \geq n_0$, we have $n_0 > v(y_{n_0}) \geq \inf(n_0, v(y_n))$, so $v(y_n) \leq v(y_{n_0}) < n_0$, and $v(y_n) \geq \inf(n, v(y_{n_0})) = v(y_{n_0})$. Finally, we get $v(y_n) = v(y_{n_0})$ for every $n \geq n_0$, and we set $v(x) = v(y_{n_0})$. Note that $v(y_n) = v(x)$ for $n > v(x)$. (Indeed, let N be big enough so that $v(y_N) = v(x)$. Then, for $v(x) < n \leq N$, $y_n = y_N \bmod m^n$, so $n > v(y_N) \geq \inf(v(y_n), n)$ and $v(y_n) \leq v(x) < n$, but then $v(y_n) \geq \inf(v(y_N), n) = v(x)$.) Note also that $y_n \in m^n$ (ie $x_n = 0$) for $n \leq v(x)$. (Indeed, let $n \leq v(x)$. Then $v(y_n) \geq \inf(n, v(x)) = n$, so $y_n \in m^n$.)

If $x \in \widehat{A}$ is the image of an element a of A, and we choose (x_n) in $A^{\mathbb{N}}$ representing x (ie $x = (x_n + m^n)$), then $x_n = a$ modulo m^n for every n. If $n > \max(v(x), v(a))$ and n is big enough that $v(x) = v(x_n)$, this implies that $v(a) = v(x_n) = v(x)$. So v does indeed extend the valuation on A.

It is clear from the definition that, for every $x, y \in \widehat{A}$, $v(xy) = v(x) + v(y)$ and $v(x+y) \geq \inf(v(x), v(y))$. (If x and y are representend by families (x_n) and (y_n) in $A^{\mathbb{N}}$, just take n big enough so that $v(x) = v(x_n)$ and $v(y) = v(y_n)$.) Also, the only element with valuation ∞ is 0. In particular, A is a domain (if $x, y \in A - \{0\}$, then

$v(xy) = v(x) + v(y) \neq \infty$, so $xy \neq 0$). Let K be the fraction field of A. Using that v sends products to sums, we can extend v to a morphism of groups $v : K^\times \to \mathbb{Z}$ satisfying the conditions of problem 2 (the surjectivity comes from the fact that A is a dicrete valuation ring, hence $v(A - \{0\}) = \mathbb{Z}_{\geq 0} \subset v(K)$). So K is a discrete valuation field, and to finish we just have to show that $\widehat{A} - \{0\} = v^{-1}(\mathbb{Z}_{\geq 0})$.

Obviously, v sends $\widehat{A} - \{0\}$ to $\mathbb{Z}_{\geq 0}$. Conversely, let $x \in K^\times$ such that $v(x) \geq 0$. We write $x = ab^{-1}$, with $a, b \in \widehat{A}$ and $N := v(a) \geq M := v(b)$. Write $a = (a_n + m^n)$ and $b = (b_n + m^n)$, with $(a_n), (b_n) \in A^{\mathbb{N}}$. We choose the a_n and b_n such that, if $a_n \in m^n$ (resp. $b_n \in m^n$), then $v(a_n) = n$ (resp. $v(b_n) = n$). This does not affect the classes $a_n + m^n$ and $b_n + m^n$. We want to show that b_n divides a_n in A for every $n \geq 0$, which obviously implies that $xy^{-1} \in \widehat{A}$. If $n > N$, then $v(x_n) = N$ and $v(Y_n) = M \leq N$, so $x_n y_n^{-1} \in A$ (as A itself is a discrete valuation ring). If $M < n \leq N$, then $v(b_n) = M$ and $v(a_n) = n$, so $a_n b_n^{-1} \in A$. If $n \leq M$, then $v(a_n) = v(b_n) = n$, so $a_n b_n^{-1} \in A$.

This finishes the proof that $\widehat{A}$ is a discrete valuation ring. Its unique maximal ideal is $I := \{0\} \cup v^{-1}(\mathbb{Z}_{\geq 1})$, let's show that $I = \widehat{m}$. First, it follows easily from the definition of v on $\widehat{A}$ that I is the set of $a = (a_n)$ in $\widehat{A}$ such that $a_1 = 0$ (in A/m). This contains the image of m in $\widehat{A}$, so I contains $\widehat{m}$. Conversely, let $x \in I - \{0\}$, let $n = v(x)$. We choose an element a of valuation a in m. Then $v(xa^{-1}) = n - 1 \geq 0$, so $xa^{-1} \in \widehat{A}$, so $x = a(xa^{-1}) \in \widehat{m}$.

Finally, we have $\widehat{m}^n = \{0\} \cup v^{-1}(\mathbb{Z}_{\geq n})$, so

$$\widehat{m}^n = \{(a = (a_n) \in \widehat{A} | a_i = 0 \text{ in } A/m^i \text{ for } 0 \leq i \leq n\}.$$

So

$$\widehat{A}/\widehat{m}^n = \{(x_0, \ldots, x_n) \in \prod_{i=0}^{n} A/m^i | \forall i \leq n-1,\ x_i + m^i = x_{i+1} + m^i\}.$$

For a family $(x_0, \ldots, x_n)$ as above, x_n determines all the x_i. So the map $A/m^n \to \widehat{A}/\widehat{m}^n$ sending a to the family $(a + m^i)_{0 \leq i \leq n}$ (which is the obvious map) is an isomorphism of rings.

(3). Let's call these three topologies $\mathscr{T}_1$, $\mathscr{T}_2$ and $\mathscr{T}_3$. Let $U := \prod_{n \geq 0} X_n$ be a generating set for $\mathscr{T}_1$ as above, ie $X_n \subset A/m^n$ and $X_n = A/m^n$ for all but a finite number of n's. Choose $N \geq 0$ such that $X_n = A/m^n$ for every $n \geq N$. Then $U + \widehat{m}^N = U$, so U is a union of classes of $\widehat{A}/\widehat{m}^n$, so there exists a family $(x_i)_{i \in I}$ of elements of $\widehat{A}$ such that U is the disjoint union of the $x_i + \widehat{m}^N$. To show that U is open in $\mathscr{T}_2$ and $\mathscr{T}_3$, it suffices to show that a set of the form $x + \widehat{m}^N$ is. For $\mathscr{T}_2$, such a set if open by definition. For $\mathscr{T}_3$, such a set if open because it's the open ball of radius c^{N-1} centered at x. (That is, $y \in x + \widehat{m}^N$ if and only if $v(x - y) \geq N$ if and only if $v(x - y) > N - 1$.)

We alredy showed that every generating open for $\mathscr{T}_2$ if open for $\mathscr{T}_3$. To finish the proof, we have to show that an open ball for $\mathscr{T}_3$ is open for $\mathscr{T}_1$. Let $x \in \widehat{A}$ and $c \in \mathbb{R}_{>0}$, and let U be

then open ball of radius r and center x. Write $x = (x_n)$, $x_n \in A/m^n$. Then $U = x + \widehat{m}^N$ where N is biggest integer such that $c^N < r$, so $U = \prod_{n\geq 0} X_n$, with $X_n = \{x_n\}$ if $n < N$ and $X_n = A/m^n$ if $n \geq N$.

Note that, using the second description of the topology of $\widehat{A}$, we see that the map v on $A - \{0\}$ is locally constant.

We now show that $\widehat{A}$ is complete. Let $(x_n)_{n\geq 0}$ be a Cauchy sequence in $\widehat{A}$. This means that for every $A \in \mathbb{R}$, there exists $N \in \mathbb{Z}_{\geq 0}$ such that $v(x_n - x_p) \geq A$ if $n, p \geq N$. We want to show that (x_n) converges. As a Cauchy sequence converges if and only if some (infinite) subsequence of it converges, we can always replace (x_n) by a subsequence. In particular, we may assume that either $x_n = 0$ for every n, or $x_n \neq 0$ for every n. In the first case, the sequence (x_n) converges to 0. In the second case, using the fact that v is locally constant on $A - \{0\}$, we may assume that all the x_n have the same valuation, say n_0; after diving all the x_n by the same element of A of valuation n_0, we may assume that $n_0 = 0$. By taking another subsequence of necessary, we may also assume that $v(x_n - x_p) \geq n$ for every $p \geq n$. For every n, we choose a family $(x_{i,n})_{i\geq 0}$ in $A^{\mathbb{N}}$ representing x_n. if $p \geq n$, then for $0 \leq i \leq n$, $x_{i,p}$ and $x_{i,p}$ are equal modulo m^n, hence modulo m^i. So, if x is the element of $\widehat{A}$ represented by $(x_{n,n})$, then $x = x_n$ modulo $\widehat{m}^n$ for every n, and the sequence (x_n) converges to x.

If k is finite, then, for every $n \geq 0$, $m^n/m^{n+1} \simeq k$ is finite, so A/m^n is finite. The fact that $\widehat{A}$ is Hausdorff compact follows from Tychonoff's theorem, because the finite discrete sets A/m^n are Hausdorff compact.

(4). First note the following fact : Let $n \geq 1$. If $x \in A$ is such that $f(x) \in m^n$ and $f'(x) \notin m$ (i.e. $f'(x) \in A^\times$), then, setting $h = -\frac{f(x)}{f'(x)}$ and $y = x + h$, we have $f(y) \in m^{2n}$ and $f'(y) \notin m$. Indeed, we have $h \in m^n$, so

$$f(y) = f(x + h) \in f(x) + hf'(x) + h^2A \subset f(x) + hf'(x) + m^{2n} = m^{2n},$$

and

$$f'(y) = f'(x + h) \in f'(x) + hA \subset f'(x) + m^n,$$

so $f'(y) \notin m$ because otherwise $f'(x)$ would be in m.

Now let's prove Hensel's lemma. We construct by induction on $n \geq 0$ a sequence $(x_n)_{n\geq 0}$ of elements of A such that :

- For every $n \geq 0$, $x_n + m = \overline{x}$, $f(x_n) \in m^{2^n}$ and $f'(x_n) \notin m$.
- For every $n \geq 0$, $x_{n+1} - x_n \in m^{2^n}$.

For x_0, we choose any lift of $\overline{x}$ in A. Then $f(x_0) \in m$ because $\overline{x}$ is a root of $\overline{f}$, and $f'(x_0) \notin m$ because it's a simple root. Suppose that we have constructed $x_0, \ldots, x_n$, and let's construct x_{n+1}. Let $h_n = -\frac{f(x_n)}{f'(x_n)} \in m^{2^n}$, and take $x_{n+1} = x_n + h_n$. Then x_{n+1} satisfies all the desired conditions by the observation above.

Now note that $(x_n)_{n\geq 0}$ is a Cauchy sequence if we use the metric defined in question (3). As A is complete for this metric, the sequence $(x_n)_{n\geq 0}$ has a limit, say x. We have $x + m = \overline{x}$ (if we choose some n such that $v(x - x_n) \geq 1$, then $x \in x_n + m$, so $x = x_n \mod m$), and $f(x)$ is the limit of the sequence $(f(x_n))_{n\geq 0}$; as $f(x_n) \in m^{2n}$ for every n, $f(x_n) \to 0$.

(5). Take $A = \mathbb{Z}[p^{-1}]$, this is a discrete valuation ring with residue field $\mathbb{Z}/p\mathbb{Z}$. If $p = 5$, then the polynomial $f = t^2+1$ has a simple root in $\mathbb{Z}/p\mathbb{Z}$ (because modulo 5, $f = (t+2)(t+3)$), but it has no root in A, because A embeds in $\mathbb{R}$.

□

VII.3.4 Witt vectors

NB : The goal of this exercise is to show that, for every algebraically closed (or even just perfect) field k of characteristic $p > 0$, there exists a complete discrete valuation ring Λ with residue field k and fraction field K of characteristic 0. To apply the results of the last sections of chapter III, we also need to be able to construct a Λ such that K contains enough roots of 1, which doesn't follow immediately from this problem. (In addition to the results of this problem, we also need to know that the integral closure of a complete discrete valuation ring in a finite extension of its ring of fractions is still a complete discrete valution ring.)

Let p be a prime number. For every $n \geq 0$, we define a polynomial $W_n \in \mathbb{Z}[X_0, \dots, X_n]$ by

$$W_n(X_0, \dots, X_n) = \sum_{i=0}^{n} p^i X_i^{p^{n-i}}.$$

These are called the *Witt polynomials*.

If A is a commutative ring, define a map $\mathscr{W} : A^{\mathbb{N}} \to A^{\mathbb{N}}$ by sending $\underline{a} = (a_0, a_1, \dots) \in A^{\mathbb{N}}$ to $(W_0(a_0), W_1(a_0, a_1), W_2(a_0, a_1, a_2), \dots)$.

(1). We say that a commutative ring A is *p-unramified* if n is not a zero divisor in A for every positive integer and there exists a ring endomorphism τ of A such that, for every $a \in A$, $\tau(a) - a^p \in pA$.

  (a) Show that $\mathbb{Z}$ is p-unramified.

  (b) If A is p-unramified and I is any set, show that the polynomial ring $A[X_i, i \in I]$ is p-unramified.

(2). Let A be a commutative ring. For every f in $A[x]$, we define a sequence $(f^{\circ(n)})_{n\geq 0}$ of elements of $A[x]$ in the following way : $f^{\circ(0)}(x) = x$ and, for every $n \geq 0$, $f^{\circ(n+1)}(x) = f^{\circ(n)}(f(x))$.

Let $h \in A[x]$, and let $f(x) = x^p + ph(x)$. Show that, for every $a, b \in A$ and $n \geq 0$:

(a) $b^n - a^n \in (b-a)A$.

(b) If $b - a \in pA$, then $b^p - a^p \in p(b-a)A$.

(c) $h(b) - h(a) \in (b-a)A$.

(d) If $b - a \in pA$, then $f^{\circ(n)}(b) - f^{\circ(n)}(a) \in p^n(b-a)A$.

(3). Let A be a p-unramified ring, and let τ be a ring endormophism of A such that $\tau(x) - x^p \in pA$ for every $x \in A$. Show that $\mathscr{W} : A^{\mathbb{N}} \to A^{\mathbb{N}}$ is injective, and that its image is the set of $(w_0, w_1, \dots) \in A^{\mathbb{N}}$ such that, for every $n \geq 1$, $w_n - \tau(w_{n-1}) \in p^n A$. *(Hint : To show that every element in the image of $\mathscr{W}$ satisfies the stated condition, use (2) with $h = 0$.)*

(4). Let $\Phi \in \mathbb{Z}[X, Y]$. Show that there is a unique sequence of polynomials $\varphi_n \in \mathbb{Z}[X_0, \dots, X_n, Y_0, \dots, Y_n]$, $n \geq 0$, such that, for every $n \geq 0$:

$$W_n(\varphi_0(X_0, Y_0), \dots, \varphi_n(X_0, \dots, X_n, Y_0, \dots, Y_n)) = \Phi(W_n(X_0, \dots, X_n), W_n(Y_0, \dots, Y_n)).$$

(Hint : Use (3) for $A = \mathbb{Z}[X_0, Y_0, X_1, Y_1, \dots]$ and τ defined by $\tau(X_i) = X_i^p$, $\tau(Y_i) = Y_i^p$. But don't forget to check that φ_n only depends on $X_0, \dots, X_n$.)

(5). Applying (4) to the polynomial $X + Y$ (resp. XY), we get a sequence of polynomials $(S_n)_{\geq 0}$ (resp. $(P_n)_{n \geq 0}$). Let A be a commutative ring. We define two operations on $A^{\mathbb{N}}$ by the formulas

$$\underline{a} + \underline{b} = (S_0(a_0, b_0), \dots, S_n(a_0, \dots, a_n, b_0, \dots, b_n), \dots)$$

$$\underline{a} \cdot \underline{b} = (P_0(a_0, b_0), \dots, P_n(a_0, \dots, a_n, b_0, \dots, b_n), \dots)$$

for $\underline{a} = (a_0, a_1, \dots)$ and $\underline{b} = (b_0, b_1, \dots)$ in $A^{\mathbb{N}}$.

We write $W(A)$ for the set $A^{\mathbb{N}}$ with these two laws.

(a) Calculate S_0, S_1, P_0 and P_1.

(b) If p is invertible in A, show that $\mathscr{W} : A^{\mathbb{N}} \to A^{\mathbb{N}}$ is bijective and sends these two laws to the usual addition and multiplication (term by term) on $A^{\mathbb{N}}$.

In particular, $W(A)$ is a ring (isomorphic to the product ring $A^{\mathbb{N}}$.)

(c) Show that, for any commutative ring A, $W(A)$ is a commutative ring with zero element $(0, 0, 0, \dots)$ and unit element $(1, 0, 0, \dots)$. (Hint : Find commutative rings $B \subset C$ such that B surjects to A and that (ii) applies to C.)

This ring is called the *ring of Witt vectors of A*. (Remember the trick that we used in this question, it will be useful again.)

(d) Show that the map $W(A) \to A$ that sends $(a_0, a_1, \dots)$ to a_0 is a morphism of rings, and that the map $A \to W(A)$, $a_0 \longmapsto [a_0] = (a_0, 0, 0, \dots)$ is multiplicative.

(e) Suppose that p is not a zero divisor in A. Let $\underline{a} = (a_0, a_1, \dots) \in W(A)$, and let $\underline{b} = (b_0, b_1, \dots)$ be the unique element of $W(A)$ such that $\mathscr{W}(\underline{b}) = \mathscr{W}(p \cdot \underline{a})$. Show that $b_0 \in pA$ and that, for every $n \geq 0$, $b_{n+1} - a_n^p \in pA$.

(f) If $p = 0$ in A, show that, for every $\underline{a} = (a_0, a_1, \dots) \in W(A)$, $p \cdot \underline{a} = (0, a_0^p, a_1^p, \dots)$.

(g) For every $n \geq 0$, let I_n be the set of $(a_0, a_1, \dots) \in W(A)$ such that $a_i = 0$ for $0 \leq i \leq n$. Show that all the I_n are ideals of $W(A)$, that two elements $\underline{a} = (a_i)$ and $\underline{b} = (b_i)$ of $W(A)$ are equal modulo I_n if and only if $a_i = b_i$ for $0 \leq n - 1$, and that the map

$$W(A) \to \varprojlim_n W(A)/I_n := \{(x_n)_{n\geq 0} \in \prod_{n\geq 0} W(A)/I_n | \forall n,\ x_{n+1} = x_n \mod I_n\}$$

sending x to the family $(x + I_n)_{n\geq 0}$ is a ring isomorphism.

From now on, we take k to be a perfect ring of characteristic p. ("Of characteristic p" means that $p = 0$ in k, and "perfect" means that that the map $k \to k$, $x \longmapsto x^p$ (which respects addition because $p = 0$ in k) is an automorphism of rings.

(h) Let I be the ideal of $W(k)$ generated by p. Show that the map

$$W(k) \to \varprojlim_n W(k)/I^n := \{(x_n)_{n\geq 0} \in \prod_{n\geq 0} W(k)/I^n | \forall n,\ x_{n+1} = x_n \mod I^n\}$$

that sends x to the family $(x + I^n)_{n\geq 0}$ is a ring isomorphism.

(i) If k is a field, show that $W(k)$ is a complete discrete valuation ring (see problem VII.3.3) with residue field k and uniformizer p, and that the fraction field of $W(k)$ is a field of characteristic 0.

Solution.

(1). (a) Show that $\mathbb{Z}$ is p-unramified.

Take $\tau = \mathrm{id}_{\mathbb{Z}}$.

(b) As A is a subring and a quotient of $A[X_i, i \in I]$, so any element of A that is a zero divisor in $A[X_i, i \in I]$ is also a zero divisor in A. In particular, positive integers cannot be zero divisors in $A[X_i, i \in I]$.

If $\tau : A \to A$ is a ring automorphism such that $\tau(a) - a^p \in pA$ for every $a \in A$, extend τ to $A[X_i, i \in I]$ by setting $\tau(X_i) = X_i^p$ for every $i \in I$. This obviously satisfies the condition.

(2). (a) If $n = 0$, $b^n - a^n = 0 \in (b - a)A$. if $n \geq 1$, $b^n - a^n = (b - a)\sum_{i=0}^{n-1} b^{n-1-i}a^i \in (b - a)A$.

(b) We have

$$b^p = ((b-a)+a)^p = (b-a)^p + a^p + \sum_{k=1}^{p-1} \binom{n}{k} (b-a)^k a^{n-k},$$

so $b^p = (b-a)^p + a^p$ modulo $p(b-a)A$, which implies the conclusion (as $p(b-a)$ divides $(b-a)^p$).

(c) This follows directly from (a).

(d) We prove the result by induction on n. It's obvious for $n = 0$. Let $n \geq 0$, and suppose the result known for n. Write $x = f^{\circ(n)}(a)$, $y = f^{\circ(n)}(b)$. Then

$$f^{\circ n+1}(a) - f^{\circ(n+1)}(b) = f(x) - f(y) = x^p - y^p + p(h(x) - h(y)).$$

By (c), $h(x) - h(y) \in (x-y)A$. By (d), $x^p - y^p \in p(x-y)A$. As $(x-y) \in p^n(b-a)A$ by the induction hypothesis, we see that $f^{\circ n+1}(a) - f^{\circ(n+1)}(b) \in p^{n+1}(b-a)A$.

(3). We show by induction on $n \geq 0$ that there exist polynomials $Z_n \in \mathbb{Z}[p^{-1}][X_0, \ldots, X_n]$ such that $Z_n(W_0, \ldots, W_n) = X_n$ and $W_n(Z_0, \ldots, Z_n) = X_n$.

We take $Z_0 = X_0$. Suppose that we have constructed $Z_0, \ldots, Z_{n-1}$, for some $n \geq 1$. Note that $W_n = p^n X_n + W_{n-1}(X_0^p, \ldots, X_{n-1}^p)$. Let $f = W_{n-1}(Z_0^p, \ldots, Z_{n-1}^p)$. Then

$$f(W_0, \ldots, W_{n-1}) = W_{n-1}(X_0^p, \ldots, X_{n-1}^p),$$

so, if $Z_n = p^{-n}(X_n - f)$, then

$$X_n = p^{-n}(W_n - W_{n-1}(X_0^p, \ldots, X_{n-1}^p)) = Z_n(W_0, \ldots, W_n).$$

On the other hand,

$$W_n(Z_0, \ldots, Z_n) = p^n Z_n + W_{n-1}(Z_0^p, \ldots, Z_{n-1}^p) = X_n - f + f = X_n.$$

Let $A' = A[p^{-1}]$. As p is not a zero divisor in A, the obvious map $A \to A'$ is injective. The family $(Z_n)_{n \geq 0}$ defines a map $(A')^{\mathbb{N}} \to (A')^{\mathbb{N}}$, that is an inverse of $\mathscr{W}$ by construction, so $\mathscr{W} : (A')^{\mathbb{N}} \to (A')^{\mathbb{N}}$ is a bijection. As A injects in A', $\mathscr{W} : A^{\mathbb{N}} \to A^{\mathbb{N}}$ is injective.

Note that, to show the statements above, we only used that p is invertible in A' and that $A \to A'$ is invertible. (Not that A is p-unramified.)

Let $\underline{a} = (a_0, a_1, \ldots) \in A^{\mathbb{N}}$, and let $(w_0, w_1, \ldots) = \mathscr{W}(\underline{a})$. We want to show that for, for every $n \geq 1$, $w_n - \tau(w_{n-1}) \in p^n A$. We apply (b)(iv) with $h = 0$, so $f^{\circ(k)} = x^{p^k}$ for every $k \geq 0$. If $a \in A$, then $a^p - \tau(a) \in pA$, so

$$a^{p^n} - \tau(a^{p^{n-1}}) = f^{\circ(n-1)}(a^p) - f^{\circ(n-1)}(\tau(a)) \in p^{n-1} pA = p^n A.$$

Applying (2)(c) with $h = W_{n-1}$, this gives

$$W_{n-1}(a_0^p, \dots, a_{n-1}^p) - W_{n-1}(\tau(a_0), \dots, \tau(a_{n-1})) \in p^n A.$$

Finally,

$$w_n - \tau(w_{n-1}) = p^n a_n + W_{n-1}(a_0^p, \dots, a_{n-1}^p) - W_{n-1}(\tau(a_0), \dots, \tau(a_{n-1})) \in p^n A.$$

Let $\underline{w} = (w_0, w_1, \dots) \in A^{\mathbb{N}}$ such that, for every $n \geq 1$, $w_n - \tau(w_{n-1}) \in p^n A$. We want to show that $\underline{w}$ is in the image of $\mathscr{W}$, by finding $\underline{a} = (a_0, a_1, \dots) \in A^{\mathbb{N}}$ such that $\mathscr{W}(\underline{a}) = \underline{w}$. We construct the a_n by induction on n. Take $a_0 = w_0$. Let $n \geq 1$, and suppose that we have found $a_0, \dots, a_{n-1} \in A$ such that $W_i(a_0, \dots, a_i) = w_i$ for $0 \leq i \leq n-1$. We wan to find $a_n \in A$ such that

$$w_n = W_n(a_0, \dots, a_n) = p^n a_n + W_{n-1}(a_0^p, \dots, a_{n-1}^p).$$

Applying (2)(c) to $h = W_{n-1}$ as above, we get that, modulo $p^n A$:

$$w_n - W_{n-1}(a_0^p, \dots, a_{n-1}^p) = w_n - W_{n-1}(\tau(a_0), \dots, \tau(a_{n-1})) = w_n - \tau(w_{n-1}) = 0,$$

so there exists $a_n \in A$ such that $p^n a_n = w_n - W_{n-1}(a_0^p, \dots, a_{n-1}^p)$.

(4). Let $A = \mathbb{Z}[X_0, Y_0, X_1, Y_1, \dots]$, and let τ be the ring endomorphism of A that sends X_i (resp. Y_i) to X_i^p (resp. Y_i^p). By (1)(b), A is a p-ring. For every $n \geq 0$, let $w_n = \Phi(W_n(X_0, \dots, X_n), W_n(Y_0, \dots, Y_n))$. Then $\underline{w} := (w_0, w_1, \dots) \in A^{\mathbb{N}}$, and finding a sequence $(\varphi_n)_{n \geq 0}$ of elements of $\mathbb{Z}[X_0, X_1, Y_0, Y_1, \dots]$ satisfying the conditions of the statement amounts to showing that $\underline{w}$ is in the image of $\mathscr{W} : A^{\mathbb{N}} \to A^{\mathbb{N}}$. Also, the uniqueness of the sequence $(\varphi_n)_{\geq n}$ follows from the injectivity of $\mathscr{W} : A^{\mathbb{N}} \to A^{\mathbb{N}}$. By (3), we just need to show that $w_n - \tau(w_{n-1}) \in p^n A$ for every $n \geq 1$. But $w_n - \tau(w_{n-1})$ is equal to

$$\Phi(W_n(X_0, \dots, X_n), W_n(Y_0, \dots, Y_n)) - \Phi(W_{n-1}(Y_0^p, \dots, Y_{n-1}^p), W_{n-1}(Y_0^p, \dots, Y_{n-1}^p))$$

As

$$W_n(X_0, \dots, X_n) - W_{n-1}(X_0^p, \dots, X_{n-1}^p) = p^n X_n \in p^n A,$$

an easy generalization of (2)(c) to polynomials with two indeterminates shows that $w_n - \tau(w_{n-1}) \in p^n A$.

It remains to show that, for every $n \geq 0$, φ is in $\mathbb{Z}[X_0, \dots, X_n, Y_0, \dots, Y_n]$ (and not just in the bigger ring A). In the proof of (3), we have constructed a family of polynomials $Z_n \in \mathbb{Z}[p^{-1}](X_0, \dots, X_n)$ such that $X_n = Z_n(W_0, \dots, W_n)$ for every $n \geq 0$. Applying this to the equation in the statement of (4) gives $\varphi_n = Z_n(w_0', \dots, w_n')$, with $w_i' = \Phi(W_i(X_0, \dots, X_i), W_i(Y_0, \dots, Y_i))$. So $w_n \in \mathbb{Z}[p^{-1}][X_0, \dots, X_n, Y_0, \dots, Y_n] \cap A = \mathbb{Z}[X_0, \dots, X_n, Y_0, \dots, Y_n]$.

(5). (a) We have $S_0 = X_0 + Y_0$, $S_1 = X_1 + Y_1 + \frac{1}{p}(X_0^p + Y_0^p - (X_0 + Y_0)^p)$, $P_0 = X_0 Y_0$ and $P_1 = X_0^p Y_1 + Y_0^p X_1 + p X_1 Y_1$.

(b) If p is invertible in A, we showed in the beginning of the proof of (3) that $\mathscr{W} : A^{\mathbb{N}} \to A^{\mathbb{N}}$ is bijective. By the very definition of the polynomials S_n and P_n, $\mathscr{W}$ sends the two laws on $W(A)$ to the usual addition and multiplication on the product ring $A^{\mathbb{N}}$. So $W(A)$ is a commutative ring, with zero element $\mathscr{W}^{-1}(0, 0, \dots) = (0, 0, \dots)$ and unit element $\mathscr{W}^{-1}(1, 1, \dots) = (1, 0, 0, \dots)$. (To check that $(1, 0, \dots)$ is indeed the inverse image of $(1, 1, \dots)$ by $\mathscr{W}$, it suffices to check that $\mathscr{W}(1, 0, \dots) = (1, 1, \dots)$, which is obvious.)

(c) Let B be the ring of polynomials $\mathbb{Z}[X_a, a \in A]$, and let $v : B \to A$ be the ring morphism sending X_a to a, for every $a \in A$. We also set $C = B[p^{-1}]$ and write $u : B \to C$ for the inclusion. Then u and v induces maps $B^{\mathbb{N}} \to A^{\mathbb{N}}$, $B^{\mathbb{N}} \to C^{\mathbb{N}}$, $W(v) : W(B) \to W(A)$ and $W(u) : W(B) \to W(C)$. The first two are maps of rings, and the second two respect the two laws defined above (simply because u and v are ring morphisms). As $W(C)$ is a ring and the map $W(u) : W(B) \to W(C)$ is injective, we see that the two laws on $W(B)$ satisfy all the conditions imposed on the addition and multiplication laws of commutative a ring, and that the zero and unit elements are the ones given above. As $W(v) : W(B) \to W(A)$ is surjective, the two laws on $W(A)$ also satisfy all the necessary conditions to make $W(A)$ a commutative ring, with the zero and unit elements described above.

(d) The first statement follows from the formulas for S_0 and P_0 given in (a).

For the second statement, choose $v : B \to A$ and $u : B \to C$ as in the proof of (c). Then we have a commutative diagram

$$\begin{array}{ccccc} A & \xleftarrow{v} & B & \xrightarrow{u} & C \\ {\scriptstyle [\,]}\downarrow & & {\scriptstyle [\,]}\downarrow & & {\scriptstyle [\,]}\downarrow \\ W(A) & \xleftarrow{W(v)} & W(B) & \xrightarrow{W(u)} & W(C) \end{array}$$

The vertical map on the ring is multiplicative, because its composition with the ring isomorphism $\mathscr{W} : W(C) \to C^{\mathbb{N}}$ is the map $a \longmapsto (a, a^p, a^{p^2}, \dots)$, which is multiplicative. As $W(u)$ is an injective ring moprhism, $[\,] : B \to W(B)$ is also multiplicative. As $v : B \to A$ and $W(v) : W(B) \to W(A)$ are ring morphisms and v is surjective, $[\,] : A \to W(A)$ is also multiplicative.

(e) We show the statement by induction on n. Write $A' = A[p^{-1}]$. By the hypothesis on A, the obvious map $A \to A'$ is injective, and we use it to identify A to a subring of A'. Let $(w_0, w_1, \dots) = \mathscr{W}(a_0, a_1, \dots)$. Let $Z_0, Z_1, \dots \in \mathbb{Z}[p^{-1}][X_0, X - 1, \dots]$ be the polynomials defined in the proof of (c). We have $b_n = Z_n(pw_0, \dots, pw_n)$ for every $n \geq 0$. Then $Z_0 = X_0$ and $Z_1 = p^{-1}(X_1 - X_0^p)$, so $b_0 = pa_0 \in pA$ and we have in A' :

$$b_1 = p^{-1}(p(a_0^p + pa_1) - (pa_0)^p) = a_0^p + pa_1 - p^{p-1}a_0^p,$$

so $b_1 - a_0^p \in pA$. Now let $n \geq 1$, and suppose that we know that $b_{m+1} - a_m^p \in pA$ for every $m < n$. Then

$$Z_{n+1} = p^{-(n+1)}(X_{n+1} - W_n(Z_0^p, \dots, Z_n^p)),$$

so, in A',

$$b_{n+1} = p^{-(n+1)}(pW_{n+1}(a_0, \dots, a_{n+1}) - W_n(b_0^p, \dots, b_n^p)) =$$

$$p^{-(n+1)}(p(p^{n+1}a_{n+1} + W_n(a_0^p, \dots, a_n^p)) - W_n(b_0^p, \dots, b_n^p))$$

The right-hand side is equal to

$$pa_{n+1} + a_n^p - p^{-(n+1)}b_0^{p^{n+1}} - p^{-(n+1)}\sum_{i=1}^{n} p^i(b_i^{p^{n-i+1}} - a_{i-1}^{p^{n-i+2}}).$$

By the induction hypothesis and (2)(b), this is equal to a_n^p modulo pA.

(f) As in the proof of (d), choose a surjective ring morphism $v : B \to A$ such that p is not a zero divisor in B, and let $W(v) : W(B) \to W(A)$ be the induced map of rings. Fix $\underline{a} = (a_0, a_1, \dots) \in W(A)$, and choose $\underline{a}' = (a_0', a_1', \dots) \in W(B)$ such that $W(v)(\underline{a}') = \underline{a}$. Let $\underline{b}' = (b_0', b_1', \dots) = p\underline{a}'$. Then, by (v), $b_0' \in pB$ and $b_n' - {a_{n-1}'}^p \in pB$ for every $n \geq 1$. Now if $\underline{b} = (v(b_0'), v(b_1'), \dots)$, then $\underline{b} = W(v)(\underline{b}') = p\underline{a}$, and we have $b_0 = 0$ and $b_n = a_{n-1}^p$ for every $n \geq 1$.

(g) For every $n \geq 0$, let J_n be the subset of the product ring $A^{\mathbb{N}}$ made up of the $\underline{a} = (a_i)$ such that $a_i = 0$ for $0 \leq i \leq n-1$. Then J_n is obviously an ideal, and two elements $\underline{a} = (a_i)$ and $\underline{b} = (b_i)$ of $A^{\mathbb{N}}$ are equal modulo J_n if and only if $a_i = b_i$ for $0 \leq n-1$.

We apply the trick of (c). Choose a surjective ring morphism $u : B \to A$ and an injective ring morphism $v : B \to C$ such that p is invertible in C. We write $I_n(A)$, $I_n(B)$ and $I_n(C)$ for the subsets of $W(A)$, $W(B)$ and $W(C)$ defined above, and we use the same convention for J_n. The map $\mathscr{W} : W(C) \to C^{\mathbb{N}}$ is an isomorphism of rings, and $\mathscr{W}(I_n(C)) = J_n(C)$, so we get the fact that the $I_n(C)$ are ideals. The second assertion in this case follows from the easy fact that, for every $n \geq 0$ and every $a_0, \dots, a_{n-1}, b_0, \dots, b_{n-1} \in C$, $a_i = b_i$ for every $0 \leq i \leq n-1$ if and only $W_i(a_0, \dots, a_{n-1}) = W_i(b_0, \dots, b_{n-1})$ for every $0 \leq i \leq n-1$.

Now note that $I_n(B) = u^{-1}(I_n(C))$ and $I_n(A) = v(I_n(B))$, so they are ideals. Fix $n \geq 0$. If $\underline{a} = (a_i)$ and $\underline{b} = (b_i)$ are in $W(A)$, $W(B)$ or $W(C)$, we write $\underline{a} \sim \underline{b}$ to indicate that $a_i = b_i$ for $0 \leq i \leq n$.

If $\underline{a}$, $\underline{b}$ are in $W(B)$, then $\underline{a} \sim \underline{b}$ if and only if $u(\underline{a}) \sim u(\underline{b})$, and $\underline{a} = \underline{b}$ modulo $I_n(B)$ if and only if $u(\underline{a}) = u(\underline{b})$ modulo $I_n(C)$. This gives the second assertion for B.

Assume that $\underline{a}, \underline{b} \in W(A)$. If $\underline{a} \sim \underline{b}$, then there exists $\underline{a}', \underline{b}' \in W(B)$ such that $v(\underline{a}') = \underline{a}$, $v(\underline{b}') = \underline{b}$ and $\underline{a}' \sim \underline{b}'$; then $\underline{a}' = \underline{b}'$ modulo $I_n(B)$, so $\underline{a} = \underline{b}$ modulo $I_n(A)$. If on the other hand $\underline{a} = \underline{b}$ modulo $I_n(A)$, then there exists $\underline{a}', \underline{b}', \underline{c}' \in W(B)$ such that

$v(\underline{a}') = \underline{a}$, $v(\underline{b}') = \underline{b}$, $\underline{c}' \in \operatorname{Ker} v$ and $\underline{a}' = \underline{b}' + \underline{c}'$ modulo $I_n(B)$; then $\underline{a}' \sim \underline{b}' + \underline{c}'$, so $\underline{a} \sim \underline{b}$.

Finally, we prove the last assertion. Call φ the morphism $A \to \varprojlim_n A/I_n$. Then $\operatorname{Ker}\varphi = \bigcap_n I_n = 0$, so φ is injective. Now we prove that φ is surjective. Let $x = (x_n)_{n\geq 0} \in \varprojlim_n A/I_n$, and write $x_n = (a_0^n, a_1^n, \dots)$. By what we just proved, the condition $x_n = x_{n+1}$ modulo I_n gives $a_i^n = a_i^{n+1}$ for $0 \leq i \leq n$. Let $\underline{a} = (a_0^0, a_1^a, a_2^2, \dots) \in W(A)$. Then $\underline{a} = x_n$ modulo I_n for every $n \geq 0$, so $\varphi(\underline{a}) = x$.

(h) By (f), for every $n \geq 0$, I^n is the set of $(0, \dots, 0, a_0^{p^n}, a_1^{p^n}, \dots)$, with n zeroes at the beginning and $a_0, a_1, \dots \in k$. As k is perfect, I^n is thus equal to the ideal I_n of (g), and the assertion follows from (g).

(i) If k is a field, then the ideal $I = pW(k)$ of (h) is a maximal ideal. We want to show that it's the only maximal ideal, ie that every element of $W(k) - I$ is invertible. By (d), for every $a \in k - \{0\}$, $[a_0] = (a_0, 0, 0, \dots)$ is invertible (with inverse $[a_0^{-1}]$). Now let $\underline{a} = (a_0, a_1, \dots) \in W(k)$ such that $a_0 \neq 0$. Then, by (g) and (h), $\underline{a} = [a_0]$ modulo I, so there exists $\underline{b} \in I$ such that $\underline{a} = [a_0](1 - \underline{b})$. For every $n \geq 0$, let $x_n = \sum_{k=0}^n \underline{b}^n$. As $\underline{b}^{n+1} \in I^{n+1}$, $x_n = x_{n+1}$ modulo I^{n+1} for every n, so $(x_n)_{n\geq 0}$ is an element of $\varprojlim_n W(k)/I^n$. By (h), there exists $\underline{c} \in W(k)$ such that $\underline{c} = x_n$ modulo I^n for every $n \geq 0$. Then, for every $n \geq 0$, $(1 - \underline{b})\underline{c} = (1 - \underline{b})x_n = 1$ modulo I^n. By (h) again, $(1 - \underline{b})\underline{c} = 1$, so $1 - \underline{b}$ is invertible, hence so is $\underline{a}$. (Note that, if we only assume that k is a perfect ring of characteristic p, the same proof show that $\underline{a} = (a_i)$ is invertible in $W(k)$ if and only if a_0 is invertible in k.)

Using (f) and the fact that k is perfect, we get that, for every $\underline{a} \in W(k) - \{0\}$, there exists a unique $n \geq 0$ and a unique $\underline{b} = (b_0, b_1, \dots) \in W(k)$ such that $\underline{a} = p^n\underline{b}$ and $b_0 \neq 0$ (in k). By what we have just seen, $\underline{b}$ is invertible. We set $v(\underline{a}) = n$. We also set $v(0) = \infty$.

If $\underline{a}, \underline{a}' \in W(k) - \{0\}$, let $n = v(\underline{a})$ and $n' = v(\underline{a}')$, and write $\underline{a} = p^n u$, $\underline{a}' = p^{n'}u'$ with $u, u' \in W(k)^\times$. Then $\underline{a}\underline{a}' = p^{n+n'}uu'$. By (f), $\underline{a}\underline{a}' \neq 0$ and $v(\underline{a}\underline{a}') = n + n'$. In particular, $W(k)$ is a domain. Let K be its fraction field. We extend v to a map $K^\times \to \mathbb{Z}$ by setting $v(xy^{-1}) = v(x) - v(y)$ if $x, y \in W(k)$. By what we just proved, this makes sense and defines a group morphism. Note that $v : K^\times \to \mathbb{Z}$ is surjective, because $v(p^n) = n$, for every $n \in \mathbb{Z}$. Let $x \in K^\times$, and write $x = yz^{-1}$, with $y, z \in W(k) - \{0\}$. If $v(y) = n$ and $v(z) = m$, we have $v(x) = n - m$ and $x = p^{n-m}u$, with $u \in W(k)^\times$. So x is in $W(k)$ if and only if $v(x) \geq 0$. (Note that p is not invertible in $W(k)$, because it is in the maximal ideal.)

To finish the proof that $W(k)$ is a discrete valuation ring, we have to show that $v(x + y) \geq \inf(v(x), v(y))$ for every $x, y \in K^\times$. We may assume that $x, y \in W(k) - \{0\}$. Let $n = v(x)$ and $m = v(y)$, and assume that $n \geq m$ (this is always true up to switching x and y.) Write $x = p^n u$ and $y = p^m u'$ with $u, u' \in W(k)^\times$. then $x + y = p^m(p^{n-m}u + u')$, so

$v(x+y) = m + v(p^{n-m}u + u') \geq m = \inf(v(x), v(y))$.

By (h), $W(k)$ is complete. It remains to show that the characteristic of K is 0. We already know that $\mathrm{char}(K) \neq p$, because p is invertible in K. If $\mathrm{char}(K) = \ell > 0$, then $p^\ell = p$ in K, which is not possible. So $\mathrm{char}(K) = 0$.

□

VII.4 Chapter IV exercises

None yet, but see problem VII.7.1.

VII.5 Chapter V exercises

VII.5.1 Existence of the Haar measure on a compact group

Let G be a Hausdorff locally compact topological group. By a measure on G, we mean a measure on the Borel σ-algebra of G. A nonzero measure μ on G is called a *left Haar measure* if, for every measurable function $f : G \to \mathbb{C}$ and every $g \in G$, $\int_G f(x)d\mu = \int_G f(gx)d\mu$. In general, left Haar measures always exist and they are unique up to scaling. Here we are only interested in the case of compact groups.

We denote by $\mathscr{C}$ the space of continuous functions with compact support $G \to \mathbb{C}$, equipped with the compact-open topology. This is the topology generated by the sets $\{f \in \mathscr{C} | f(K) \subset U\}$, for $K \subset G$ compact and $U \subset \mathbb{C}$ open.

Let $\mathscr{C}^*$ be the space of continuous linear maps $\mathscr{C} \to \mathbb{C}$, equipped with the weak topology. Remember that the weak topology is the coarsest topology such that the maps $\mathscr{C}^* \to \mathbb{C}$, $\lambda \longmapsto \lambda(f)$, are continuous for every $f \in \mathscr{C}$. So a base of opens is given by the sets

$$\{\lambda \in \mathscr{C}^* | |(\lambda - \lambda_1)(f_1)| < r_1, \ldots, |(\lambda - \lambda_n)(f_n)| < r_n\},$$

for $\lambda_1, \ldots, \lambda_n \in \mathscr{C}^*$, $f_1, \ldots, f_n \in \mathscr{C}$ and $r_1, \ldots, r_n \in \mathbb{R}_{>0}$.

Now assume that G is compact and either is metrizable, or has a countable basis for its topology. This hypothesis is just here to guarantee the following fact (that you don't have to prove) : for every probability measure μ on G, the linear form $f \longmapsto \int_G f(x)d\mu$ on $\mathscr{C}$ is continuous, hence in $\mathscr{C}^*$. So we identify the set $\mathscr{P}(G)$ of probability measures on G with a subset of $\mathscr{C}^*$.

(1). Show that the compact-open topology on $\mathscr{C}$ is the same as the topology of uniform convergence, i.e. the topology induced by the norm $\| \cdot \|_\infty$ given by

$$\|f\|_\infty = \sup_{x \in G} |f(x)|.$$

(2). Let ρ be the representation of G on $\mathscr{C}$ defined by $(\rho(g)f)(x) = f(g^{-1}x)$, for $f \in \mathscr{C}$ and $x, g \in G$. Show that the map $G \times \mathscr{C} \to \mathscr{C}$, $(g, f) \longmapsto \rho(g)f$ is continuous (ie ρ is a continuous representation of G).

(3). Show that the contragredient representation ρ^* (defined as usual by $\rho^*(g)\lambda = \lambda \circ \rho(g^{-1})$, if $g \in G$ and $\lambda \in \mathscr{C}^*$) is also continuous.

(4). Show that $\mathscr{P}(G)$ is a convex compact G-invariant subset of $\mathscr{C}^*$. *(Hint : Tychonoff's theorem.)*

(5). Show that G has a left Haar measure. *(Hint : An appropriate form of the Markov-Kakutani fixed point theorem might help. For example theorem 5.11 of Rudin's* Functional analysis*.)*

Solution.

(1). Let $f_0 \in \mathscr{C}$. We want to show that every open neighbourhood of f_0 in the compact-open topology contains an open neighbourhood of f_0 in the topology of uniform convergence, and vice versa.

For the first direction, we may assume that the open neighbourhood of f_0 if of the form $X := \{f \in \mathscr{C} | f(K_1) \subset U_1, \ldots, f(K_m) \subset U_m\}$ where $K_1, \ldots, K_m \subset G$ are compact and $U_1, \ldots, U_m \subset \mathbb{C}$ are open. For every $\varepsilon > 0$, we write

$$V_\varepsilon = \{f \in \mathscr{C} | \|f - f_0\|_\infty < \varepsilon\}.$$

Let $i \in \{1, \ldots, m\}$. As $f_0(K_i) \subset U_i$ is compact, there exists $\varepsilon_i > 0$ such that $\{x \in \mathbb{C} | \exists y \in f_0(K_i),\ |x - y| < \varepsilon_i\} \subset U_i$. [12] Let $\varepsilon = \min(\varepsilon_1, \ldots, \varepsilon_m)$. Then $V_\varepsilon \subset X$. Indeed, if $f \in V_\varepsilon$ and $i \in \{1, \ldots, m\}$, then for every $x \in K_i$, $|f(x) - f_i(x)| < \varepsilon$, so $f(x) \in U_i$ by the choice of ε.

Conversely, let $\varepsilon > 0$ and let V_ε be defined as above. We have to show that V_ε contains an open neighbourhood of f_0 in the compact-open topology. Let $\eta = \varepsilon/2$. For every $x \in G$, let

$$U_x = \{y \in G | |f_0(x) - f_0(y)| < \eta\} \subset K_x = \{y \in G | |f_0(x) - f_0(y)| \leq \eta\}.$$

Then U_x is open and K_x is compact. As $G = \bigcup_{x \in G} U_x$ and G is compact, there exist $x_1, \ldots, x_n \in G$ such that $G = \bigcup_{i=1}^n U_{x_i} = \bigcup_{i=1}^n K_{x_i}$. For every $i \in \{1, \ldots, n\}$, write $K_i = K_{x_i}$ and let $U_i = \{a \in \mathbb{C} | |f_0(x_i) - a| < \eta\}$. Then $X := \{f \in \mathscr{C} | \forall i \in \{1, \ldots, n\},\ f(K_i) \subset U_i\} \subset V_\varepsilon$. Indeed, let $f \in X$ and let $x \in G$. Then there exists $i \in \{1, \ldots, n\}$ such that $x \in K_i$, and we have

$$|f(x) - f_0(x)| \leq |f(x) - f_0(x_i)| + |f_0(x_i) - f_0(x)| < \eta + \eta = \varepsilon.$$

[12] This is a standard compactness argument. For every $x \in f_0(K_i)$, choose $\varepsilon_x > 0$ such that $B(x, \varepsilon_x) \subset U_i$, where $B(x, r) = \{y \in \mathbb{C} | |x - y| < r\}$. Then $f_0(K_i) \subset \bigcup_{x \in f_0(K_i)} B(x, \varepsilon_x/2)$, so there exists $x_1, \ldots, x_r \in f_0(K_i)$ such that $f_0(K_i) \subset \bigcup_{j=1}^r B(x_j, \varepsilon_{x_j}/2)$. Then $\varepsilon_i = \frac{1}{2}\min(\varepsilon_{x_1}, \ldots, \varepsilon_{x_r})$ will work.

(2). Denote by μ the map $G \times \mathscr{C} \to \mathscr{C}$. Let $g_0 \in G$ and $f_0 \in \mathscr{C}$. We want to prove that, for every $\varepsilon > 0$, there exists a neighbourhood W of g_0 in G and $\eta > 0$ such that, if $g \in W$ and $\|f - f_0\|_\infty < \eta$, then $\|gf - g_0 f_0\|_\infty < \varepsilon$. Fix $\varepsilon > 0$.

As G is compact, $f_0 : G \to \mathbb{C}$ is uniformely continuous, so there exists two collections $U_1, \dots, U_n$ and $V_1, \dots, V_n$ of open subsets of G such that $K_i := \overline{U_i} \subset V_i$ for every i, $G = U_1 \cup \cdots \cup U_n$ and, for every i and every $x, y \in V_i$, $|f_0(x) - f_0(y)| < \varepsilon/2$. Let $i \in \{1, \dots, n\}$. For every $x \in K_i$, choose j such that $g_0^{-1}x$ is in V_j and open neighbourhoods U_x of x in G and W_x of g_0 in G such that, for every $g \in W_x$ and $y \in U_x$, $g^{-1}y \in V_j$. As K_i is compact, we can choose $x_1, \dots, x_m$ such that $K_i \subset U_{x_1} \cup \cdots \cup U_{x_m}$. Then the open neighbourhood $W_i := W_{x_1} \cap \cdots \cap W_{x_m}$ of g_0 has the property that, for every $g \in W_i$ and every $x \in K_i$, there exists $j \in \{1, \dots, n\}$ such that both $g_0^{-1}x$ and $g^{-1}x$ are in V_j. Finally, let $W = W_1 \cap \cdots \cap W_n$. Let $g \in W$ and $f \in \mathscr{C}$ such that $\|f - f_0\|_\infty < \varepsilon/2$. We want to show that $\|gf - g_0 f_0\|_\infty < \varepsilon$. Let $x \in G$. Choose $i \in \{1, \dots, n\}$ such that $x \in K_i$. Then there exists $j \in \{1, \dots, n\}$ such that $g_0^{-1}x, g^{-1}x \in V_j$. Then

$$(gf - g_0 f_0)(x) = f(g^{-1}x) - f_0(g_0^{-1}x) = (f(g^{-1}x) - f_0(g^{-1}x)) + (f_0(g^{-1}x) - f_0(g_0^{-1}x)),$$

so

$$|(gf - g_0 f_0)(x)| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

(3). Let $g_0 \in G$, $\lambda_0 \in \mathscr{C}^*$, $f \in \mathscr{C}$ and $r \in \mathbb{R}_{>0}$. We want to find a neighbourhood W of g_0 in G and a neighbourhood U of λ in $\mathscr{C}^*$ such that, for every $g \in W$ and $\lambda \in U$, $|(g\lambda - g_0\lambda_0)(f)| < r$. Let

$$\|\lambda_0\| = \sup\{\|\lambda_0(f_1)\|_\infty, f_1 \in \mathscr{C}, \|f_1\|_\infty = 1\}.$$

Then $\|\lambda_0\|$ is finite because λ_0 is continuous, and $\|\lambda_0(f_1)\|_\infty \leq \|\lambda_0\|\|f_1\|_\infty$ for every $f_1 \in \mathscr{C}$.

By (2) (or just the uniform continuity of f), there exists a neighbourhood W of g_0 in G such that, for every $g \in W$, $\|g^{-1}f - g_0^{-1}f\|_\infty < r/(2(1 + \|\lambda_0\|))$. Let $U = \{\lambda \in \mathscr{C}^* | |(\lambda - \lambda_0)(g^{-1}f)| < r/2\}$. Then U is a neighbourhood of λ_0 in $\mathscr{C}^*$, and we have, for every $g \in W$ and $\lambda \in U$,

$$|(g\lambda - g_0\lambda_0)(f)| = |\lambda(g^{-1}f) - \lambda_0(g^{-1}f)| + |\lambda_0(g^{-1}f) - \lambda_0(g_0^{-1}f)| < r.$$

(4). The set $\mathscr{P}(G)$ is obviously convex and G-invariant, the harder thing is proving that it is compact.

We denote by B its unit ball of $\mathscr{C}$ (for the norm $\|\cdot\|_\infty$). Let B_+ be the set of $f \in B$ such that $f(G) \subset \mathbb{R}_{\geq 0}$. Then B^+ generates $\mathscr{C}$ as a vector space,[13] so the map $\mathscr{C}^* \to \prod_{B^+} \mathbb{C}$,

[13]Obviously B generates $\mathscr{C}$. If $f \in \mathscr{C}$, then $f = f_1 + if_2$, with $f_1, f_2 : G \to \mathbb{R}$ uniquely determined, and we have $f_1, f_2 \in B$ if $f \in B$. Finally, if $f \in \mathscr{C}$ has real values, then $f = f^+ - f^-$ with $f^+(x) = \sup(f(x), 0)$ and $f^-(x) = -\inf(f(x), 0)$, and $f^+, f^- \in B^+$ if $f \in B$.

$\lambda \longmapsto (\lambda(f))_{f\in\mathscr{C}}$, is injective, and we use it to identify $\mathscr{C}^*$ to a subspace of $\prod_{B^+}\mathbb{C}$. By definition, the weak topology on $\mathscr{C}^*$ is the topology induced by the product topology on $\prod_B \mathbb{C}$.

Let $K = \prod_B [0,1] \subset \prod_B \mathbb{C}$. By Tychonoff's theorem, K is compact. By the Riesz representation theorem, $\mathscr{C}^* \cap K$ is closed in K. Indeed, if $(a_f)_{f\in B^+}$ is in the closure of $\mathscr{C}^* \cap K$, then the map $B^+ \to \mathbb{C}$, $f \longmapsto a_f$ extends to a linear map $\mathscr{C} \to \mathbb{C}$. This linear map is positive because the a_f are in $\mathbb{R}_{\geq 0}$, so it is of the form $f \longmapsto \int f d\mu$, where μ is a regular Borel measure on G, and hence it is a continuous (for the compact-open topology on $\mathscr{C}$), ie an element of $\mathscr{C}^*$.

So $K \cap \mathscr{C}^*$ is compact, and this implies that $\mathscr{P}(G)$ is compact, because $\mathscr{P}(G) = \{\lambda \in K \cap \mathscr{C}^* | \lambda(1) = 1\}$ (where 1 the constant function 1 on G) is closed in $K \cap \mathscr{C}^*$.

(5). We want to apply the Markov-Kakutani fixed point theorem in the form of theorem 5.11 of Rudin's *Functional analysis* to the group G acting as above on the space $\mathscr{C}^*$, and to the invariant compact convex subset $\mathscr{P}(G)$ of $\mathscr{C}^*$. First, note that $\mathscr{P}(G)$ is nonempty because the linear functional $f \longmapsto f(1)$ is in $\mathscr{P}(G)$ (that's the Dirac measure at $1 \in G$).

Second, $\mathscr{C}^*$ is locally convex because the sets $\{\lambda \in \mathscr{C}^* | |\lambda(f_1) < r_1, \ldots, |\lambda(f_s)| < r_s\}$, for $f_1, \ldots, f_s \in \mathscr{C}$ and $r_1, \ldots, r_s \in \mathbb{R}_{>0}$, form a basis of neighbourhoods of 0 in $\mathscr{C}^*$ and they are convex.

Third, we have to check that the action of G on $\mathscr{P}(G)$ (not $\mathscr{C}^*$!) is equicontinuous. This means that, for every neighbourhood W of 0 in $\mathscr{C}^*$, there exists a neighbourhood V of 0 in $\mathscr{C}^*$ such that, for every $\mu_1, \mu_2 \in \mathscr{P}(G)$ such that $\mu_1 - \mu_2 \in V$, we have $g.(\mu_1 - \mu_2) \in W$ for every $g \in G$.

It is obviously enough to prove this for a neighbourhood of 0 of the form $W = \{\lambda \in \mathscr{C}^* | |\lambda(f)| < r\}$, where $f \in \mathscr{C}$ and $r \in \mathbb{R}_{>0}$ are fixed and $\|f\|_\infty = 1$.

As the action of G on $\mathscr{C}$ is continuous, the open sets $U_h = \{g \in G | \|(g^{-1}f - h^{-1}f)\|_\infty < r/3\}$, $h \in G$, cover G. As G is compact, we can find $h_1, \ldots, h_m \in G$ such that $G = \bigcup_{i=1}^m U_{h_i}$. Now let $V = \{\lambda \in \mathscr{C}^* | \forall i \in \{1, \ldots, m\},\ |\lambda(h_i^{-1}f)| < r/3\}$. This is a neighbourhood of 0 in $\mathscr{C}^*$. Let $\mu_1, \mu_2 \in \mathscr{P}(G)$ such that $\lambda := \mu_1 - \mu_2 \in V$. We want to prove that $g\lambda \in W$ for every $g \in G$. First notice that, as μ_1 and μ_2 are probability measures, $|\lambda(f_1)| \leq 2\|f_1\|_\infty$ for every $f_1 \in \mathscr{C}$. Let $g \in G$. There exists $i \in \{1, \ldots, m\}$ such that $g \in U_{h_i}$, ie $\|g^{-1}f - h_i^{-1}f\|_\infty < r/3$. Then

$$|(g\lambda)(f)| = |\lambda(g^{-1}f)| \leq |\lambda(g^{-1}f - h_i^{-1}f)| + |\lambda(h_i^{-1}f)| < 2r/3 + r/3 = r,$$

so $g\lambda \in W$.

Finally, the fixed point theorem gives an element $\mu \in \mathscr{P}(G)$ such that $g\mu = \mu$ for every $g \in G$. But this is the same as saying that μ is a left Haar measure.

□

VII.5.2 Haar measures are unique

Let G be a Hausdorff locally compact topological group. We denote by $\mathscr{C}(G,\mathbb{C})$ the space of continuous functions with compact support from G to $\mathbb{C}$. Remember (from remark V.1.3 of chapter V) that a left Haar measure on G is a nonzero positive $\mathbb{C}$-linear map $\lambda : \mathscr{C}(G,\mathbb{C}) \to \mathbb{C}$ such that, if μ is the measure on the σ-algebra of Borel sets of G corresponding to λ, for every continuous function $f : G \to \mathbb{C}$ and every $g \in G$, $\int_G f(x)d\mu = \int_G f(gx)d\mu$.

For simplicity, we will assume in this problem that the group G is a *normal* topological space. This is the case for example if we assume that G is *σ-compact* (i.e., G is a countable union of compact subsets).

Remember Urysohn's lemma :

Theorem. *If X is a normal topological space, then, for every closet subsets $Y, Z \subset X$ such that $Y \cap Z = \varnothing$, there exists a continuous function $f : X \to [0,1]$ such that $f(Y) = \{0\}$ and $f(Z) = \{1\}$.*

Let $\lambda_1, \lambda_2 : \mathscr{C}(G,\mathbb{C}) \to \mathbb{C}$ be two left Haar measures on G, and let μ_1, μ_2 be the corresponding measures on the σ-algebra of Borel sets of G. We want to show that λ_1 and λ_2 are equal up to a real positive scalar.

(1). Show that it is enough to see that $\lambda_1(f)\lambda_2(g) = \lambda_1(g)\lambda_2(f)$ for every $f, g \in \mathscr{C}(G,\mathbb{C})$.

(2). Show that, for every non-empty open subset $U \subset G$, there exists $\psi \in \mathscr{C}(G,\mathbb{C})$ taking only non-negative values, supported in U, and such that $\lambda_1(\psi) = 1$.

(3). Let $A \subset \mathscr{C}(G,\mathbb{C})$ be a finite subset and $\varepsilon > 0$. Find $0 \neq \psi_\varepsilon \in \mathscr{C}(G,\mathbb{C})$ taking non-negative values and such that, for every $f \in A$,

$$\lambda_2(f) = \lambda_1(f) \int_G (\psi_\varepsilon(x^{-1}))d\mu_2(x) + O(\varepsilon)$$

(where "$O(\varepsilon)$" means "bounded by $C\varepsilon$, for $C \in \mathbb{R}_{\geq 0}$ depending on A but not on ε").

(Hint : Find ψ_ε by applying (2) to a well-chosen neighborhood of 1 in G. To show the property of ψ_ε, calculate $\int_G \int_G f(xy)\psi_\varepsilon(y)\, d\mu_1(y)d\mu_2(x)$ in two different ways.)

(4). Conclude.

Solution.

(1). Suppose that $\lambda_1(f)\lambda_2(g) = \lambda_1(g)\lambda_2(f)$ for any $f, g \in \mathscr{C}(G,\mathbb{C})$. Let $A_1, A_2 \subset G$ be two compact subsets such that $\mu_1(A_1) > 0$ and $\mu_2(A_2) > 0$. (These exist because μ_1 and

μ_2 are regular.) By Urysohn's lemma, we can find $g : G \to [0,1]$ continuous such that $g_{|A_1 \cup A_2} = 1$. Then $\lambda_i(g) \in \mathbb{R}_{>0}$ for $i = 1, 2$, and we have $\lambda_1 = \frac{\lambda_1(g)}{\lambda_2(g)}\lambda_2$.

(2). Let $U \subset G$ be open nonempty. Then any compact subset of G can be covered by a finite number of left translates of U, which all have volume equal to that of U for the measure μ_1 (thanks to left invariance). As there exists compact subsets of G with nonzero volume, $\mu_1(U) \neq 0$. As μ_1 is regular, there exists a compact subset K of U such that $\mu_1(K) > 0$. Then $\lambda_1(\mathbb{1}_K) > 0$. By theorem 3.14 of Rudin's book [25], there exsist a function $f_1 \in \mathscr{C}(U, \mathbb{R}_{\geq 0})$ such that $\lambda_1(\mathbb{1}_K - f) < \frac{1}{2}\lambda_1(\mathbb{1}_K)$, hence $\lambda_1(f) > 0$. If we extend f to G by taking $f(x) = 0$ for $x \notin U$, this is still continuous. Now we just take $\psi = \frac{1}{\lambda_1(f)}f$.

(3). For every $\varepsilon > 0$, choose a neighborhood U_ε of 1 in G such that, for every $f \in A$, every $x \in G$ and every $y \in U_\varepsilon$, $|f(x) - f(xy)| \leq \varepsilon$. Let $\psi_\varepsilon : G \to \mathbb{R}_{\geq 0}$ be a continuous function with compact support included in U_ε such that $\lambda_1(\psi_\varepsilon) = 1$. (Such a function exists by (2).) We may assume that the supports of all the ψ_ε, for $0 < \varepsilon < 1$, are contained in some compact subset K_1 of G.

Choose a compact subset K_2 of G containing the supports of all the elements of A, and let $K = K_1 \cup K_1K_2$. (Another compact subset of G.) Let $f \in A$ and $\varepsilon \in]0,1[$. Then, for every $x \in G$,

$$|f(x) - \int_G f(xy)\psi_\varepsilon(y)d\mu_1(y)| = |\int_G (f(x) - f(xy))\psi_\varepsilon(y)d\mu_1(y)| \leq \varepsilon \int_G \psi_\varepsilon(y)d\mu_1(y) = \varepsilon,$$

so

$$|\int_G \int_G f(xy)\psi_\varepsilon(y)\, d\mu_1(y)d\mu_2(x) - \lambda_2(f)| \leq \int_K \varepsilon d\mu_2(x) \leq \varepsilon\mu_2(K).$$

On the other hand, using the change of variables $z = xy$ and the left invariance of μ_1, we get

$$\int_G \int_G f(xy)\psi_\varepsilon(y)\, d\mu_1(y)d\mu_2(x) = \int_G f(z)(\int_G \psi_\varepsilon(x^{-1}z)d\mu_2(x))d\mu_1(z).$$

Using the left invariance of μ_2 (and the change of variables $x' = z^{-1}x$), we see that, for every $z \in G$,

$$\int_G \psi_\varepsilon(x^{-1}z)d\mu_2(x) = \int_G \psi_\varepsilon(x^{-1})d\mu_2(x).$$

So

$$\int_G \int_G f(xy)\psi_\varepsilon(y)\, d\mu_1(y)d\mu_2(x) = \lambda_1(f) \int_G \psi_\varepsilon(x^{-1})d\mu_2(x),$$

which gives the result.

(4). Let $f, g \in \mathscr{C}(G, \mathbb{C})$, and take $A = \{f, g\}$. Then we have found in (3) a constant $C \in \mathbb{R}_{>0}$ and functions $\psi_\varepsilon \in \mathscr{C}(G, \mathbb{R}_{\geq 0})$, for every $\varepsilon \in]0,1[$, such that

$$|\lambda_2(f) - \lambda_1(f) \int_G \psi_\varepsilon(x^{-1})d\mu_2(x)| \leq C\varepsilon$$

and

$$|\lambda_2(g) - \lambda_1(g) \int_G \psi_\varepsilon(x^{-1}) d\mu_2(x)| \leq C\varepsilon.$$

Note that any of these two relations implies that there exist constants $0 < A < B$ such that $A \leq \int_G \psi_\varepsilon(x^{-1}) d\mu_2(x) \leq B$ for ε small enough. In particular, we get

$$\lambda_1(f)\lambda_2(g) = \lambda_2(f)\lambda_1(g) + O(\varepsilon).$$

Making ε tend to 0, we get $\lambda_1(f)\lambda_2(g) = \lambda_2(f)\lambda_1(g)$, as desired.

□

VII.5.3 Unimodular groups

We use the notation and definitions of problem VII.5.2. Let G be a locally compact Hausdorff topological group, let μ be a left Haar measure on G. We also admit the fact that left Haar measure on G are unique up to mutliplication by a scalar (if G is a normal topological space, this was proved in problem VII.5.2.)

(1). Show that there exists a function $c : G \to \mathbb{R}_{>0}$ such that, for every $g \in G$ and $f \in \mathscr{C}(G, \mathbb{C})$, $\int_G f(xg) d\mu(x) = c(g) \int_G f(x) d\mu(x)$.

(2). Show that $c : G \to (\mathbb{R}_{>0}, \times)$ is a continuous morphism of groups.

The function c is called the *modular function* of G, and we say that G is *unimodular* if $c(G) = 1$ (i.e. if μ is also a right Haar measure).

(3). If G is compact, show that it is unimodular.

(4). Let $f \in \mathscr{C}(G, \mathbb{C})$. Show that $\int_G f(x^{-1})\, d\mu(x) = \int_G c(x) f(x) d\mu(x)$. [14]

(Hint : Take another function $g \in \mathscr{C}(G, \mathbb{C})$ and calculate $\int_G \int_G g(yx) c(x)^{-1} f(x^{-1}) d\mu(x) d\mu(y)$ in two different ways.)

(5). Let G be the topological group $\mathbb{R}_{>0} \times \mathbb{R}$, with the multiplication given by $(a, b)(a', b') = (aa', ab' + b)$. Find a left Haar measure on G. Is G unimodular ?

Solution.

(1). Let $g \in G$. Then $d\mu(x)$ and $d\mu(xg^{-1})$ are both left Haar measures on G, so, by problem VII.5.2, there exists $c(g) \in \mathbb{R}_{>0}$ such that $d\mu(xg^{-1}) = c(g) d\mu(x)$, i.e., for every $f \in \mathscr{C}(G, \mathbb{C})$, $\int_G f(xg) d\mu(x) = \int_G f(y) d\mu(yg^{-1}) = c(g) \int_G f(x) d\mu(x)$.

[14] Cette formule est-elle juste ?

(2). By problem VII.5.2, all left Haar measures on G are proportional, and so we would have gotten the same function c in question (1) if we had used another left Haar measure on G to define it. In particular, for all $g_1, g_2 \in G$,

$$c(g_1g_2)d\mu(x) = d\mu(x(g_1g_2)^{-1}) = d\mu(xg_2^{-1}g_1^{-1}) = c(g_1)d\mu(xg_2^{-1}) = c(g_1)c(g_2)d\mu(x),$$

so $c(g_1g_2) = c(g_1)c(g_2)$. Hence $c : G \to (\mathbb{R}_{>0}, \times)$ is a morphism of groups.

Now we show that c is continuous. Let $\varepsilon > 0$. Choose $f \in \mathscr{C}(G, \mathbb{R}_{\geq 0})$ such that $\int_G f(x)d\mu(x) = 1$. Let U be a neighborhood of 1 in G such that, for every $x \in G$ and every $y \in U$, $|f(x) - f(xy)| \leq \varepsilon$, and let K be the support of f. Then, if $g, g' \in G$ are such that $g^{-1}g' \in U$, we have

$$|c(g')-c(g)| = |(c(g')-c(g)) \int_G f(x)d\mu(x)| = | \int_G f(xg')d\mu(x) - \int_G f(xg)d\mu(x)| \leq \varepsilon\mu(K).$$

(3). If G is compact, then $c(G)$ is a compact subgroup of $\mathbb{R}_{>0}$. The only compact subgroup of $\mathbb{R}_{>0}$ is $\{1\}$, so $c(G) = \{1\}$, i.e., G is unimodular.

(4). Let $g \in \mathscr{C}(G, \mathbb{C})$. Then we have

$$\int_G \int_G g(yx)c(x)^{-1}f(x^{-1})d\mu(x)d\mu(y) = \int_G f(x^{-1})c(x)^{-1}(\int_G g(yx)d\mu(y))d\mu(x).$$

By definition of c, $c(x)^{-1} \int_G g(yx)d\mu(y) = \int_G g(y)d\mu(y)$, and so

$$\int_G \int_G g(yx)c(x)^{-1}f(x^{-1})d\mu(x)d\mu(y) = \int_G f(x^{-1})d\mu(x) \int_G g(y)d\mu(y).$$

On the other, the left invariance of μ and the change of variables $z = yx$ give

$$\int_G \int_G g(yx)c(x)^{-1}f(x^{-1})d\mu(x)d\mu(y) = \int_G \int_G g(z)c(y^{-1}z)^{-1}f(z^{-1}y)d\mu(z)d\mu(y).$$

Using the left invariance of μ again and the change of variables $t = z^{-1}y$, this becomes equal to

$$\int_G g(z)d\mu(z) \int_G c(t)f(t)d\mu(t).$$

Choosing a function g such that $\int_G g(x)d\mu(x) = 1$, we get the result.

(5). Note that G is isomorphic to the closed subgroup $\{\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}, a \in \mathbb{R}_{>0}, b \in \mathbb{R}\}$ of $\mathrm{GL}_2(\mathbb{R})$. Let dx be the Lebesgue measure on $\mathbb{R}$. Then it is very easy to check that $d\mu(a, b) := \frac{da}{a^2}db$ is a left Haar measure on G. It's not a right Haar measure, because $d\mu((a, b)(x, y)) = x^{-1}d\mu(a, b)$, so the modular function c of G is given by $c(x, y) = x$.

□

VII.5.4 Some examples of topological groups

Among the following closed subgroups of $\mathrm{GL}_n(\mathbb{C})$, which ones are connected ? Which ones are compact ?

(1). $\mathrm{GL}_n(\mathbb{C})$

(2). $\mathrm{SL}_n(\mathbb{C}) := \{A \in M_n(\mathbb{R}) | \det(A) = 1\}$

(3). $\mathrm{GL}_n(\mathbb{R})$

(4). $\mathrm{SL}_n(\mathbb{R}) := \{A \in M_n(\mathbb{R}) | \det(A) = 1\}$

(5). $\mathrm{O}(q) := \{A \in M_n(\mathbb{R}) | {}^tAqA = q\}$, where $q \in M_n(\mathbb{R})$ is an invertible symmetric matrix (corresponding to a non-degenerate quadratic form on $\mathbb{R}^n$)

Warning : I am not assuming that the quadratic form is positive definite. Feel free to use the fact that any non-degenerate quadratic form on $\mathbb{R}^n$ is equivalent to a form of the type $q(x_1, \ldots, x_n) = x_1^2 + \cdots + x_r^2 - x_{r+1}^2 - \cdots - x_n^2$ (the pair of integers $(r, n-r)$ is called the *signature* of the form).

(6). $\mathrm{SO}(q) := \{A \in \mathrm{O}(q) | \det(A) = 1\}$

(7). $\mathrm{O}(n, \mathbb{C}) := \{A \in M_n(\mathbb{C}) | {}^tAA = I_n\}$

(8). $\mathrm{SO}(n, \mathbb{C}) := \{A \in \mathrm{O}(n, \mathbb{C}) | \det(A) = 1\}$

(9). $\mathrm{U}(q) := \{A \in M_n(\mathbb{C}) | A^*qA = q\}$, where $q \in M_n(\mathbb{C})$ is an invertible Hermitian matrix (corresponding to a non-degenerate Hermitian form on $\mathbb{C}^n$)

Warning : I am not assuming that the Hermitian form is positive definite. Feel free to use the fact that any non-degenerate Hermitian form on $\mathbb{C}^n$ is equivalent to a form of the type $q(z_1, \ldots, z_n) = |z_1|^2 + \cdots + |z_r|^2 - |z_{r+1}|^2 - \cdots - |z_n|^2$ (the pair of integers $(r, n-r)$ is called the *signature* of the form).

(10). $\mathrm{SU}(q) := \{A \in \mathrm{U}(q) | \det(A) = 1\}$

(11). $\mathrm{Sp}(q, \mathbb{R}) := \{A \in M_n(\mathbb{R}) | {}^tAqA = q\}$, where $q \in M_n(\mathbb{R})$ is an invertible skew-symmetric matrix (corresponding to a symplectic form on $\mathbb{R}^n$)

Hint : First, show that n has to be even and that you can take the matrix q to be $J_n := \begin{pmatrix} 0 & I_{n/2} \\ -I_{n/2} & 0 \end{pmatrix}$, where $I_{n/2}$ is the identity matrix in $GL_{n/2}(\mathbb{R})$. We write $\mathrm{Sp}_n(\mathbb{R}) = Sp(J_n, \mathbb{R})$ and denote by $\langle ., . \rangle$ the symplectic form on $\mathbb{R}^n$ defined by J_n (ie $\langle x, y \rangle = {}^txJ_ny$).

Then there are (at least) three ways to proceed to prove that $\mathrm{Sp}_n(\mathbb{R})$ is connected :

(a) Argument by induction : Let $Z = \mathbb{R}^n - \{0\}$. Consider the obvious action of $\mathrm{Sp}_n(\mathbb{R})$ on $\mathbb{R}^n$.

(i). Show that $\mathrm{Sp}_n(\mathbb{R})$ preserves Z and acts transitively on Z.

(ii). Show that the stabilizer in $\mathrm{Sp}_n(\mathbb{R})$ of $(1, 0, \ldots, 0) \in Z$ is connected.

(iii). Show that Z is connected and conclude.

(b) Argument by Iwasawa decomposition : Let $K = \mathrm{Sp}_n(\mathbb{R}) \cap \mathrm{O}(n, \mathbb{R})$. Let L be the set of $\mathbb{R}$-vector subspaces V of $\mathbb{R}^n$ such that $\dim V = n/2$ and that the restriction of $\langle ., . \rangle$ to V is zero (such a subspace is called a maximal totally isotropic subspace, or a Lagrangian subspace).

(i). Make $\mathrm{Sp}_n(\mathbb{R})$ act on L by $(g, V) \mapsto gV$ (the image of V by the linear transformation g). Show that this action is well defined, and that K acts transitively on L.

(ii). Show that every matrix in K is of the form $\begin{pmatrix} M & -N \\ N & M \end{pmatrix}$ (blocks of size $n/2 \times n/2$), with some condition on M and N, and that the map $\varphi : K \to U(n)$ that sends $\begin{pmatrix} M & -N \\ N & M \end{pmatrix}$ to $M + iN$ is an isomorphism.

(iii). Calculate the stabilizer P of $V_0 = \mathbb{R}^{n/2} \oplus 0 \in L$ (in $\mathrm{Sp}_n(\mathbb{R})$). Is P connected ?

(iv). Show that $\mathrm{Sp}_n(\mathbb{R}) = KP^0$ (ie, every element of $\mathrm{Sp}_n(\mathbb{R})$ is the product of an element of K and an element of P^0), where P^0 is the connected component of the unit element in P.

(v). Conclude.

(The fact that $\mathrm{Sp}_n\mathbb{R}) = KP$ is a particular case of the Iwasawa decomposition.)

(c) Argument by symplectic polar decomposition : Let $K = \mathrm{Sp}_n(\mathbb{R}) \cap \mathrm{O}(n, \mathbb{R})$ and let S be the set of elements of $\mathrm{Sp}_n(\mathbb{R})$ that are symmetric positive definite.

(i). Show that every matrix in K is of the form $\begin{pmatrix} M & -N \\ N & M \end{pmatrix}$ (blocks of size $n/2 \times n/2$), with some condition on M and N, and that the map $\varphi : K \to U(n)$ that sends $\begin{pmatrix} M & -N \\ N & M \end{pmatrix}$ to $M + iN$ is an isomorphism.

(ii). Show that S is connected.

(iii). Show that every element g of $\mathrm{Sp}_n(\mathbb{R})$ can be written in a unique way as $g = us$, with $u \in K$ and $s \in S$.

(iv). Conclude.

(12). $\mathrm{Sp}(q, \mathbb{C}) := \{A \in M_n(\mathbb{C}) | {}^tAqA = q\}$, where $q \in M_n(\mathbb{C})$ is an invertible skew-symmetric matrix (corresponding to a symplectic form on $\mathbb{C}^n$)

Hint : Try to adapt one of the methods for $Sp(q, \mathbb{R})$.

(13). Why am not adding $\mathrm{SSp}(q,\mathbb{R}) := \{A \in \mathrm{Sp}(q,\mathbb{R}) | \det(A) = 1\}$ and $\mathrm{SSp}(q,\mathbb{C})$ to the list ?

Solution.

(1). Let's show that $\mathrm{GL}_n(\mathbb{C})$ is connected and not compact.

By $g \in \mathrm{GL}_n(\mathbb{C})$. By Jordan reduction, there exists $h \in \mathrm{GL}_n(\mathbb{C})$ such that $hgh^{-1} = D+N$, where

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

and

$$N = \begin{pmatrix} 0 & a_1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & a_{n-1} \\ 0 & & & 0 \end{pmatrix},$$

with $\lambda_1, \ldots, \lambda_n \in \mathbb{C}^\times$ and $a_1, \ldots, a_{n-1} \in \{0,1\}$. Write $\lambda_r = \alpha_r e^{i\theta_r}$, with $\alpha_r \in \mathbb{R}_{>0}$ and $\theta_r \in \mathbb{R}$. If $t \in \mathbb{R}$, we set

$$\lambda_r(t) = e^{t\log(\alpha_r)} e^{it\theta_r},$$

$$D = \begin{pmatrix} \lambda_1(t) & & 0 \\ & \ddots & \\ 0 & & \lambda_n(t) \end{pmatrix},$$

$$N = \begin{pmatrix} 0 & ta_1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & ta_{n-1} \\ 0 & & & 0 \end{pmatrix},$$

and $g(t) = h^{-1}(D(t) + N(t))h$. Then $t \longmapsto g(t)$ is a continuous function from $\mathbb{R}$ to $\mathrm{GL}_n(\mathbb{C})$, $g(t) \in \mathrm{GL}_n(\mathbb{C})$ for every $t \in \mathbb{R}$, $g(1) = g$ and $g(0) = I_n$. So $\mathrm{GL}_n(\mathbb{C})$ is connected (and even path-connected). Note also that if $g \in \mathrm{SL}_n(\mathbb{C})$, then $\alpha_1 \ldots \alpha_n = 1$ and $\theta_1 + \cdots + \theta_n \in 2\pi\mathbb{Z}$. After modifying θ_n by an element of $2\pi\mathbb{Z}$, we can assume that $\theta_1 + \cdots + \theta_n = 0$. Then, for every $t \in \mathbb{R}$,

$$\det(g(t)) = e^{t\log(\alpha_1 \ldots \alpha_n)} e^{it(\theta_1 + \cdots + \theta_n)} = 1,$$

that is, $g(t) \in \mathrm{SL}_n(\mathbb{C})$. So the same proof shows that $\mathrm{SL}_n(\mathbb{C})$ is connected.

On the other hand, we have a continuous surjective function $\det : \mathrm{GL}_n(\mathbb{C}) \to \mathbb{C}^\times$, and $\mathbb{C}^\times$ is not compact, so $\mathrm{GL}_n(\mathbb{C})$ cannot be compact.

(2). We already saw in the proof of (1) that $\mathrm{SL}_n(\mathbb{C})$ is connected. Let's show that $\mathrm{SL}_n(\mathbb{C})$ is not compact. For example, we can consider the set U of matrices of the form

$$\begin{pmatrix} 1 & & a \\ & \ddots & \\ 0 & & 1 \end{pmatrix},$$

with $a \in \mathbb{C}$ (all the entries except the ones on the diagonal and the upper left-hand one are zero). This is a closed subset (actually a subgroup) of $\mathrm{SL}_n(\mathbb{C})$, and it is homemorphic (actually isomorphic as a topological group) to $\mathbb{C}$, which is not compact. So $\mathrm{SL}_n(\mathbb{C})$ cannot be compact.

(3). The map $\det : \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^\times$ is continuous and surjective. As $\mathbb{R}^\times$ is neither connected nor compact, $\mathrm{GL}_n(\mathbb{R})$ is neither connected nor compact.

(4). Let's show that $\mathrm{SL}_n(\mathbb{R})$ is connected and not compact.

We can prove that it is not compact just as in the case of $\mathrm{SL}_n(\mathbb{C})$: if U is the closed subset of $\mathrm{SL}_n(\mathbb{C})$ defined in the proof of (2), then $U \cap \mathrm{SL}_n(\mathbb{R}) \simeq \mathbb{R}$, which is not compact, so $\mathrm{SL}_n(\mathbb{R})$ is not compact.

To prove that $\mathrm{SL}_n(\mathbb{R})$ is connected, we can for example use the fact that it is generated by the transvections (a.k.a. shear transformations). Remember that, for any field K, a transvection in $M_n(K)$ is a matrix of the form $I_n + A$, where $\mathrm{rk}(A) \leq 1$. Such a matrix is automatically in $\mathrm{SL}_n(K)$, and $\mathrm{SL}_n(K)$ is generated by matrices of this form. [15] Now let's take $g \in \mathrm{SL}_n(\mathbb{R})$, and write $g = (I_n + A_1) \dots (I_n + A_r)$, where $\mathrm{rk}(A_i) \leq 1$ for every $i \in \{1, \dots, r\}$. For every $t \in \mathbb{R}$, let $g(t) = (I_n + tA_1) \dots (I_n + tA_r)$. Then $t \longmapsto g(t)$ is a continuous map from $\mathbb{R}$ to $\mathrm{SL}_n(\mathbb{R})$, and we have $g(1) = g$ and $g(0) = I_n$. So $\mathrm{SL}_n(\mathbb{R})$ is connected.

(5). Let $(r, n-r)$ be the signature of q. Then there exists a matrix $g \in \mathrm{GL}_n(\mathbb{R})$ such that $q = {}^t g I_{r,n-r} g$, where $I_{r,n-r}$ is the diagonal matrix whose first r diagonal entries equal to 1 and whose last $n-r$ diagonal entries are equal to -1. If $A \in \mathrm{GL}_n(\mathbb{R})$, we see easily that $A \in \mathrm{O}(q)$ if and only if $gAg^{-1} \in \mathrm{O}(I_{r,n-r})$. So the map $A \longmapsto gAg^{-1}$ is a homemorphism from $\mathrm{O}(q)$ to $\mathrm{O}(I_{r,n-r})$, and so we may assume that $q = I_{r,n-r}$.

Fix $r \in \{0, \dots, n\}$ and write G_r for $\mathrm{O}(I_{r,n-r})$. First we show that G_r is not connected. Consider the continuous map $\det : G_r \to \mathbb{R}^\times$. If $A \in G_r$, then $\det({}^t A I_{r,n-r} A) = \det(I_{r,n-r})$, so $\det(A)^2 = 1$, so $\det(A) \in \{\pm 1\}$. On the other hand,

$$A := \begin{pmatrix} I_{n-1} & 0 \\ 0 & -1 \end{pmatrix}$$

is in G_r and $\det(A) = -1$. So $\det : G_r \to \{\pm 1\}$ is a surjective continuous map. As $\{\pm 1\}$ is not connected, G_r cannot be connected.

[15] ref ?

If $1 \leq r \leq n-1$, let's show that G_r is not compact. Consider the subset X of G_r of matrices of the form

$$\begin{pmatrix} a & 0 & 0 & b \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ c & 0 & 0 & d \end{pmatrix},$$

where $a, b, c, d \in \mathbb{R}$ and all the diagonal terms except for the first and last one are equal to 1. This is a closed subset, and an easy calculation shows that every element of X is of the form

$$\begin{pmatrix} \sqrt{1+t^2} & 0 & 0 & t \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ t & 0 & 0 & \sqrt{1+t^2} \end{pmatrix},$$

for a uniquely determined $t \in \mathbb{R}^\times$. So X is homemorphic to $\mathbb{R}^\times$. As it is closed in G_r and $\mathbb{R}^\times$ is not compact, G_r cannot be compact.

If $r = 0$, then $I_{r,n-r} = -I_n$, so $\mathrm{O}(I_{r,n-r}) = \mathrm{O}(I_n)$. This group is usually denoted by $\mathrm{O}(n)$. Let's show that it is compact. Let $A \in M_n(\mathbb{R})$, and let $v_1, \ldots, v_n$ be the column vectors of A. Then $A \in \mathrm{O}(n)$ if and only of $(v_1, \ldots, v_n)$ is an orthonormal basis of $\mathbb{R}^n$. Let $S^n = \{v \in \mathbb{R}^n | \|v\| = 1\}$, this is the unit sphere in $\mathbb{R}^n$ and it is compact (because it's closed and bounded). The set of orthonormal bases $(v_1, \ldots, v_n)$ of $\mathbb{R}^n$ is a closed subset of $(S^n)^n$, so we get a homemorphism between $\mathrm{O}(n)$ and a closed subset of $(S^n)^n$, and this implies that $\mathrm{O}(n)$ is compact.

(6). As in (5), we may assume that $q = I_{r,n-r}$, with $0 \leq r \leq n-r$. If $r \in \{1, \ldots, n-r\}$, then the closed subset X of $\mathrm{O}(I_{r,n-r})$ constructed in (5) is actually contained in $\mathrm{SO}(I_{r,n-r})$. As X is homemorphic to $\mathbb{R}^\times$, which is not compact, $\mathrm{SO}(I_{r,n-r})$ is not compact.

If $r = 0$ or $r = n$, then $\mathrm{SO}(I_{r,n-r}) = \mathrm{SO}(n) := \{A \in \mathrm{O}(n) | \det(A) = 1\}$. This is a closed subgroup of the compact group $\mathrm{O}(n)$ (cf. (5)), so it is compact.

Now let's show that $\mathrm{SO}(q)$ is connected for every non-degenerate q. First we note the following two lemmas.

Lemma. *Let q be any non-degenerate quadratic form on $\mathbb{R}^n$, and let $(v_1, \ldots, v_r)$ be an orthogonal family in $\mathbb{R}^n$ such that $q(v_i) \in \{\pm 1\}$ for every $i \in \{1, \ldots, r\}$. Then $(v_1, \ldots, v_r)$ can be extended to an orthogonal basis $(v_1, \ldots, v_n)$ of $\mathbb{R}^n$ such that $q(v_i) \in \{\pm 1\}$ for every $i \in \{1, \ldots, n\}$.*

Proof. Denote by $\langle ., . \rangle$ the symmetric bilinear form corresponding to q. Let $W = \mathrm{Span}(v_1, \ldots, v_r)$. We prove the lemma by induction on $n - \dim W = n - r$. If $n = r$, we're done, so let's assume that $r < n$. Then $W^\perp \neq \{0\}$, and $q_{|W^\perp}$ is a non-degenerate quadratic form, so there exists $v_{r+1} \in W^\perp$ such that $q(v_{r+1}) \neq 0$. After multiplying v_{r+1} by a scalar, we may assume that $q(v_{r+1}) = \pm 1$. Now we just have to

apply the induction hypothesis to $(v_1, \dots, v_{r+1})$.

□

Lemma. *Let X and Y be topological spaces, and let $\pi : X \to Y$ be a continuous surjective map. Suppose that :*

(a) For every $y \in Y$, $\pi^{-1}(y)$ is path-connected.

(b) For every $y \in Y$, there exists an open neighborhood U of y in Y and a continuous map $s : U \to X$ such that $\pi \circ s = \mathrm{id}_U$.

Then X is path-connected if and only if Y is path-connected.

Proof. If X is path-connected, then Y is obviously path-connected.

Conversely, assume that Y is path-connected. Let $x_1, x_2 \in X$, and write $y_1 = \pi(x_1)$, $y_2 = \pi(x_2)$. By (a), there exists a continuous map $\gamma : [0,1] \to Y$ such that $\gamma(0) = y_1$ and $\gamma(1) = y_2$. Using (c) and the compactness of $\gamma([0,1])$, we get a sequence $0 = 1_0 < a_1 \dots a_n = 1$, open subsets $U_1, \dots, U_n$ and continuous maps $s_i : U_i \to X$ such that $\pi \circ s_i = \mathrm{id}_{U_i}$ and $\gamma([a_{i-1}, a_i]) \subset U_i$ for every $i \in \{1, \dots, n\}$. For every $i \in \{1, \dots, n\}$, let $\delta_i : [a_{i-1}, a_i] \to X$ be $s_i \circ \gamma_{|[a_{i-1}, a_i]}$. This is a continuous path on X connecting $s_i(\gamma(a_{i-1}))$ and $s_i(\gamma(a_i))$. Also, by condition (b), we can find a continuous paths connecting x_1 and $s_1(\gamma(0))$, x_2 and $s_n(\gamma(1))$, and $s_i(\gamma(a_i))$ and $s_{i+1}(\gamma(a_i))$ for every $i \in \{1, \dots, n-1\}$. So we have connected x_1 and x_2 by a continuous path.

□

Now we come back to the problem. Let q be a non-degenerate quadratic form on $\mathbb{R}^n$, we want to show that $\mathrm{SO}(q)$ is path-connected, unless $n = 2$ and the signature of the form is $(1,1)$. We proceed by induction on n. If $n = 1$, then $\mathrm{SO}(q) = \{1\}$.

Suppose that $n = 2$. If $r = 2$ or $r = 0$, then $SO(q) \simeq \mathrm{SO}(2)$, so $\mathrm{SO}(q)$ is homeomorphic to the unit circle in $\mathbb{R}^2$, and this is path-connected. If $r = n - r = 1$, then an easy calculation show that

$$\mathrm{SO}(I_{1,1}) = \{ \begin{pmatrix} a & b \\ b & a \end{pmatrix} | a^2 - b^2 = 1 \},$$

so $\mathrm{SO}(I_{1,1})$ has two connected components.

Assume that $n \geq 3$. After replacing q by $-q$, we may assume that there exists a $v_0 \in \mathbb{R}^n$ such that $q(v_0) = 1$. Let $\langle ., . \rangle$ be the bilinear symmetric form attached to q, and let $S = \{v \in V | q(v) = 1\}$. We consider the continous map $\pi : \mathrm{SO}(q) \to S$, $A \longmapsto Av_0$. We'll show that $\mathrm{SO}(q)$ is path-connected by checking the conditions of the second lemma :

- Let's show that π is surjective. Let $v_1 \in S$. By the first lemma, there exists orthonomal bases $(w_1 = v_0, \dots, w_n)$ and $(v_1, \dots, v_n)$ such that $q(v_i), q(w_i) \in \{\pm 1\}$ for every i. The number of 1's and -1's among $(q(w_1), \dots, q(w_n))$ and $(q(v_1), \dots, q(v_n))$ must be the same (it's the signature of q), and $q(w_1) = q(v_1)$, so after changing the

order of $v_2, \ldots, v_n$ we may assume that $q(v_i) = q(w_i)$ for every $i \in \{1, \ldots, n\}$. This means that the unique $A \in \mathrm{GL}_n(\mathbb{R})$ sending $(w_1, \ldots, w_n)$ to $(v_1, \ldots, v_n)$ is in $\mathrm{O}(q)$. After replacing w_n by $-w_n$ (which changes the sign of $\det(A)$), we may assume that $A \in \mathrm{SO}(q)$. So we have found $A \in \mathrm{SO}(q)$ such that $Av_0 = v_1$, i.e., $\pi(A) = v_1$.

- Let's show that S is path-connected. We may assume that q is given by the matrix $I_{r,n-r}$. As $I_{r,n-r}$ and $I_{n-r,r}$ give rise to isomorphic groups, we may assume that $r \geq n/2$. Then

$$S = \{(x_1, \ldots, x_n) \in \mathbb{R}^n | x_1^2 + \cdots + x_r^2 - x_{r+1}^2 - \cdots - x_n^2 = 1\},$$

and the vectors $e_1 := (1, 0, \ldots, 0)$, $-e_1 = (-1, 0, \ldots, 0)$ are in S. Let's show how to connected every point of S to e_1 or $-e_1$. Let $p = (x_1, \ldots, x_n)$. If $x_1 = 1$, then $x_2^2 + \cdots + x_r^2 - x_{r+1}^2 - \cdots - x_n^2 = 0$. Consider the continuous map $p : [0, 1] \to \mathbb{R}^n$ sending t to $p(t) := (1, tx_2, \ldots, tx_n)$. By the observation just made, $p(t) \in S$ for every t, and we also have $p(0) = e_1$, $p(1) = p$, so we are done. Suppose that $x_1 \geq 0$ and $x_1 \neq 1$. For every $t \in [0, 1]$, set $x_1(t) = 1 - t + tx_1$. Then, if $t > 0$, $1 - x_1(t)^2$ is nonzero and has the same sign as $1 - x_1^2$. Consider the continuous map $p : [0, 1] \to \mathbb{R}^n$ sending t to

$$(x_1(t), \sqrt{\frac{1 - x_1(t)^2}{1 - x_1^2}} x_2, \ldots, \sqrt{\frac{1 - x_1(t)^2}{1 - x_1^2}} x_n).$$

We check easily that $p(t) \in S$ for every t, and we have $p(0) = e_1$, $p(1) = p$, so we have connected p and e_1. Similarly, if $x_1 \leq 0$, then we can connected p and $-e_1$. To finish the proof, we just need to find a continuous path on S between e_1 and $-e_1$. As $n \geq 3$, we have $r \geq 2$, so we can use the path $p : [0, 1] \to S$ sending t to

$$((1 - 2t), \sqrt{1 - (1 - 2t)^2}, 0, \ldots, 0),$$

which makes sense because $(1 - 2t)^2 \leq 1$ if $0 \leq t \leq 1$.

- Let's show that the fibers of π are path-connected. We may assume that $q = I_{r,n-r}$, that $r \geq 1$, and that $r \neq 2$ if $n = 3$ (if $(r, n - r) = (2, 1)$, we just switch r and $n - r$, which doesn't change the group up to isomorphism). Let $e_1 = (1, 0, \ldots, 0)$, we have $e_1 \in S$ because $r \geq 1$. Let $G' = \{A \in \mathrm{SO}(q) | Ae_1 = e_1\}$ (this is a closed subgroup of $\mathrm{SO}(q)$). Let $v_1 \in S$. We have seen that π is surjective, so there exist $g, h \in \mathrm{SO}(q)$ such that $v_1 = gv_0$ and $e_1 = hv_0$. Then

$$\pi^{-1}(v_1) = \{A \in \mathrm{SO}(q) | Av_0 = v_1\} = \{A \in \mathrm{SO}(q) | Ah^{-1}e_1 = gh^{-1}e_1\} = gh^{-1}G'h,$$

so it suffices to show that G' is path-connected. But G' is isomorphic to $\mathrm{SO}(I_{r-1,n-r})$, so this follows from the induction hypothesis.

- Let's show condition (b) of the lemma, i.e. the fact that π admits a continuous section locally on S. We may assume that $q = I_{r,n-r}$ with $r \geq 1$, so that $e_1 := (1, 0, \ldots, 0)$ is in S. Let $v_1 \in S$, let $A \in \mathrm{SO}(q)$ such that $Ae_1 = v_1$ (this exists by the surjectivity

of π), and let $v_1, \ldots, v_n$ be the columns of A. Then $(v_1, \ldots, v_n)$ is an orthogonal basis $(v_1, \ldots, v_n)$ of $\mathbb{R}^n$, and we have $q(v_i) = \pm 1$ for every $i \in \{1, \ldots, n\}$. Suppose that $v_1' \in S$, if v_1' is close enough to v_1 then $(v_1', v_2, \ldots, v_n)$ is still a basis of $\mathbb{R}^n$ (because the determinant of the matrix with columns $v_1', v_2, \ldots, v_n$ will be close to $\det(A) \neq 0$), and we want to apply the Gram-Schmidt process to this basis. Of course, this is not always possible, because q is not definite. Let's ignore this problem for now. The Gram-Schmidt process gives an orthogonal basis $(v_1', v_2', \ldots, v_n')$ by the inductive formula

$$v_i' = v_i - \sum_{j=1}^{i-1} \frac{\langle v_j', v_i \rangle}{\langle v_j', v_j' \rangle} v_i,$$

for $2 \leq i \leq n$. The elements $v_2', \ldots, v_n'$, if they make sense, vary continuously with v_1' and equal $v_2, \ldots, v_n$ if $v_1' = v_1$. So, for v_1' close enough to v_1, we will have $\langle v_1', v_1' \rangle \neq 0$ and then v_2' will make sense, and then $\langle v_2', v_2' \rangle$ will be close to $\langle v_2, v_2 \rangle$, hence nonzero, and then v_3' will make sense, etc. So there exists a neighborhood U of v_1 in S such that the Gram-Schmidt process will work for $v_1' \in U$, and will produce an orthogonal basis $(v_1', v_2', \ldots, v_n')$ with $q(v_i') \neq 0$ for every i. After shrinking U, we may also assume that, for $v_1' \in U$, $q(v_i)$ and $q(v_i')$ have the same sign for every i, and so the matrix $B(v_1')$ with columns $\frac{1}{q(v_1')} v_1', \ldots, \frac{1}{q(v_n')} v_n'$ is in $\mathrm{O}(q)$. Also, $\det(B(v_1'))$ is a continuous function of v_1' and can only take the values 1 and -1, so, after shrinking U, we may assume that $\det(B(v_1')) = 1$ (i.e. $B(v_1') \in \mathrm{SO}(q)$) for every $v_1' \in U$. We have $B(v_1')e_1 = \frac{1}{q(v_1')} v_1' = v_1'$. Choose $g \in \mathrm{SO}(q)$ such that $v_0 = ge_1$ (this is possible by the surjectivity of π). Then $v_1' = B(v_1')e_1 = B(v_1')gv_0$, so the fonction $s : U \to \mathrm{SO}(q)$, $v_1' \longmapsto B(v_1')g$, is continuous and satisfies $\pi \circ s = \mathrm{id}_U$.

(7). Note that all the non-degenerate quadratic forms on $\mathbb{C}^n$ are equivalent, so all the associated orthogonal groups are isomorphic to $\mathrm{O}(n, \mathbb{C})$. This is why we don't vary the form as in (5).

Just as in (5), we see that $\det : \mathrm{O}(n, \mathbb{C}) \to \{\pm 1\}$ is a surjective continuous map, so $\mathrm{O}(n, \mathbb{C})$ is not connected.

If $n = 1$, then $\mathrm{O}(n, \mathbb{C}) = \{\pm 1\}$ is compact. Suppose that $n \geq 2$, and choose $r \in \{1, \ldots, n-1\}$. Then the quadratic forms on $\mathbb{C}^n$ given by I_n and $I_{r,n-r}$ are equivalent, so $\mathrm{O}(n, \mathbb{C})$ is isomorphic to

$$G := \{A \in M_n(\mathbb{C}) | {}^A I_{r,n-r} A = I_{r,n-r}\}.$$

But the closed subset $G \cap M_n(\mathbb{R})$ of G is clearly equal to $\mathrm{O}(I_{r,n-r})$, and we have seen in (5) that this is not compact, so $\mathrm{O}(n, \mathbb{C})$ is not compact.

(8). If $n = 1$, then $\mathrm{SO}(n, \mathbb{C}) = \{1\}$ is compact and connected. If $n \geq 2$, we show that $\mathrm{SO}(n, \mathbb{C})$ is not compact as in (7) (this time using the fact, proved in (6), that $\mathrm{SO}(I_{r,n-r})$ is not compact for $1 \leq r \leq n-1$).

Let's show that $\mathrm{SO}(n,\mathbb{C})$ is always connected. We use the same method as in (6). First note the following lemma, analogous to the first lemma of (6) :
Lemma. *Let q be any non-degenerate quadratic form on $\mathbb{C}^n$, and let $(v_1,\dots,v_r)$ be an orthonormal family in $\mathbb{C}^n$ Then $(v_1,\dots,v_r)$ can be extended to an orthonormal basis $(v_1,\dots,v_n)$ of $\mathbb{C}^n$.*

The proof is exactly the same as in (6), the only difference being that, if $v\in\mathbb{C}^n$ is such that $q(v)\neq 0$, then we can always find a $\lambda\in\mathbb{C}$ such that $q(\lambda v)=1$. So we can normalize any orthogonal basis to make it an orthonormal basis, as long as the basis vectors are not in the set $\{v\in\mathbb{C}^n | q(v)=0\}$.

Now let's prove that $\mathrm{SO}(n,\mathbb{C})$ is connected by induction on n. We know the case $n=1$, so suppose that $n\geq 2$. Let

$$S=\{(z_1,\dots,z_n)\in\mathbb{C}^n | z_1^2+\cdots+z_n^2=1\},$$

let $e_1=(1,0,\dots,0)$, and define a continuous map $\pi:\mathrm{SO}(n,\mathbb{C})\to S$ by $\pi(A)=Ae_1$. We want to check the conditions of the second lemma of (6). The proofs are similar but simpler.

- Let's show that π is surjective. If $v_1\in S$, the lemma says that it is possible to complete it to an orthonomal basis $(v_1,\dots,v_n)$ of $\mathbb{C}^n$. If A is the matrix with columns $v_1,\dots,v_n$, then $A\in\mathrm{SO}(n,\mathbb{C})$ and $\pi(A)=v_1$.
- Let's show that S is path-connected. Let $(x_1,\dots,x_n)\in S$. Choose a continuous function $t\longmapsto x_1(t)$ from $[0,1]$ to $\mathbb{C}$ such that, for every $t\in[0,1]$, $x_1(t)^2=(1-t^2)+t^2x_1^2$. Then $1-x_1(t)^2=t^2(1-x_1^2)=t^2(x_2^+\cdots+x_n^2)$, so the continuous path $p:[0,1]\to\mathbb{C}^n$, $t\longmapsto(x_1(t),tx_2,\dots,tx_n)$, has image in S, and it connects $(x_1,\dots,x_n)$ and $(1,0,\dots,0)$.
- Let's show that the fibers of π are path-connected. As in (6), using the surjectivity of π, we see that all the fibers are homemorphic to $\pi^{-1}(e_1)\simeq\mathrm{SO}(n-1,\mathbb{C})$, so we can apply the induction hypothesis.
- Let's show that π admits a continuous section locally on S. Let $v_1\in S$. The lemma gives an orthonomal basis $(v_1,\dots,v_n)$ of $\mathbb{C}^n$. Just as in (6), we can find a neighborhood U of v_1 in S such that, if $v_1'\in U$, then $(v_1',v_2,\dots,v_n)$ is still a basis of $\mathbb{C}^n$ and applying the Gram-Schmidt process to it will make sense and produce an orthonomal basis $(v_1',\dots,v_n')$. Denote by $B(v_1')$ the matrix with columns $v_1',\dots,v_n'$, then $B(v_1')e_1=v_1'$ and, after shrinking U (again, just like in (6)), we get $B(v_1')\in\mathrm{SO}(n,\mathbb{C})$ for every $v_1'\in U$. Then the function $s:U\to\mathrm{SO}(n,\mathbb{C})$, $v_1'\longmapsto B(v_1')$, is continuous and satisfies $\pi\circ s=\mathrm{id}_U$.

(9). This will be very similar to (5) and (6) (except that all the groups are connected here). First suppose that $q=I_{r,n-r}$ with $1\leq r\leq n-1$, and let's show that neither $\mathrm{SU}(q)$ nor $\mathrm{U}(q)$ are compact. As $\mathrm{SU}(q)$ is closed in $\mathrm{U}(q)$, it suffice to show that $\mathrm{SU}(q)$ is not compact.

Consider the intersection of $\mathrm{SU}(q)$ with the closed subset $X := \begin{pmatrix} * & 0 & \dots & 0 & * \\ 0 & 1 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & 1 & 0 \\ * & 0 & \dots & 0 & * \end{pmatrix}$ of $M_n(\mathbb{C})$. An easy calculation show that an element $A = (a_{ij})$ of X is in $\mathrm{SU}(q)$ if and only we can find $r \in \mathbb{R}_{\geq 0}$ and $t, u \in \mathbb{R}$ such that $a_{11} = \sqrt{1+r^2}e^{it}$, $a_{1n} = re^{iu}$, $a_{n1} = re^{-iu}$ and $a_{nn} = \sqrt{1+r^2}e^{-it}$. So $X \cap \mathrm{SU}(q)$ contains a copy of $\mathbb{R}_{\geq 0}$ as a closed subset, hence it cannot be compact, and neither can $\mathrm{SU}(q)$.

Suppose that $q = I_n$ or $q = I_{0,n} = -I_n$. Then $\mathrm{U}(q)$ (resp. $\mathrm{SU}(q)$) is the usual unitary (resp. special unitary) group in $M_n(\mathbb{C})$, which is denoted by $\mathrm{U}(n)$ (resp. $\mathrm{SU}(n)$). Let's show that both $\mathrm{U}(n)$ and $\mathrm{SU}(n)$ are both compact. As $\mathrm{SU}(n)$ is a closed subgroup of $\mathrm{U}(n)$, it suffices to show that $\mathrm{U}(n)$ is compact. Consider the homeomorphism $M_n(\mathbb{C}) \to (\mathbb{C}^n)^n$ sending a matrix to the list of its column vectors. Then the image of $\mathrm{U}(n)$ is a closed subset of S^n, where $S = \{(z_1, \dots, z_n) \in \mathbb{C}^n || z_1|^2 + \dots + |z_n|^2\}$. As S is compact (as a closed and bounded subset of $\mathbb{C}^n$), $\mathrm{U}(n)$ is compact.

To show that $\mathrm{SU}(q)$ and $\mathrm{U}(q)$, we will use the same method as in (6). First note the following lemma, which is proved exactly as th first lemma of (6) :
Lemma. *Let q be any non-degenerate Hermitian form on $\mathbb{C}^n$, and let $(v_1, \dots, v_r)$ be an orthogonal family in $\mathbb{R}^n$ such that $q(v_i) \in U(1) := \{z \in \mathbb{C} || z| = 1\}$ for every $i \in \{1, \dots, r\}$. Then $(v_1, \dots, v_r)$ can be extended to an orthogonal basis $(v_1, \dots, v_n)$ of $\mathbb{R}^n$ such that $q(v_i) \in \mathrm{U}(1)$ for every $i \in \{1, \dots, n\}$.*

Now we come back to the case $q = I_{r,n-r}$. We may assume that $r \geq 1$. Let's show that both $\mathrm{SU}(q)$ and $\mathrm{U}(q)$ are connected. We reason by induction on n. If $n = 1$, then $\mathrm{U}(q) = \mathrm{U}(1)$ is the unit circle in $\mathbb{C}$ and $\mathrm{SU}(q) = \{1\}$, so both are connected. Assume that $n \geq 2$, and let $S = \{z \in \mathbb{C}^n | q(z) = 1\}$ and $e_1 = (1, 0, \dots, 0)$; as $r \geq 1$, $e_1 \in S$. We have a continuous map $\pi : \mathrm{U}(q) \to S$, $A \longmapsto Ae_1$; we denote by π' its restriction to $\mathrm{SU}(q)$. Let's check the conditions of the second lemma of (6).

- Let's show that π and π' are surjective. It suffices to treat the case of π'. If $v_1 \in S$, the above lemma says that it is possible to complete it to an orthogonal basis $(v_1, \dots, v_n)$ of $\mathbb{C}^n$ such that $q(v_i) \in \mathrm{U}(1)$ for every $i \in \{2, \dots, n\}$. Let A be the matrix with columns $v_1, \dots, v_n$. Then $A \in \mathrm{U}(q)$, and in particular $\det(A) \in \mathrm{U}(1)$. Let A' be the matrix with columns $v_1, \dots, v_{n-1}, \overline{\det(A)}v_n$. Then A' is also in $\mathrm{U}(q)$, and $\det(A') = |\det(A)|^2 = 1$, so $A' \in \mathrm{SU}(q)$. By construction, we have $\pi'(A') = v_1$.
- Let's show that S is path-connected. Let $(z_1, \dots, z_n) \in S$. Write $z_1 = re^{i\theta}$ with $r \in \mathbb{R}_{\geq 0}$ and $\theta \in \mathbb{R}$, and let $\lambda = |z_2|^2 + \dots + |z_r|^2 - |z_{r+1}|^2 - \dots - |z_n|^2$. Then $\lambda \in \mathbb{R}$, and $\lambda \leq 1$ (because $\lambda = 1 - |z_1|^2$). Consider the function $[0,1] \to \mathbb{C}$, $t \longmapsto z_1(t) := \sqrt{1-\lambda t^2}e^{it\theta}$. Note that $z_1(1) = \sqrt{|z_1|^2}e^{i\theta} = z_1$. Let $\gamma : [0,1] \to \mathbb{C}^n$ be defined by
$$\gamma(t) = (z_1(t), tz_2(t), \dots, tz_n(t)).$$

Then, for every $t \in [0,1]$,

$$q(\gamma(t)) = (1-\lambda t^2) + t^2\lambda = 1,$$

i.e., $\gamma(t) \in S$. As $\gamma(0) = e_1$ and $\gamma(1) = (z_1, \dots, z_n)$, this shows that S is path-connected.

- Let's show that the fibers of π and π' are path-connected. As in (6), using the surjectivity of π (resp. π'), we see that all the fibers of π (resp. π') are homemorphic to $\pi^{-1}(e_1) \simeq \mathrm{U}(I_{r-1,n-r})$ (resp. $\pi'^{-1}(e_1) = \mathrm{SU}(I_{r-1,n-r})$), so we can apply the induction hypothesis.
- Let's show that π and π' admit continuous sections locally on S. It suffices to show it for π' (because any section of π' will also give a section of π.) Let $v_1 \in S$. The lemma gives an orthogonal basis $(v_1, \dots, v_n)$ of $\mathbb{C}^n$ such that $q(v_i) \in \mathrm{U}(1)$ for every $i \in \{2, \dots, n\}$. Just as in (6), we can find a neighborhood U of v_1 in S such that, if $v_1' \in U$, then $(v_1', v_2, \dots, v_n)$ is still a basis of $\mathbb{C}^n$ and applying the Gram-Schmidt process to it will make sense and produce an orthonomal basis $(v_1', \dots, v_n')$. Denote by $B(v_1')$ the matrix with columns $v_1', \dots, v_n'$, then $B(v_1') \in \mathrm{U}(q)$ and $B(v_1')e_1 = v_1'$. Let $A(v_1')$ be the matrix with columns $v_1', \dots, v_{n-1}', \overline{\det(B(v_1'))}v_n'$, then $A(V_1') \in \mathrm{SU}(q)$ and $\pi'(A(v_1')) = v_1'$. The function $s : U \to \mathrm{SU}(q)$, $v_1' \longmapsto Q(v_1')$, is continuous and satisfies $\pi' \circ s = \mathrm{id}_U$.

(10). The group $\mathrm{SU}(q)$ is always connected, and it is compact if and only if q has signature $(n,0)$ or $(0,n)$. See (9) for proofs.

(11). We start by proving a useful lemma.

Lemma VII.5.1. *Let K be a field, let V be a finite-dimensional vector space, and let $\langle .,. \rangle$ be a non-degenerate symplectic form on V. Let $v_1, \dots, v_{2r} \in V$ be such that $\langle v_1, v_2 \rangle = \langle v_3, v_4 \rangle = \cdots = \langle v_{2r-1}, v_{2r} \rangle = 1$, and $\langle v_i, v_j \rangle = 0$ if $i = j$ or $\{i,j\}$ is not of the form $\{2m-1, 2m\}$.*

Then :

(i) the family $(v_1, \dots, v_{2r})$ is linearly independent;

(ii) if $W = \mathrm{Span}(v_1, \dots, v_{2r})$, then $V = W \oplus W^\perp$ (where $W^\perp = \{v \in V | \forall w \in W,\ \langle v, w \rangle = 0\}$);

(iii) we can complete it to a basis $(v_1, \dots, v_{2n})$ satisfying a similar condition (i.e. $\langle v_1, v_2 \rangle = \langle v_3, v_4 \rangle = \cdots = \langle v_{2n-1}, v_{2n} \rangle = 1$, and $\langle v_i, v_j \rangle = 0$ if $i = j$ or $\{i,j\}$ is not of the form $\{2m-1, 2m\}$).

In particular, we see that every finite-dimensional K-vector space having a non-degenerate symplectic form must be of even dimension.

Proof. Let show (i) Let $\lambda_1, \dots, \lambda_{2r} \in K$ such that $\lambda_1 v_1 + \cdots + \lambda_{2r} v_{2r} = 0$. Let

$i \in \{1, \dots, 2r\}$. If i is odd, then we have

$$0 = \langle \lambda_1 v_1 + \cdots + \lambda_{2r} v_{2r}, v_{i+1} \rangle = \lambda_i.$$

If i is even, then we have

$$0 = \langle \lambda_1 v_1 + \cdots + \lambda_{2r} v_{2r}, v_{i-1} \rangle = \lambda_i.$$

Now we show (ii). Let $W = \mathrm{Span}(v_1, \dots, v_{2r})$. As the form is non-degenerate, $\dim(W) + \dim(W^\perp) = \dim(V)$. Let's show that $W \cap W^\perp = \{0\}$. Let $v = \lambda_1 v_1 + \cdots + \lambda_{2r} v_{2r} \in W$. If $v \in W^\perp$, we see by looking at all the $\langle v, v_i \rangle$ as above that $\lambda_1 = \cdots = \lambda_{2r} = 0$. Finally, we get that $V = W \oplus W^\perp$.

Now let's show (iii). We proceed by induction on $\dim V - \dim W$. If $\dim V - \dim W = 0$, then $V = W$ and we are done, so suppose that $\dim W < \dim V$. By (ii), we just need to find a basis of $W^\perp$ satisfying the conditions of the lemma (that is, we just need to treat the case $r = 0$). Let $v_{2r+1} \in W^\perp - \{0\}$. As the form is alternating, $\langle v_{2r+1}, v_{2r+1} \rangle = 0$. As the restriction of the form to $W^\perp$ is non-degenerate (because $V = W \oplus W^\perp$ and W and $W^\perp$ are orthogonal), there exists $v_{2r+2} \in W^\perp$ such that $\langle v_{2r+1}, v_{2r+2} \rangle \neq 0$, and we may assume after rescaling that $\langle v_{2r+1}, v_{2r+2} \rangle = 1$. Then we can apply the induction hypothesis to $(v_1, \dots, v_{2r+2})$.

□

Now we come back to the problem. By the lemma, n has to be even, say $n = 2m$, and we can find a basis $(v_1, \dots, v_n)$ of $\mathbb{R}^n$ as in the lemma. Then, in the basis $(v_1, v_3, \dots, v_{2m-1}, v_{2m}, \dots, v_4, v_2)$, the matrix of the symplectic form is J_n. As in (5), this implies that $\mathrm{Sp}(q, \mathbb{R})$ and $\mathrm{Sp}(J_n, \mathbb{R})$ are isomorphic (as topological groups), so we may assume that $q = J_n$. We will also denote the group $\mathrm{Sp}(J_n, \mathbb{R})$ by $\mathrm{Sp}_n(\mathbb{R})$.

Let's first show that $\mathrm{Sp}_n(\mathbb{R})$ is not compact. Consider the the closed subset $X := \begin{pmatrix} * & 0 & \dots & 0 & * \\ 0 & 1 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & 1 & 0 \\ * & 0 & \dots & 0 & * \end{pmatrix}$ of $M_n(\mathbb{R})$. An easy calculation show that a matrix $A = (a_{ij})$ of $M_n(\mathbb{R})$ is in the closed subset $\mathrm{Sp}_n(\mathbb{R}) \cap X$ of $\mathrm{Sp}_n(\mathbb{R})$ if and only if $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ is in $\mathrm{SL}_2(\mathbb{R})$. As $\mathrm{SL}_2(\mathbb{R})$ is not compact, neither is $\mathrm{Sp}_n(\mathbb{R})$.

Now let's show that $\mathrm{Sp}_n(\mathbb{R})$ is connected. We follow the hint, in fact we'll follow all three hints.

(a) The first method is similar to the method used in (6)-(10). We do an induction on n. If $n = 2$, we saw above that $\mathrm{Sp}_n(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R})$, and this is connected by (2). So let's suppose that $n \geq 4$.

Let $\langle .,. \rangle$ be the symplectic form on $\mathbb{R}^n$ with matrix J_n. We consider the continuous map $\pi : \mathrm{Sp}_n(\mathbb{R}) \to Z$, $A \longmapsto Ae_1$ (where $e_1 = (1, 0, \ldots, 0)$), and we try to check the hypotheses of the second lemma of (6).

- Let's show that π is surjective. If $v_1 \in Z$, complete it to a basis $(v_1, \ldots, v_{2m})$ of $\mathbb{R}^n$ as in the lemma above. Let A be the matrix with columns $v_1, v_3, \ldots, v_{2m-1}, v_{2m}, \ldots, v_4, v_2$. Then $A \in \mathrm{Sp}_n(\mathbb{R})$ and $\pi(A) = v_1$.

- As $n \geq 2$, Z is path-connected.

- Let's show that the fibers of π are path-connected. As in (6), using the surjectivity of π, we see that all the fibers of π are homemorphic to $\pi^{-1}(e_1)$. Let's show that $\pi^{-1}(e_1) \simeq \mathrm{Sp}_{2n-2}(\mathbb{R})$ so that we can apply the induction hypothesis.

  Let $A \in \pi^{-1}(e_1)$, i.e. $Ae_1 = e_1$. Let $v = Ae_n$. Then v is orthogonal to $e_2, \ldots, e_n$, and $\langle e_1, v \rangle = 1$. If we write $v = \lambda_1 e_1 + \cdots + \lambda_n e_n$, this implies easily that $\lambda_1 = \cdots = \lambda_{n-1} = 0$ and $\lambda_n = 1$, i.e. that $v = e_n$. So the matrix A is of the form $\begin{pmatrix} 1 & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & 1 \end{pmatrix}$ with $B \in M_{2n-2}(\mathbb{R})$, and it is easy to see that B is actually in $\mathrm{Sp}_{2n-2}(\mathbb{R})$. Conversely, every matrix of this form is clearly in $\pi^{-1}(e_1)$.

- Let's show that π admits continuous sections locally on Z. Let $v_1 \in Z$, and complete it to a basis $(v_1, \ldots, v_{2m})$ of $\mathbb{R}^{2m}$ satisfying the condition of the lemma. As in (6), the main point is to show that, if v_1' is close enough to v_1, then we can complete to a basis $(v_1', \ldots, v_{2m}')$ satisfying the condition of the lemma that depends continuously on v_1'. The inspiration is again the Gram-Schmidt process. Let v_1' be close enough to v_1 so that $\langle v_1', v_2 \rangle \neq 0$. We set $v_2' = \frac{\langle v_1', v_2 \rangle}{v}{}_2$, so that $\langle v_1', v_2' \rangle = 1$. We set $v_3' = -\langle v_2', v_3 \rangle v_1' - \langle v_1', v_3 \rangle v_2' + v_3$, so that $\langle v_i', v_3' \rangle = 0$ for $i = 1, 2$. If v_1' is close enough to v_1, then v_3' is close to v_3 and so we have $\langle v_3', v_4 \rangle \neq 0$. We set $v_4' = -\langle v_2', v_4 \rangle v_1' - \langle v_1', v_4 \rangle v_2' + \frac{1}{\langle v_3', v_4 \rangle} v_4$, so that $\langle v_i', v_4' \rangle = 0$ if $i = 1, 2$ and $\langle v_3', v_4' \rangle = 1$. We continue the construction of $v_1', \ldots, v_{2m}'$ in the same way : If we already have $v_1', \ldots, v_{2r}'$, then we set

$$v_{2r+1}' = v_{2r+1} - \sum_{s=1}^{r} (\langle v_{2s}', v_{2r+1} \rangle v_{2s-1}' + \langle v_{2s-1}', v_{2r+1} \rangle v_{2s}')$$

  and

$$v_{2r+2}' = \frac{1}{\langle v_{2r+1}', v_{2r+2} \rangle} v_{2r+2} - \sum_{s=1}^{r} (\langle v_{2s}', v_{2r+2} \rangle v_{2s-1}' + \langle v_{2s-1}', v_{2r+2} \rangle v_{2s}').$$

  The second expression makes sense if v_1' is close enough to v_1, because then v_{2r+1}' will be close to v_{2r+1} and so $\langle v_{2r+1}', v_{2r+2} \rangle$ will be nonzero. Finally, we denote by $B(v_1')$ the matrix with columns $v_1', v_3', \ldots, v_{2m-1}', v_{2m}', \ldots, v_4', v_2'$. Then

$B(v_1') \in \mathrm{Sp}_{2m}(\mathbb{R})$, $\pi(B(v_1')) = v_1'$ and $B(v_1')$ depends continuously on v_1'. So we are done.

(b) We use the notation of (b) of the problem. Remember that n is even by the lemma, write $n = 2m$.

(i). The action is well-defined because $\mathrm{Sp}_n(\mathbb{R})$ preserves the form $\langle ., . \rangle$, so the image of a Lagrangian subspace is a Lagrangian subspace. Let $(e_1, \dots, e_n)$ be the canonical basis of $\mathbb{R}^n$, and let $V_0 = \mathrm{Span}(e_1, \dots, e_{n/2})$. Then V_0 is a Lagrangian subspace of $\mathbb{R}^n$. To show that the action of K on L is transitive, we just need to show that, for every $V \in L$, there exists $g \in K$ such that $V = gV_0$ (or, in other words, such that the first $n/2$ columns of g generate V).

Let $(v_1, \dots, v_m)$ be a basis of V that is orthonormal for the usual scalar product on $\mathbb{R}^n$, and consider the matrix with columns $v_1, \dots, v_m$. This is a $2m \times m$ matrix, and we write it $M = \begin{pmatrix} A \\ B \end{pmatrix}$, with $A, B \in M_m(\mathbb{R})$. The condition that the basis is orthonormal is equivalent to $I_m = {}^tMM = {}^tAA + {}^tBB$, and the fact that V is Lagrangian is equivalent to $0 = {}^tMJ_nM = -{}^tBA + {}^tAB$. By (ii) (or by an easy calculation), the matrix $g := \begin{pmatrix} A & -B \\ B & A \end{pmatrix}$ is in K. By the choice of A and B, $gV_0 = V$.

(ii). Let $g \in K$, write $g = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ with $A, B, C, D \in M_{n/2}(\mathbb{R})$. Then ${}^tgJ_ng = J_n$ because $g \in \mathrm{Sp}_n(\mathbb{R})$, so $g^tgJ_ng = gJ_n$. As $g^tg = I_n$ (because $g \in \mathrm{O}_n(\mathbb{R})$), we get $gJ_n = J_ng$, which is equivalent to $A = D$ and $B = -C$. If $g = \begin{pmatrix} A & B \\ -B & A \end{pmatrix}$, then the conditions that $g \in \mathrm{O}(n)$ and $g \in \mathrm{Sp}_n(\mathbb{R})$ actually become equivalent to each other, and they are both equivalent to the conditions that ${}^tAA + {}^tBB = I_n$ and ${}^tAB = {}^tBA$. So K is the group of matrices $\begin{pmatrix} A & B \\ -B & A \end{pmatrix}$ satisfying these two conditions on A and B.

Now we consider the map $\varphi : K \to M_n(\mathbb{C})$ sending $\begin{pmatrix} A & B \\ -B & A \end{pmatrix}$ to $A - iB$. For $A, B \in M_n(\mathbb{R})$, we have

$$(A - iB)^*(A - iB) = ({}^tA + i{}^tB)(A - iB) = ({}^tAA + {}^tBB) + i({}^tBA - {}^tAB).$$

This show that $\mathrm{U}(n)$ is exactly the image of φ. As it is clear that φ is a homemorphism onto its image, it just remains to show that φ is a morphism of groups. But this is a straightforward calculation.

(iii). First, a matrix $M \in M_n(\mathbb{R})$ stabilizes V_0 if and only if it is of the form $M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$ with $A, B, D \in M_m(\mathbb{R})$. Such a matrix is in $\mathrm{Sp}_n(\mathbb{R})$ if and

only if ${}^tAD = I_m$ and ${}^tBD = {}^tDB$. Let Sym be the set of symmetric matrices in $M_m(\mathbb{R})$, it's a $\mathbb{R}$-vector of $M_m(\mathbb{R})$ and in particular connected. The map $\begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \longmapsto (D, {}^tDB)$ induces a homeomorphism $\psi : P \xrightarrow{\sim} \mathrm{GL}_m(\mathbb{R}) \times Sym$. As $\mathrm{GL}_m(\mathbb{R})$ is not connected (see (3)), P is not connected.

(iv). First let's calculate P^0. I claim that $\mathrm{GL}_m(\mathbb{R})^0$ (the connected component of I_m in $\mathrm{GL}_m(\mathbb{R})$) is equal to $\mathrm{GL}_m(\mathbb{R})^+ := \{g \in \mathrm{GL}_m(\mathbb{R}) | \det(g) > 0\}$. First, $\mathrm{GL}_m(\mathbb{R})^+$ is the inverse image by the continuous surjectif map $\det : \mathrm{GL}_m(\mathbb{R}) \to \mathbb{R}^\times$ of a connected component of $\mathbb{R}^\times$, so it is a union of connected components of $\mathrm{GL}_m(\mathbb{R})$. So it is enough to show that $\mathrm{GL}_m(\mathbb{R})^+$ is connected. The maps $\mathrm{GL}_m(\mathbb{R})^+ \to \mathrm{SL}_m(\mathbb{R}) \times \mathbb{R}_{>0}$, $g \longmapsto (\det(g)^{-1/m} g, \det(g))$, and $\mathrm{SL}_m(\mathbb{R}) \times \mathbb{R}_{>0} \to \mathrm{GL}_m(\mathbb{R})^+$, $(g, \lambda) \longmapsto \lambda^{1/m} g$, are continuous and inverse of each other, so $\mathrm{GL}_m(\mathbb{R})^+$ is homeomorphic to $\mathrm{SL}_m(\mathbb{R}) \times \mathbb{R}_{>0}$. As $\mathrm{SL}_m(\mathbb{R})$ is connected (see question (4)), $\mathrm{GL}_m(\mathbb{R})^+$ is connected.

Now, using the homemorphism ψ of (iii), we get $P^0 = \psi^{-1}(\mathrm{GL}_m(\mathbb{R})^+ \times Sym)$, so P^0 is the set of $\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$ in P such that $\det(D) > 0$ (or equivalently $\det(A) > 0$).

We come back to the problem. Let $g \in \mathrm{Sp}_n(\mathbb{R})$. Then $V = gV_0$ is a Lagrangian subspace of $\mathbb{R}^n$, so, by (i), there exists $h \in K$ such that $V = hV_0$. We get $hV_0 = gV_0$, i.e., $h^{-1}g \in P$, which means that $g = hp$ with $p \in P$. We still need to show that we can choose $p \in P^0$. Let A be the diagonal matrix $\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & -1 \end{pmatrix}$ in $\mathrm{GL}_m(\mathbb{R})$, let $q = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix}$. Then $q^2 = I_n$, $q \in K \cap P$, and exactly one of p and qp is in P^0 (this is obvious on the description of P^0 we obtained above). So either $p \in P^0$ and we're done, or we write $g = (hq)(qp)$, and we have $hq \in K$ and $qp \in P^0$.

(v). By (ii) (and question (9)), K is path-connected. By the calculation of P^0 in (iv), P^0 is path-connected. Let $g \in \mathrm{Sp}_n(\mathbb{R})$; by (iv), we can write $g = hp$, with $h \in K$ and $p \in P^0$. Choose continuous maps $t \longmapsto p(t)$ (resp. $t \longmapsto h(t)$) from $[0, 1]$ to P^0 (resp. K) such that $p(0) = h(0) = I_n$ and $p(1) = p$, $h(1) = h$. Then $t \longmapsto h(t)p(t)$ is a continuous path on $\mathrm{Sp}_n(\mathbb{R})$ that connects I_n and g.

(c) (i). This is identical to (b)(ii).

(ii). Let

$$\mathfrak{s} = \{A \in M_n(\mathbb{R}) | {}^tA = A \text{ and } AJ + JA = 0\}.$$

This is vector subspace of $M_n(\mathbb{R})$, and in particular it is path-connected. We'll show that the matrix exponential sends $\mathfrak{s}$ onto S, which implies that S is path-

connected.

Let $g \in S$. As g is symmetric positive definite, there exists $h \in \mathrm{O}(n)$ and $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \in M_n(\mathbb{R})$, with $\lambda_1, \dots, \lambda_n > 0$, such that $g = h^{-1}Dh$. Let $E = \begin{pmatrix} \log \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \log \lambda_n \end{pmatrix}$, so that $e^E = D$ and $e^{h^{-1}Eh} = g$. Also, $h^{-1}Eh$ is symmetric because E is symmetric and h is orthogonal. On the other hand, the function $f : \mathbb{R} \to M_n(\mathbb{R})$, $t \longmapsto e^{th^{-1}Eh}Je^{th^{-1}Eh} - J$ is real analytic, and it sends every $t \in \mathbb{N}$ to 0 (if $t \in \mathbb{N}$, then $e^{th^{-1}Eh} = g^t \in \mathrm{Sp}_n(\mathbb{R})$), so it is identically 0 by the identity theorem. [16] Also, we have

$$f'(t) = h^{-1}Ehf(t) + f(t)h^{-1}Eh.$$

In particular, $0 = f'(0) = h^{-1}EhJ + Jh^{-1}Eh$, so $h^{-1}Eh \in \mathfrak{s}$.

(iii). Let's show the uniqueness. Let $g \in \mathrm{Sp}_n(\mathbb{R})$, suppose that $g = us = u's'$, with $u, u' \in K$ and $s, s' \in S$. Then ${}^t gg = s^2 = (s')^2$. As s and s' are symmetric positive definite, $s^2 = (s')^2$ implies that $s = s'$,[17] and then we also get $u = u'$.

Let's show existence. Let $g \in \mathrm{Sp}_n(\mathbb{R})$. Then ${}^t gg$ is in S, so, by (ii), there exists $A \in \mathfrak{s}$ such that ${}^t gg = e^A$. Let $s = e^{\frac{1}{2}A}$, we have $s \in S$. Let $u = gs^{-1}$. Then

$${}^t uu = s^{-1\,t}ggs^{-1} = s^{-1}s^2s^{-1} = I_n,$$

so $u \in \mathrm{Sp}_n(\mathbb{R}) \cap \mathrm{O}_n(\mathbb{R}) = K$.

(iv). We know that K is path-connected by (i) and question (9), and we know that S is path-connected by (ii). We conclude that $\mathrm{Sp}_n(\mathbb{R})$ is path-connected as in (b)(v).

(12). First, the lemma of (11) show that we just need to consider the case where $q = J_n$. We write $\mathrm{Sp}_n(\mathbb{C}) = \mathrm{Sp}(J_n, \mathbb{C})$. As the closed subgroup $\mathrm{Sp}_n(\mathbb{R}) = \mathrm{Sp}_n(\mathbb{C}) \cap M_n(\mathbb{R})$ of $\mathrm{Sp}_n(\mathbb{C})$ is not compact (by (11)), $\mathrm{Sp}_n(\mathbb{C})$ cannot be compact. To show that $\mathrm{Sp}_n(\mathbb{C})$ is connected, we can adapt any of the methods of (11), but the easiest is to use method (a) goes through with almost no change. In methods (b) and (c), we have to use $K = \mathrm{Sp}_n(\mathbb{C}) \cap \mathrm{U}(n)$,

[16] See corollary 1.2.6 of Krantz and Parks's book [19].

[17] This is a standard exercise. Up to conjugating by an orthogonal matrix, we may assume that s' is diagonal. Then we want to show that s is also diagonal, which is enough because the eigenvalues of s have to be equal to the eigenvalues of s', as they're both the square roots of the eigenvalues of $s^2 = (s')^2$. So we are reduced to the following statement : Let s be a symmetric definite positive matrix such that s^2 is diagonal, then s is also diagonal. Let $\lambda_1, \dots, \lambda_n \in \mathbb{R}_{>0}$ be the eigenvalues of s. As taking the square is an injective operation on $\mathbb{R}_{>0}$, we have $\lambda_i^2 = \lambda_j^2$ if and only if $\lambda_i = \lambda_j$, and so there exists a polynomial $P \in \mathbb{R}[X]$ such that $P(\lambda_i^2) = \lambda_i$ for every $i \in \{1, \dots, n\}$. As s is diagonalizable, $P(s^2) = s$. But s^2 is diagonal, so $P(s^2)$ is also diagonal.

which is isomorphic to the group of "unitary" matrices in $M_{n/2}(\mathbb{H})$, and to prove that this is path-connected. (The rest of methods (b) and (c) adapts very easily, but the previous part requires more work.)

□

VII.5.5 Representations of compact commutative groups

(1). If G is an abelian compact Hausdorff topological group, show that every irreducible continuous finite-dimensional representation of G is of dimension 1.

(2). Find all the continuous 1-dimensional representations of $S^1 := \{z \in \mathbb{C} || z| = 1\}$ and their images in $L^2(S^1)$.

Solution.

(1). It's exactly the same proof as for finite groups. Let (V, ρ) be an irreducible continuous representation of G. Then for every $g \in G$, $\rho(g) \in \mathrm{End}_G(V)$, so, by Schur's lemma, there exists $\chi(g) \in \mathbb{C}^\times$ such that $\rho(g) = \chi(g)\mathrm{id}_V$. This show that every vector subspace of V is invariant by G. As V is irreducible, the only G-invariant subspaces of V are 0 and V, so $\dim_{\mathbb{C}} V = 1$.

(2). For every $n \in \mathbb{Z}$, let $\chi_n : S^1 \to \mathbb{C}^\times$ be the map $z \longmapsto z^n$. This is a continuous morphism of groups, hence a continuous 1-dimensional representation of S^1. Its image in $L^2(S^1)$ is also 1-dimensional, and it is generated by the function $z \to \mathrm{Tr}(\chi_n(z)^{-1}) = z^{-n}$.

Now let's show that every continuous 1-dimensional representation of S^1 is of the form χ_n (and hence, by (a), every continuous irreducible finite-dimensional representation of S^1). Let $\chi : S^1 \to \mathbb{C}^\times$ be such a representation, ie a continuous morphism of groups.

Composing χ with the map $\pi : \mathbb{R} \to S^1$, $x \longmapsto \exp(2\pi i x)$, we get a continuous morphism of groups $\psi : \mathbb{R} \to \mathbb{C}^\times$. As ψ is continuous, $\psi(x) \to 1$ as $x \to 1$. Hence we can find $c > 0$ such that $a := \int_0^c \psi(x)dx \neq 0$. (Just choose $c > 0$ such that $1/2 \leq \psi(x)$ for $0 \leq x \leq c$.) Now for every $x \in \mathbb{R}$,

$$\int_x^{x+c} \psi(t)dt = \int_0^c \psi(x+t)dt = \psi(x)\int_0^c \psi(t)dt = a\psi(x),$$

because ψ is a morphism of groups. So we get

$$\psi(x) = a^{-1}\int_x^{x+c} \psi(t)dt.$$

This shows that ψ is derivable, and also that

$$\psi'(x) = a^{-1}(\psi(x+c) - \psi(x)) = \psi(x)a^{-1}(\psi(c) - 1).$$

So $\psi(x) = \exp(\alpha x)$ (with $\alpha = a^{-1}(\psi(c) - 1)$, though we don't care).

Now as ψ factors through $\pi : \mathbb{R} \to S^1$, we have $\exp(2i\pi\alpha) = 1$, hence $\alpha \in \mathbb{Z}$, and we get $\chi = \chi_\alpha$.

(We could also have used the fact that the functions χ_n, $n \in \mathbb{Z}$, form a Hilbert basis of $L^2(S^1)$ by the theory of Fourier series. So if S^1 has any 1-dimensional representation χ that were not of the χ_n, this χ would have to be in $L^2(S^1)$, nonzero and orthogonal to all the χ_n, and that's impossible.)

□

VII.5.6 Complex representations of profinite groups

(1). Let n be a positive integer. Put some norm $\|.\|$ on $M_n(\mathbb{C})$. (They are all equivalent, so you can choose your favourite one.) Show that there exists $\varepsilon > 0$ such that the only subgroup of $\mathrm{GL}_n(\mathbb{C})$ contained in $\{g \in \mathrm{GL}_n(\mathbb{C}) | \|g - 1\| < \varepsilon\}$ is $\{1\}$. *(Hint : Start with the case $n = 1$, and then don't do an induction on n.)*

(2). We say that a topological group is *profinite* if

$$\Gamma = \varprojlim_{\Delta} \Gamma/\Delta := \{(x_\Delta) \in \prod_{\Delta} \Gamma/\Delta | \forall \Delta' \subset \Delta,\ x_\Delta = x_{\Delta'}\Delta\},$$

where we take the limit over all normal subgroups Δ of finite index of Γ, and if the topology of Γ is induced by the topology of $\prod_\Delta \Gamma/\Delta$, where we put the discrete topology on each Γ/Δ. Examples of profinite groups are $\mathbb{Z}_p$, $\mathbb{Z}_p^\times$ and the Galois group of a possibly infinite Galois extension of fields.

(a) Show that a profinite group is compact Hausdorff.

(b) Suppose that Γ is a profinite group, and let (V, ρ) be a continuous finite-dimensional representation of Γ on a $\mathbb{C}$-vector space. Show that $\mathrm{Ker}\,\rho$ is a subgroup of finite index of Γ.

Solution.

(1). We do the case $n = 1$. Let G be a subgroup of $\mathbb{C}^\times$ such that, for every $g \in G$, $|g-1| < 1/2$. First we show that $G \subset S^1$. Indeed, if there is a $g \in G$ such that $|g| \neq 1$, then either $|g|^n \to 0$ as $n \to +\infty$, or $|g^n| \to +\infty$ as $n \to +\infty$ In both cases $|g^n - 1|$ eventually becomes bigger than $1/2$, which is impossible. Now suppose that $G = \{1\}$, and let $g \in G - \{1\}$. Write $g = \exp(2\pi i\alpha)$, $\alpha \in \mathbb{R}$. We may assume that $0 < \alpha < \pi/2$ (if this does not work for g, it will for g^{-1}). There exists $n \in \mathbb{Z}_{\geq 1}$ such that $\pi/2 < n\alpha < \pi$, and then $|g^n - 1| > 1/2$, contradicting our hypothesis on G. So $G = \{1\}$.

Now let n be any positive integer, and choose $\varepsilon > 0$ such that, if $g \in \mathrm{GL}_n(\mathbb{C})$ is such that $\|g-1\| < \varepsilon$, then for every eigenvalue λ of g satisfies $|\lambda - 1| < 1/2$. Let G be a subgroup of $\mathrm{GL}_n(\mathbb{C})$ such that $\|g-1\| < \varepsilon$ for every $g \in \mathrm{GL}_n(\mathbb{C})$.

Let $g \in G$, and let $\lambda_1, \ldots, \lambda_n$ be its eigenvalues. For every $k \in \mathbb{Z}$, the eigenvalues of g^k are $\lambda_1^k, \ldots, \lambda_n^k$. By the first part, this forces all the λ_i to be equal to 1.

We have shown that all the eigenvalues of g are equal to 1, so $g = 1 + N$ with $N \in M_n(\mathbb{C})$ nilpotent. If $N \neq 0$, choose $e \in \mathbb{C}^n$ such that $N(e) \neq 0$ but $N^2(e) = 0$. Then, for every $m \in \mathbb{Z}_{\geq 1}$, $g^m(e) = (1+N)^m(e) = e + mN(e)$, so $\|g^m(e) - e\| = m\|N(e)\| \to +\infty$ as $m \to +\infty$, which contradicts the hypothesis on G. So the only element of G is 1.

Another way to see that every element of G has to be semisimple is the following : Take ε small enough so that every $g \in M_n(\mathbb{C})$ with $\|g-1\| \leq \varepsilon$ is invertible, and consider the closure $\overline{G}$ of G in $M_n(\mathbb{C})$. By the choice of ε, this is also the closure of G in $\mathrm{GL}_n(\mathbb{C})$, hence it's a subgroup of $\mathrm{GL}_n(\mathbb{C})$. But it's also a compact subset of $M_n(\mathbb{C})$ because it's closed and bounded. So $\overline{G}$ is a compact subgroup of $\mathrm{GL}_n(\mathbb{C})$. By theorem V.3.1.6 of chapter V, there exists a Hermitian inner product on $\mathbb{C}^n$ for which every element of $\overline{G}$ is unitary. In other words, there exists $g \in \mathrm{GL}_n(\mathbb{C})$ such that $\overline{G} \subset gU(n)g^{-1}$. As every element of $U(n)$ is diagonalizable, so is every element of $\overline{G}$.

(2). (a) This follows from Tychonoff's theorem, because finite sets with the discrete topology are compact Hausdorff.

(b) After choosing a basis of V, we can see ρ as a continuous morphism of groups $G \to \mathrm{GL}_n(\mathbb{C})$. Let $\varepsilon > 0$ be as in (1), and let $U = \{g \in \mathrm{GL}_n(\mathbb{C}) | \|g-1\| < \varepsilon\}$. Then U is an open neighbourhood of 1 in $\mathrm{GL}_n(\mathbb{C})$, so $\rho^{-1}(U)$ is an open neighbourghood of 1 in G. By the fact that $\Gamma = \varprojlim_{\Delta} \Gamma/\Delta$ and the definition of the product topology, this implies that $\rho^{-1}(U)$ contains a normal subgroup Δ of Γ of finite index. In particular, $\rho(\Delta)$ is a subgroup of $\mathrm{GL}_n(\mathbb{C})$ contained in U. By (1), $\rho(\Delta) = 1$, ie $\Delta \subset \mathrm{Ker}(\rho)$, so $[\Gamma : \mathrm{Ker}(\rho)] < \infty$.

□

VII.5.7 A unitary representation of G such that $G \to \mathrm{U}(V)$ is not continuous

Let G be a non-discrete compact group. Show that the morphism $\rho : G \to U(L^2(G)),\ x \mapsto R_x$ is not continuous, where the topology on $U(L^2(G))$ is the one induced by $\|.\|_{op}$. (Note : you don't really need G to be compact for this, G locally compact would be enough, but non-discreteness is of course necessary.)

Solution. We will show that, for every $x \neq 1$ in G, we have $\|\mathrm{id} - R_x\|_{op} \geq \frac{1}{\sqrt{2}}$.[18] As G is not discrete, we can find a sequence $(x_n)_{n\geq 0}$ converging to 1 such that $x_n \neq 1$ for every n, so this will imply the result.

Let $x \in G - \{1\}$. We choose open subsets $1 \in U \subset V$ in G such that $\overline{U}$ is compact and contained in V, $\mathrm{vol}(V) \leq 2\,\mathrm{vol}(\overline{U})$ et $\overline{U} \cap \overline{U}x = \varnothing$. By Urysohn's lemma (see problem VII.5.2), there exists a continuous function $f : G \to [0,1]$ such that $f_{|\overline{U}} = 1$ and $f_{G-V} = 0$. We have

$$\|f\|^2_{L^2(G)} = \int_G |f(g)|^2 dg = \int_V |f(g)|^2 dg \leq \mathrm{vol}(V) \leq 2\,\mathrm{vol}(\overline{U}).$$

On the other hand,

$$\|f - R_x(f)\|^2_{L^2(G)} = \int_G |f(g) - f(gx)|^2 dg \geq \int_{\overline{U}} |f(g) - f(gx)|^2 dg = \mathrm{vol}(\overline{U}).$$

So

$$\|\mathrm{id} - R_x\|^2_{op} \geq \frac{\|f - R_x(f)\|^2_{L^2(G)}}{\|f\|^2_{L^2G}} \geq \frac{\mathrm{vol}(\overline{U})}{2\,\mathrm{vol}(\overline{U})} = \frac{1}{2}.$$

□

VII.5.8 A compact group with no faihful representation

Find a compact (Hausdorff) topological group that doesn't have any faithful finite-dimensional representation.

Solution. By problem VII.5.6, any infinite profinite group will do. Take for example $\widehat{\mathbb{Z}} := \varprojlim_{n\geq 1} \mathbb{Z}/n\mathbb{Z}$, where the integers n are ordered by the divisibility relation and, if $n|m$, the map $\mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ is the obvious projection.

□

VII.5.9 Uniqueness of the inner product making an irreducible representation unitary

Let G be a compact Hausdorff group, let V be a finite-dimensional $\mathbb{C}$-vector space and $\rho : G \to \mathrm{GL}(V)$ be a continuous representation of G on V. Assume that ρ is an irreducible representation of G.

[18] We could actually show with the same methods that $\|R_x - \mathrm{id}|_{op} \geq 1$.

If $\langle .,. \rangle_1$ and $\langle .,. \rangle_2$ are two Hermitian inner products making ρ a unitary representation, show that there exists $\lambda \in \mathbb{R}_{>0}$ such that $\langle .,. \rangle_2 = \lambda \langle .,. \rangle_1$.

Solution. Let S be the set of bi-additive forms $V \times V \to \mathbb{C}$ that are $\mathbb{C}$-linear in the first variable and $\mathbb{C}$-antilinear in the second; that is, if $f \in S$, then, for all $v, v', w, w' \in V$ and $\lambda, \mu \in \mathbb{C}$, we have

$$f(\lambda v + v', \mu w + w') = \lambda\overline{\mu} f(v,w) + \lambda f(v,w') + \overline{\mu} f(v',w) + f(w,w').$$

This is a finite-dimensional $\mathbb{C}$-vector space in the obvious way, and we make G act on it by : if $g \in G$ and $f \in S$, then $g.f$ is the form $(v,w) \longmapsto f(\rho(g)^{-1}v, \rho(g)^{-1}w)$. It is easy to see that this makes S a continuous representation of G, but this will also follow from the next paragraph. Note also that $\dim_{\mathbb{C}} S = (\dim_{\mathbb{C}} V)^2$. (If $(e_1, \dots, e_n)$ is a $\mathbb{C}$-basis of V, then sending $f \in S$ to the matrix $(f(e_i, e_j))_{1 \leq i,j \leq n}$ gives a $\mathbb{R}$-linear isomorphism $S \xrightarrow{\sim} M_n(\mathbb{C})$, so $\dim_{\mathbb{C}} S = \frac{1}{2} \dim_{\mathbb{R}} S = \frac{1}{2} \dim_{\mathbb{R}} M_n(\mathbb{C}) = \frac{1}{2}(2n^2) = n^2$.)

Let $\varphi : V \to V^*$ be the map $v \longmapsto (w \longmapsto \langle w, v \rangle_1)$. It is an isomorphisme of $\mathbb{R}$-vector spaces, and we have $\varphi(\lambda v) = \overline{\lambda} v$ for all $\lambda \in \mathbb{C}$ and $v \in V$. Also, φ is G-equivariant, because $\langle \rho(g)v, w \rangle_1 = \langle v, \rho(g)^{-1} w \rangle_1$ for all $v, w \in V$. Now write $W = V \otimes_{\mathbb{C}} V^*$, and consider the map $\psi : W^* \to S$ sending $u : W \to \mathbb{C}$ to the form $(v,w) \longmapsto u(v \otimes \varphi(w))$. I claim that this well-defined and a G-equivariant $\mathbb{C}$-linear isomorphism. Indeed, it is straighforward to check that ψ is well-defined, $\mathbb{C}$-linear and G-equivariant; it is injective because the map $\mathrm{id}_V \otimes \varphi : V \otimes_{\mathbb{C}} V \to V \otimes_{\mathbb{C}} V^*$ is a $\mathbb{R}$-linear isomorphism, hence in particular surjective; and it is an isomorphism because its source and target have the same dimension, which is $(\dim_{\mathbb{C}} V)^2$. So the representations of G on W^* and on S are isomorphic.

Now note that Hermitian inner products on V are elements of S, and that saying that a Hermitian inner product on V makes ρ a unitary representations is equivalent to saying that the corresponding element of S is in S^G (by the definition of the action of G on S). So if we show that $\dim_{\mathbb{C}}(S^G) = 1$, we will be very close to solving the problem. Let's use characters to calculate $\dim_{\mathbb{C}}(S^G)$. By decomposing S into irreducible representations and using corollary V.3.3.3 of chapter V, we see (as in the proof of lemma II.1.2.4 of chapter II) that

$$\dim_{\mathbb{C}}(S^G) = \int_G \chi_S(g) dg.$$

We can calculate the character of S using the G-equivariant isomorphism $S \simeq W^*$. Let $g \in G$. By propositions II.1.1.3 and II.1.1.11 of chapter II (which don't require the group G to be finite), we have

$$\chi_{W^*}(g) = \chi_W(g^{-1}) = \chi_V(g^{-1})\chi_{V^*}(g^{-1}) = \chi_V(g^{-1})\chi_V(g).$$

Moreover, as V has a Hermitian inner product that makes $\rho(g)$ unitary, all the eigenvalues of $\rho(g)$ are complex numbers of module 1, and so $\chi_V(g^{-1}) = \overline{\chi_V(g)}$. Finally, we get

$$\chi_S(g) = \chi_{W^*}(g) = |\chi_v(g)|^2.$$

As V is irreducible, Schur orthogonality (corollary V.3.3.3 of chapter V) gives

$$\int_G \chi_S(g) dg = \int_G |\chi_V(g)|^2 dg = 1.$$

So we have proved that $\dim_{\mathbb{C}}(S^G) = 1$.

As $\langle .,. \rangle_1$ and $\langle .,. \rangle_2$ are both nonzero elements of the 1-dimensional $\mathbb{C}$-vector space S^G, there exists $\lambda \in \mathbb{C}$ such that $\langle .,. \rangle_2 = \lambda \langle .,. \rangle_1$. Taking v to be any nonzero element of V, we get $\lambda = \frac{\langle v,v \rangle_2}{\langle v,v \rangle_1} \in \mathbb{R}_{>0}$.

□

VII.6 Chapter VI exercises

VII.6.1 Surjectivity of the exponential map

(1). Show that the exponential map $\exp : M_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$ is surjective.

(2). Is $\exp : M_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})$ surjective ? If not, what is its image ?

(3). Let $\mathrm{SO}(n) = \{A \in M_n(\mathbb{R}) | {}^tAA = I_n \text{ and } \det(A) = 1\}$ and $\mathfrak{so}(n) = \{A \in M_n(\mathbb{R}) | {}^tA + A = 0\}$. Show that $\exp(\mathfrak{so}(n)) = \mathrm{SO}(n)$.

(4). Let $\mathrm{O}(n) = \{A \in M_n(\mathbb{R}) | {}^tAA = I_n\}$. Can you find a subspace E of $M_n(\mathbb{R})$ such that $\exp(E) = \mathrm{O}(n)$?

Solution.

(1). Let $g \in \mathrm{GL}_n(\mathbb{C})$. We write $g = su$, with s diagonalizable, u unipotent and $su = us$ (this is the Chevalley-Jordan decomposition). The idea is to find "logarithms" of s and u separately, and to choose them so that they will also commute.

As s is diagonalizable, we have $s = hdh^{-1}$ with $h \in \mathrm{GL}_n(\mathbb{C})$ and d a diagonal matrix. Call $\lambda_1, \ldots, \lambda_n$ the diagonal terms of d, and choose $\mu_1, \ldots, \mu_n \in \mathbb{C}$ such that $e^{\mu_i} = \lambda_i$ for every i (this is possible because the λ_i are nonzero). If $\lambda_i = \lambda_j$, we can arrange that $\mu_i = \mu_j$; so we may assume that there exists a polynomial $P \in \mathbb{C}[X]$ such that $P(\lambda_i) = \mu_i$ for every i. Let D be the diagonal matrix with diagonal entries $\mu_1, \ldots, \mu_n$, let $S = hDh^{-1}$. We have $e^D = d$, hence $e^S = he^Dh^{-1} = s$. Also, as $P(d) = D$, we also have $P(s) = S$, and so S and u commute.

On the other hand, u is unipotent, so $n := u - 1$ is nilpotent. Let

$$N = \sum_{r \geq 1} (-1)^{r-1} \frac{1}{r} n^r.$$

This sum is finite because n is nilpotent, and, by proposition VI.3.4 of chapter VI, we have $e^N = n$. Also, N is by definition a polynomial in u, so it commutes with every matrix that commutes with u, and in particular it commutes with S.

Finally, using the fact that S and N commute, we get $e^{S+N} = e^S e^N = su = g$.

(2). We know that $\exp$ is a continuous map and that $M_n(\mathbb{R})$ is connected, so $\exp(M_n(\mathbb{R}))$ is also connected. As $\mathrm{GL}_n(\mathbb{R})$ is not connected (see (2) of VII.5.4), $\exp(M_n(\mathbb{R}))$ cannot be equal to $\mathrm{GL}_n(\mathbb{R})$.

Before calculating the image of $\exp$, we prove two lemmas.
Lemma. *Let $A, B \in M_n(\mathbb{R})$, and suppose that there exists $g \in \mathrm{GL}_n(\mathbb{C})$ such that $gAg^{-1} = B$. Then there exists $h \in \mathrm{GL}_n(\mathbb{R})$ such that $hAh^{-1} = B$.*

The lemma is actually true when we replace $\mathbb{C}/\mathbb{R}$ by any field extension, but the proof that we will give here only works for infinite fields. [19]

Solution. Let $g \in \mathrm{GL}_n(\mathbb{C})$ be such that $gAg^{-1} = B$, write $g = X + iY$ with $X, Y \in M_n(\mathbb{R})$. Then we have $XA + iYA = BX + iBY$, hence, as A and B are in $M_n(\mathbb{R})$, $XA = BX$ and $YA = YB$. If we knew that X or Y is invertible, we would be done, but this is not necessarily true. However, notice that, for every $t \in \mathbb{C}$, $(X+tY)A = B(X+tY)$. Consider the function $f : t \longmapsto \det(X+tY)$. This is a degree $\leq n$ polynomial (with coefficients in $\mathbb{R}$), and $f(i) = \det(g) \neq 0$. So f is not the zero polynomial, and so it cannot be identically 0 on $\mathbb{R}$ (because $\mathbb{R}$ is infinite), i.e. there exists $\lambda \in \mathbb{R}$ such that $f(\lambda) \neq 0$. Now let $h = X + \lambda Y$. We have $h \in M_n(\mathbb{R})$, $\det(h) = f(\lambda) \neq 0$ so h is even in $\mathrm{GL}_n(\mathbb{R})$, and $hA = Bh$.

□

Before stating the second lemma, we introduce some notation. For $A \in M_n(\mathbb{C})$, $\lambda \in \mathbb{C}$ and $r \in \mathbb{N}^*$, we set

$$w_{r,\lambda}(A) = \dim(\mathrm{Ker}(A - \lambda I_n)^r) - \dim(\mathrm{Ker}(A - \lambda I_n)^{r-1}).$$

It is easy to check (using the Jordan normal) that $w_{r,\lambda}(A)$ is the number of Jordan blocks of size $\geq r$ in the Jordan normal form of A.

For $\lambda \in \mathbb{C}$ and $r \geq 1$, we also denote by $J_r(\lambda) \in M_r(\mathbb{C})$ the Jordan block $\begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$.

Lemma. *Let $A \in M_n(\mathbb{C})$. Then A is similar to a matrix with real entries if and only if, for every $\lambda \in \mathbb{C}$ and every $r \in \mathbb{N}^*$, $w_{r,\lambda}(A) = w_{r,\bar{\lambda}}(A)$.*

Of course, the condition is empty for $\lambda \in \mathbb{R}$.

Solution. Suppose that $w_{r,\lambda}(A) = w_{r,\bar{\lambda}}(A)$ for every $\lambda \in \mathbb{C}$ and every $r \in \mathbb{N}^*$, and let's

[19] For the general case, see [28], proposition X.1.3; there is also a generalization of the lemma in the exercises of section X.1 of the same book.

show that A is similar to a real matrix. We may assume that A is in Jordan normal form. The condition says that, for every $\lambda \in \mathbb{C}$ and every $r \in \mathbb{N}^*$, if A has a Jordan block of the form $J_r(\lambda)$, then it must also have a Jordan block of the form $J_r(\overline{\lambda})$. So it is enough to show that every matrix of the form $\begin{pmatrix} J_r(\lambda) & 0 \\ 0 & J_r(\overline{\lambda}) \end{pmatrix}$, for $\lambda \in \mathbb{C}$ and $r \in \mathbb{N}^*$, is similar to a real matrix. But if B is the matrix of the previous sentence and $P = \begin{pmatrix} I_r & iI_r \\ iI_r & I_r \end{pmatrix}$, then a straightforward calculation shows that $P^{-1}BP \in M_{2r}(\mathbb{R})$.

Conversely, suppose that A is similar matrix, and let's show that $w_r(\lambda) = w_r(\overline{\lambda})$ for every $\lambda \in \mathbb{C}$ and $r \in \mathbb{N}^*$. We may assume that $M_n(\mathbb{R})$. Let $V = \mathbb{C}^r$, seen as a $\mathbb{R}$-vector space, and let T be the element of $\mathrm{End}(V)$ given by A. We denote by σ the $\mathbb{R}$-linear automorphism of V given by applying complex conjugation to all the coordinates. Then $T \circ \sigma = \sigma \circ T$, and $(\lambda \mathrm{id}_V) \circ \sigma = \sigma \circ (\overline{\lambda}\mathrm{id}_V)$ for every λ. So

$$w_{r,\overline{\lambda}}(A) = \tfrac{1}{2}\dim_{\mathbb{R}}(\mathrm{Ker}(\sigma \circ (T - \lambda \mathrm{id}_V)^r \circ \sigma^{-1})) - \tfrac{1}{2}\dim_{\mathbb{R}}(\mathrm{Ker}(\sigma \circ (T - \lambda \mathrm{id}_V)^{r-1} \circ \sigma^{-1})),$$

and this equal to $w_{r,\lambda}(A)$ because σ is an automorphism.

□

Let's come back to the problem.

Let X be the set of $g \in \mathrm{GL}_n(\mathbb{R})$ such that, for every $\lambda \in \mathbb{R}_{<0}$ and $r \in \mathbb{N}^*$, $w_r(\lambda)$ is even. We want to show that $X = \exp(M_n(\mathbb{R}))$. First let's see what happens when we exponentiate Jordan blocks. Let $\lambda \in \mathbb{C}$ and $r \in \mathbb{N}^*$. We have $J_r(\lambda) = \lambda I_r + N$, with $N \in M_r(\mathbb{R})$ a nilpotent matrix such that $N^{r-1} \neq 0$. So $e^{J_r(\lambda)} = e^{\lambda}e^N = e^{\lambda}(I_r + N')$, with $N' = N + \frac{1}{2}N^2 + \cdots + \frac{1}{(r-1)!}N^{r-1}$. The matrix N' is nilpotent, and we have $N'^{r-1} = N^{r-1} \neq 0$ (because $N^r = 0$), so $e^{J_r(\lambda)}$ is similar to $J_r(e^{\lambda})$. Using the interpretation of $w_{r,\lambda}(A)$ in terms of Jordan blocks in the Jordan normal form of A, this implies immediately that, for every $\lambda \in \mathbb{C}$ and $r \in \mathbb{N}^*$,

$$w_{r,\lambda}(e^A) = \sum_{\mu \in \mathbb{C} | e^{\mu} = \lambda} w_{r,\mu}(A).$$

In particular, if A is similar to a real matrix, then, for every $\lambda \in \mathbb{R}_{<0}$ and $r \in \mathbb{N}^*$, we have

$$w_{r,\lambda}(e^A) = \sum_{m \in \mathbb{Z}} w_{r,\log|\lambda| + i(2m+1)\pi}(A) = \sum_{m \geq 0} 2w_{r,\log|\lambda| + i(2m+1)\pi}(A),$$

so $w_{r,\lambda}(e^A)$ and $e^A \in X$.

Conversely, let's show that $X \subset \exp(M_n(\mathbb{R}))$. Let $g \in X$. We may assume that g is in Jordan normal form, so we have $g = \begin{pmatrix} g_1 & & 0 \\ & \ddots & \\ 0 & & g_m \end{pmatrix}$, with each g_j of the form $J_r(\lambda)$ with

$\lambda \in \mathbb{R}_{>0}$ or $\begin{pmatrix} J_r(\lambda) & 0 \\ 0 & J_r(\overline{\lambda}) \end{pmatrix}$ with $\lambda \in \mathbb{C} - \mathbb{R}_{\geq 0}$ It suffices to show that each of the g_j is the exponential of a real matrix, so we may assume that g is one of the g_j.

If $g = J_r(\lambda)$ with $\lambda \in \mathbb{R}_{>0}$, then $g = \lambda I_r + N$ with N nilpotent, and so $g = e^A$ with $A = \log(\lambda) \sum_{m \geq 1} (-1)^{m-1} \frac{1}{m} N^m$ by proposition VI.3.4.

If $g = \begin{pmatrix} J_r(\lambda) & 0 \\ 0 & J_r(\overline{\lambda}) \end{pmatrix}$ with $\lambda \in \mathbb{C} - \mathbb{R}_{\geq 0}$, choose $\mu \in \mathbb{C}$ such that $e^\mu = \lambda$ and let $A = \begin{pmatrix} J_r(\mu) & 0 \\ 0 & J_r(\overline{\mu}) \end{pmatrix}$. Note that $e^{\overline{\mu}} = \overline{e^\mu}$, so, by what we have seen above, e^A is similar to $\begin{pmatrix} J_r(\lambda) & 0 \\ 0 & J_r(\overline{\lambda}) \end{pmatrix} = g$. Also, by the second lemma, A is similar to a matrix $B \in M_{2r}(\mathbb{R})$. Then e^B and g are similar and they are both in $M_{2r}(\mathbb{R})$, so, by the first lemma, there exists $h \in \mathrm{GL}_{2r}(\mathbb{R})$ such that $he^Bh^{-1} = g$. Finally, we get $g = e^{hBh^{-1}}$, with $hBh^{-1} \in M_{2r}(\mathbb{R})$.

(3). Let's show that $\exp(\mathfrak{so}(n)) \subset \mathrm{SO}(n)$. Let $A \in \mathfrak{so}(n)$, then ${}^tA = -A$, so A and ${}^tA$ commute, so (by proposition VI.3.2 of chapter VI) $e^{{}^tA+A} = e^{{}^tA}e^A$. As ${}^tA + A = 0$ and $e^{{}^tA} = {}^te^A$, this gives $I_n = e^0 = {}^te^Ae^A$, i.e., $e^A \in \mathrm{O}(n)$. Also, $\det(e^A) = e^{\mathrm{Tr}A}$ (again by proposition VI.3.2 of chapter VI). But $0 = \mathrm{Tr}({}^tA + A) = 2\mathrm{Tr}(A)$, so $\mathrm{Tr}A = 0$ and $\det(e^A) = 1$. This shows that $e^A \in \mathrm{SO}(n)$.

Conversely, let $g \in \mathrm{SO}(n)$, and let's show that there exists $A \in \mathfrak{so}(n)$ such that $e^A = g$. We can find $h \in \mathrm{O}(n)$ such that $hgh^{-1} = \begin{pmatrix} g_1 & & 0 \\ & \ddots & \\ 0 & & g_m \end{pmatrix}$, with each g_j equal to either 1 or to $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$, $\theta \in \mathbb{R}$. As conjugating by a matrix of $\mathrm{O}(n)$ preserves $\mathfrak{so}(n)$, we may assume that $h = I_n$, and it suffices to treat the case $g = g_j$. If $g = 1$, we can take $A = 0$. If $g = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ with $\theta \in \mathbb{R}$, then we have $g = PhP^{-1}$ with $P = \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$ and $h = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$, so $g = e^A$ with $A = P \begin{pmatrix} i\theta & 0 \\ 0 & -i\theta \end{pmatrix} P^{-1}$, and it is easy to check that $A \in \mathfrak{so}(n)$.

(4). No. If A is the diagonal matrix with diagonal entries $-1, 1, \ldots, 1$, then $A \in \mathrm{O}(n)$ and $\det(A) = -1$. But, by proposition VI.3.2 of chapter VI, we know that, for every $B \in M_n(\mathbb{R})$, $\det(e^B) = e^{\mathrm{Tr}B} > 0$. So A cannot be the exponential of a real matrix.

□

VII.6.2 Kernel of the adjoint representation

Let G be a closed subgroup of $\mathrm{GL}_n(\mathbb{C})$ and $\mathfrak{g}$ be its Lie algebra. What is the kernel of the continuous group morphism $\mathrm{Ad} : G \to GL(\mathfrak{g})$?

Solution. If $g \in G$, then $\mathrm{Ad}(g)$ is the differential (see theorem VI.5.2 of chapter VI) of the continuous group morphism $\mathrm{Int}(g) : G \to G,\ h \longmapsto ghg^{-1}$. Also, we know (see the remarks below the theorem we just quoted) that $\mathrm{Ad}(g)$ determines $\mathrm{Int}(g)_{|G^0}$ (this is just because $\exp(\mathfrak{g})$ contains a neighborhood of 1 in G), so $\mathrm{Ad}(g) = 0$ if and only if $\mathrm{Int}(g)_{|G^0} = \mathrm{id}_{G^0}$. So the kernel of Ad is the centralizer of G^0 in G. If G is connected, this is just the center of G, but in general it could be bigger (for example if G is isomorphic to the direct product of a connected group and a noncommutative finite group).

□

VII.6.3 Lie algebras of compact groups

(1). Let G a compact closed subgroup of $\mathrm{GL}_n(\mathbb{C})$ and $\mathfrak{g} = \mathrm{Lie}(G)$. Prove that there exists an inner product $\langle ., \rangle$ on $\mathfrak{g}$ such that, for every $X, Y, Z \in \mathfrak{g}$,

$$\langle [X, Y], Z\rangle + \langle Y, [X, Z]\rangle = 0.$$

(2). Show that $\mathfrak{sl}_n(\mathbb{R})$ cannot be the Lie algebra of a compact closed subgroup of $\mathrm{GL}_n(\mathbb{C})$ if $n \geq 2$.

Solution.

(1). Consider the continuous representation of G on $\mathfrak{g}$ given by the map $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$. As G is compact, there exists (by theorem V.3.1.6 of chapter V)[20] an inner product $\langle ., .\rangle$ on $\mathfrak{g}$ that makes Ad a unitary representation. Let $X, Y, Z \in \mathfrak{g}$. Then, for every $t \in \mathbb{R}$, we have

$$\langle e^{tX} Y e^{-tX}, Z\rangle = \langle Y, e^{-tX} Z e^{tX}\rangle.$$

Taking the derivative of this equality and evaluating at $t = 0$ gives $\langle [X, Y], Z\rangle = \langle Y, [Z, W]\rangle$, which is what we wanted.

(2). Suppose that $\mathfrak{sl}_n(\mathbb{R})$ is the Lie algebra of a compact subgroup of $\mathrm{GL}_n(\mathbb{C})$, and choose an inner product $\langle ., .\rangle$ on $\mathfrak{sl}_n(\mathbb{R})$ satisfying the condition of (1). As $n \geq 2$, $\mathfrak{sl}_n(\mathbb{R})$ contains a copy of $\mathfrak{sl}_2(\mathbb{R})$, so we can find $E, F, H \in \mathfrak{sl}_n(\mathbb{R})$ such that $[E, F] = H$, $[H, E] = 2E$ and $[H, F] = -2F$. We have $\langle [E, H], E\rangle + \langle H, [E, E]\rangle = 0$, so $\langle -2E, E\rangle + \langle H, 0\rangle = 0$, hence $\langle E, E\rangle = 0$, which contradicts the fact that $\langle ., .\rangle$ is definite positive.

□

[20] This theorem applies to complex representations, but the exact same proof works for real representations.

VII.6.4 Some Lie algebras, and the adjoint representation

In this problem, k is a commutative ring (with unit). If V is a k-module, we write $\mathfrak{gl}(V) = \mathrm{End}_k(V)$. The *Lie bracket* $[.,.] : \mathfrak{gl}(V)^2 \to \mathfrak{gl}(V)$ is defined by $[X,Y] = XY - YX$. A *Lie subalgebra* of $\mathfrak{gl}(V)$ is a k-submodule $\mathfrak{g}$ that is stable by $[.,.]$.

If $V = k^n$, we also write $\mathfrak{gl}_n(k)$ for $\mathfrak{gl}(V)$.

(1). Let n be a positive integer, and let $J \in \mathfrak{gl}_n(k)$. Show that

$$\mathfrak{sl}_n(k) := \{X \in \mathfrak{gl}_n(k) | \mathrm{Tr}(X) = 0\}$$

and

$$\mathfrak{o}(J,k) := \{X \in \mathfrak{gl}_n(k) | XJ + JX^t = 0\}$$

are Lie subalgebras of $\mathfrak{gl}_n(k)$.

(2). Let A be a k-algebra (not necessarily associative). A k-linear map $\delta : A \to A$ is called a *derivation* if $\delta(ab) = a\delta(b) + \delta(a)b$, for every $a, b \in A$.

Show that $\mathrm{Der}(A)$ is a Lie subalgebra of $\mathfrak{gl}(A)$.

(3). Let $\mathfrak{g}$ be a Lie algebra, and consider the map $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ sending $X \in \mathfrak{g}$ to the linear endomorphism $Y \longmapsto [X,Y]$ of $\mathfrak{g}$. Show that ad is a morphism of Lie algebras, that $\mathrm{ad}(\mathfrak{g}) \subset \mathrm{Der}(\mathfrak{g})$ and that $\mathrm{ad}(\mathfrak{g})$ is an ideal of the Lie algebra $\mathrm{Der}(\mathfrak{g})$.

Solution. The proofs of (1) and (2) are easy calculations. The fact that ad is a morphism of Lie algebras and that $\mathrm{ad}(X)$ is a derivation for every $X \in \mathfrak{g}$ are both equivalent to the Jacobi identity in a straightforward way. Let's show that $\mathrm{ad}(\mathfrak{g})$ is an ideal of $\mathrm{Der}(\mathfrak{g})$. Let $X \in \mathfrak{g}$ and $\delta \in \mathrm{Der}(\mathfrak{g})$. Then, for every $Y \in \mathfrak{g}$,

$$[\delta, \mathrm{ad}(X)](Y) = -\mathrm{ad}(X)(\delta(Y)) + \delta(\mathrm{ad}(X)(Y)) = -[X, \delta(Y)] + \delta([X,Y]).$$

As δ is a derivation, $\delta([X,Y]) = [\delta(X), Y] + [X, \delta(Y)]$, so $\delta, \mathrm{ad}(X)](Y) = [\delta(X), Y]$, and finally $[\delta, \mathrm{ad}(X)] = \mathrm{ad}(\delta(X)) \in \mathrm{ad}(\mathfrak{g})$.

□

VII.6.5 Lie algebra of a linear algebraic group

Let k be a commutative ring. Fix a positive integer n and a family of polynomials $(P_\alpha)_{\alpha \in I}$ in $k[X_{ij}, 1 \leq i,j \leq n]$. We say that $(P_\alpha)_{\alpha \in I}$ defines an algebraic group G over k if, for every map of commutative rings $k \to k'$, the set $\mathrm{G}(k')$ of zeros of $(P_\alpha)_{\alpha \in I}$ in $\mathrm{GL}_n(k')$ is a subgroup of $\mathrm{GL}_n(k')$. (An element $g = (x_{ij})$ of $\mathrm{GL}_n(k')$ is called a zero of $(P_\alpha)_{\alpha \in I}$ if $P_\alpha(x_{ij}) = 0$ for every α.)

Examples of such algebraic groups are GL_n, SL_n, $\mathrm{O}(n)$ and Sp_{2n}, where, for every k' as above,

$$\mathrm{O}(n,k') = \{g \in \mathrm{GL}_n(k') | gg^t = I_n\}$$

and

$$\mathrm{Sp}_{2n}(k') = \{g \in \mathrm{GL}_{2n}(k') | gJg^t = I_{2n}\},$$

with $J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$.

Let $k[\varepsilon] = k \oplus k\varepsilon$, with the multiplication given by $\varepsilon^2 = 0$. Then the Lie algebra of G is by definition

$$\mathfrak{g} = \{X \in \mathfrak{gl}_n(k) | I_n + \varepsilon X \in \mathrm{G}(k[\varepsilon])\}.$$

(1). Show that $\mathfrak{g}$ is indeed a Lie subalgebra of $\mathfrak{gl}_n(k)$.

(2). If $\mathrm{G} = \mathrm{GL}_n$ (resp. SL_n, $\mathrm{O}(n)$, Sp_{2n}), show that $\mathfrak{g} = \mathfrak{gl}_n(k)$ (resp. $\mathfrak{sl}_n(k)$, $\mathfrak{o}_n(k) := \mathfrak{o}(I_n, k)$, $\mathfrak{sp}_{2n}(k) := \mathfrak{o}(J, k)$).

(3). Suppose that k is a field, and let A be a k-algebra. [21] Let G be a linear algebraic group over k, defined by a family of polynomials $(P_\alpha)_{\alpha \in I}$ in $k[X_{ij}, 1 \leq i, j \leq n]$. We denote by G_A the linear algebraic group over A defined by the same polynomials, now seen as polynomials with coefficients in A.

Show that the map $G(k[\varepsilon]) \to G(A[\varepsilon])$ induces an isomorphism of A-Lie algebras $(\mathrm{Lie}\, G) \otimes_k A \xrightarrow{\sim} \mathrm{Lie}\, G_A$.

(4). Suppose that $k = \mathbb{C}$, and let G be a linear algebraic group over k, defined by a family of polynomials $(P_\alpha)_{\alpha \in I}$ in $k[X_{ij}, 1 \leq i, j \leq n]$. Show that $G(\mathbb{C})$ is a closed subgroup of $\mathrm{GL}_n(\mathbb{C})$, and that we have $\mathrm{Lie}(G(\mathbb{C})) \subset \mathrm{Lie}\, G$.[22] (The first Lie algebra is the one defined in question (1), and the second Lie algebra is the one defined in definition VI.4.1 of chapter VI.)

Solution.

(1). There are three things to check : that $\mathfrak{g}$ is stable by multiplication by elements of k, that it's stable by addition and that it's stable by the bracket.

Let $X, Y \in \mathfrak{g}$ and $\lambda \in k$.

First, we have a (unique) k-algebra map $u : k[\varepsilon] \to k[\varepsilon]$ that sends ε to $\lambda\varepsilon$. It induces a map $u_* : G(k[\varepsilon]) \to G(k[\varepsilon])$. As $X \in \mathfrak{g}$, $I_n + \varepsilon X \in G(k[\varepsilon])$, and we have $u_*(I_n + \varepsilon X) = I_n + \varepsilon(\lambda X)$. So $\lambda X \in \mathfrak{g}$.

[21] We actually only need the fact that A is a flat k-algebra, which happens to be automatic if k is a field.
[22] This is actually an equality, but I couldn't figure out an elementary proof.

Second, as $G(k[\varepsilon])$ is a subgroup of $\mathrm{GL}_n(k[\varepsilon])$ and $I_n+\varepsilon X, I_n+\varepsilon Y \in G(k[\varepsilon])$, the product $(I_n+\varepsilon X)(I_n+\varepsilon Y) = I_n+\varepsilon(X+Y)$ is also in $G(k[\varepsilon])$. So $X+Y \in \mathfrak{g}$.

Finally, we consider $k' = k[\varepsilon]\otimes_k k[\varepsilon] = k[\varepsilon_1,\varepsilon_2]$, where $\varepsilon_1 = \varepsilon\otimes 1$ and $\varepsilon_2 = 1\otimes\varepsilon$. We have maps $u_1, u_2, u : k[\varepsilon] \to k'$ defined by $u_1(\varepsilon) = \varepsilon_1$, $u_2(\varepsilon) = \varepsilon_2$, and $u(\varepsilon) = \varepsilon_1\varepsilon_2 = \varepsilon\otimes\varepsilon$. These give maps $u_{1*}, u_{2*}, u_* : \mathrm{GL}_n(k[\varepsilon]) \to \mathrm{GL}_n(k')$. We have $g_1 = u_{1*}(I_n+\varepsilon X)$, $g_2 = u_{2*}(I_n+\varepsilon Y) \in G(k')$. Using the fact that $(I_n+\varepsilon X)^{-1} = (I_n-\varepsilon X)$ (and similarly for $I_n+\varepsilon Y$), we see that

$$g_1 g_2 g_1^{-1} g_2^{-1} = I_n + \varepsilon_1\varepsilon_2[X,Y] = u_*(I_n+\varepsilon[X,Y]) \in G(k').$$

But because u is injective, an element g of $GL_n(k[\varepsilon])$ is in $G(k[\varepsilon])$ if and only if $u_*(g)$ is in $G(k')$. So $I_n+\varepsilon[X,Y] \in G(k[\varepsilon])$, and $[X,Y] \in \mathfrak{g}$.

(2). The result is clear for GL_n.

Let's calculate $\mathrm{Lie}(\mathrm{SL}_n)$ (where $\mathrm{Lie}(G)$ means "the Lie algebra of G"). Let $X \in \mathfrak{gl}_n(k)$, and let $\chi_X \in k[t]$ be its characteristic polynomial ($\chi_X(t) = \det(1-tX)$). Then

$$\det(I_n+\varepsilon X) = \chi_X(-\varepsilon) = 1+\varepsilon\mathrm{Tr}(X),$$

because $\varepsilon^2 = 0$. So

$$I_n+\varepsilon X \in \mathrm{SL}_n(k[\varepsilon]) \Leftrightarrow \det(I_n+\varepsilon X) = 1 \Leftrightarrow \mathrm{Tr}(X) = 0 \Leftrightarrow X \in \mathfrak{sl}_n(k).$$

For $\mathrm{Lie}(\mathrm{O}(n))$, note that $(I_n+\varepsilon X)^t = I_n+\varepsilon X^t$, so

$$(I_n+\varepsilon X)(I_n+\varepsilon X)^t = I_n+\varepsilon(X+X^t).$$

So obviously $I_n+\varepsilon X \in \mathrm{O}(n)(k[\varepsilon])$ if and only if $X+X^t = 0$, ie $X \in \mathfrak{so}_n(k)$.

The calculation for $\mathfrak{sp}_n(k)$ is the same, mutatis mutandis.

(3). If k is any commutative ring, $f \in k[t_1,\dots,t_m]$ is a polynomial and $a = (a_1,\dots,r_m) \in k^m$, denote by $df(a)$ the k-linear map from k^m to k given by

$$(x_1,\dots,x_m) \longmapsto \sum_{r=1}^{m} x_r \frac{\partial f}{\partial t_r}(a_1,\dots,a_m).$$

Then we have, for all $x_1,\dots,x_m \in k$,

$$f(a_1+\varepsilon x_1,\dots,a_m+\varepsilon x_m) = f(a_1,\dots,a_n) + \varepsilon df(a)(x_1,\dots,x_m).$$

(This is easy to check for monomials, and f is a linear combination of monomials.)

Applying this to the family of polynomials $(P_\alpha)_{\alpha\in I}$ defining the algebraic group G, we get that $\mathrm{Lie}\,G$ is, by definition, the intersection of the kernels of all the linear forms $dP_\alpha(1)$

on $M_n(k)$. This is true without any assumption on k. Now, if A is a k-algebra, we always have a natural map

$$\left(\bigcap_{\alpha\in I}\mathrm{Ker}(dP_\alpha(1):M_n(k)\to k)\right)\otimes_k A\to\left(\bigcap_{\alpha\in I}\mathrm{Ker}(dP_\alpha(1):M_n(A)\to A)\right),$$

i.e., a map $(\mathrm{Lie}\,G)\otimes_k A\to\mathrm{Lie}\,G_A$, but this map is not an isomorphism in general, because taking the tensor product by A is not an exact operation. If for example k is a field, there is no problem. More generally, if A is a flat k-algebra,[23] then the above map is an isomorphism.

(4). We know that $G(\mathbb{C})$ is a subgroup of $\mathrm{GL}_n(\mathbb{C})$ (by definition of a linear algebraic group), and it is closed because it is the set of zeros of a family of continuous functions from $\mathrm{GL}_n(\mathbb{C})$ to $\mathbb{C}$ (the functions given by the P_α).

Let $X\in\mathfrak{gl}_n(\mathbb{C})$ and $\alpha\in I$, and define $c:\mathbb{R}\to\mathbb{C}$ by by $c(t)=P_\alpha(e^{tX})$. Then we have $c'(0)=dP_\alpha(1)(X)$. If $X\in\mathrm{Lie}\,G(\mathbb{C})$, then $e^{tX}\in G$ for every t, so c is identically 0 and $c'(0)=0$. This shows that

$$\mathrm{Lie}\,G(\mathbb{C})\subset\bigcap_{\alpha\in I}\mathrm{Ker}(dP_\alpha(1)).$$

By the proof of (3), the right-hand side is $\mathrm{Lie}\,G$.

□

VII.6.6 Group of automorphisms of a k-algebra

This problem uses the definitions of problems VII.6.4(2) and VII.6.5.

Let A be a k-algebra (not necessarily associative), and assume that A is free of finite type as a k-module. For every map of commutative rings $k\to k'$, let $\mathrm{Aut}(A)(k')$ be the subgroup of $\mathrm{GL}(A\otimes_k k')$ whose elements are k'-algebra automorphisms.

(1). Show that $\mathrm{Aut}(A)$ is an algebraic group over k, i.e., identify A to k^n (as a k-module) by choosing a basis of A and show that there exists a family of polynomials in $k[X_{ij},1\le i,j\le n]$ that defines an algebraic group over k and such that $\mathrm{Aut}(A)(k')$ is the set of zeros of that family.

(2). Show that the Lie algebra of $\mathrm{Aut}(A)$ is $\mathrm{Der}(A)$.

Solution.

[23] Which meabs exactly that $\otimes_k A$ preserves exact sequences.

(1). Let $(e_1, \dots, e_n)$ be a basis of A as a k-module, and write, for every $i, j \in \{1, \dots, n\}$,

$$e_i e_j = \sum_{k=1}^{n} \alpha_{ijk} e_j,$$

with $\alpha_{ijk} \in k$. We also write $1_A = \sum_{i=1}^{n} \beta_i e_i$, where 1_A is the unit of A and the β_i are in k.

Let $k \to k'$ be a map of commutative rings, and let $g = (g_{ij}) \in \mathrm{GL}(A) \simeq \mathrm{GL}_n(k')$ (using the fixed basis). Then g is in $\mathrm{Aut}(A)(k')$ if and only g sends 1_A to 1_A and preserves the product of A. Because g is k'-linear, to check the second condition, we only need to check that $g(e_i e_j) = g(e_i)g(e_j)$, for every i, j. We have

$$g(e_i) = \sum_{j=1}^{n} g_{ij} e_j$$

and

$$g(1_A) = \sum_{i=1}^{n} \alpha_i g(e_i) = \sum_{i,j=1}^{n} \alpha_i g_{ij} e_j.$$

So we see that $g \in \mathrm{Aut}(A)(k')$ if and only if the following conditions are satisfied :

(A) For every $i \in \{1, \dots, n\}$,

$$\alpha_i = \sum_{j=1}^{n} \alpha_j g_{ji}.$$

(B) For every $i, j, k \in \{1, \dots, n\}$,

$$\sum_{a,b=1}^{n} g_{ia} g_{jb} \alpha_{abk} = \sum_{c=1}^{n} \alpha_{ijc} g_{ck}.$$

(These conditions are supposed to correspond to the conditions $g(1_A) = 1_A$ and $g(e_i)g(e_j) = g(e_i e_j)$.)

As (A) and (B) are obviously polynomial conditions in the entries of g, this gives the result.

(2). We identify $\mathrm{GL}(A)$ and GL_n as above. Let $X \in \mathfrak{gl}_n(k)$. Then $X \in \mathrm{Lie}(\mathrm{Aut}(A))$ if and only if $g := \mathrm{id}_A + \varepsilon X \in \mathrm{Aut}(A)(k[\varepsilon])$, that is, if and only if $g(1_A) = g(1_A)$ and $g(ab) = g(a)g(b)$ for every $a, b \in A \otimes_k k[\varepsilon]$.

The first condition is equivalent to $X(1_A) = 0$. For the second condition, we write $a = a_1 + \varepsilon a_2$, $b = b_1 + \varepsilon b_2$, with $a_1, a_2, b_1, b_2 \in A$. Then $g(a) = a_1 + \varepsilon(a_2 + X(a_1))$, $g(b) = b_1 + \varepsilon(b_2 + X(b_1))$ and $ab = a_1 b_1 + \varepsilon(a_1 b_2 + a_2 b_1)$. So

$$g(a)g(b) = a_1 b_1 + \varepsilon(a_2 b_1 + X(a_1) b_1 + a_1 b_2 + a_1 X(b_1)$$

and

$$g(ab) = a_1a_2 + \varepsilon(a_1b_2 + a_2b_1 + X(a_1b_1)).$$

So we have

$$g(a)g(b) = g(ab) \Leftrightarrow X(a_1b_1) = a_1X(b_1) + X(a_1)b_1.$$

From this, it's now obvious that $\mathrm{id}_A + \varepsilon X \in \mathrm{Aut}(A)(k[\varepsilon])$ if and only X is a derivation. (Notice we have $\delta(1_A) = 1_A$ for any derivation $\delta : A \to A$, because $\delta(1_A) = \delta(1_A^2) = \delta(1_A) + \delta(1_A)$.)

□

VII.6.7 Symmetric algebra and symmetric powers of a representation

Let k be a commutative ring.

Let V be a k-module and $n \in \mathbb{Z}_{\geq 0}$. Remember that we write T^nV or $V^{\otimes n}$ for the n-fold tensor product of V by itself (over k); by convention, if $n = 0$, $T^0V = k$. By definition of the tensor product, for every other k-vector space W, $\mathrm{Hom}_k(T^nV, W)$ is the space of multilinear maps from V^n to W.

Let I_n be the subspace of T^nV generated by all $v_1 \otimes \cdots \otimes v_n - v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)}$, for $v_1, \ldots, v_n \in V$ and $\sigma \in \mathfrak{S}_n$. We set $S^nV = T^nV/I_n$ and call it the nth symmetric power of V. (If $n = 0$, $S^0V = k$.)

(1). Show that the multiplication $T^nV \otimes T^mV \to T^{n+m}V$ defined in class sends $T^nV \otimes_k I_m$ and $I_n \otimes_k T^mV$ to I_{n+m}. So we get a k-algebra structure on $S^*V := \bigoplus_{n\geq 0} S^nV$. Show that this k-algebra is commutative. (This is called the *symmetric algebra* of V.)

Now suppose that V is a free k-module of finite rank and choose a basis $(e_1, \ldots, e_d)$ of V.

(2). Find a basis of T^nV and calculate $\dim_k(T^nV)$.

(3). Find a basis of S^nV.

Stop assuming that V is free of finite rank.

(4). If W is another k-vector space, show that $\mathrm{Hom}_k(S^nV, W) \subset \mathrm{Hom}_k(T^nV, W)$ is the subspace of symmetric multilinear maps from V^n to W.

(5). Let $\mathfrak{g}$ be a Lie algebra over k and $u : \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation of $\mathfrak{g}$ on V. Consider the induced representation of T^nV. Show that I_n is stable by $\mathfrak{g}$. (If we have representations of $\mathfrak{g}$ on V_1 and V_2, the action of $\mathfrak{g}$ on $V_1 \otimes_k V_2$ is given by $X(v \otimes w) = (Xv) \otimes w + v \otimes (Xw)$, for every $X \in \mathfrak{g}$, $v \in V_1$ and $w \in V_2$.)

Hence we get a representation of $\mathfrak{g}$ on S^nV, called the *nth symmetric power* of the representation (V, u).

Solution.

(1). Let $v_1, \dots, v_n, w_1, \dots, w_m \in V$ and let $\sigma \in \mathfrak{S}_m$. Define $x_1, \dots, x_{n+m} \in V$ and $\tau \in \mathfrak{S}_{n+m}$ by $x_i = v_i$ $\tau(i) = i$ if $1 \leq i \leq n$, and $v_{n+j} = w_j$ and $\tau(n+j) = n + \sigma(j)$ if $1 \leq j \leq m$. Then

$$v_1 \otimes \cdots \otimes v_n \otimes (w_1 \otimes \cdots \otimes w_m - w_{\sigma(1)} \otimes \cdots \otimes w_{\sigma(m)}) = x_1 \otimes \cdots \otimes x_{n+m} - x_{\tau(1)} \otimes \cdots \otimes x_{\tau(n+m)}.$$

So the multiplication sends $T^nV \otimes I_m$ to I_{n+m}.

The case of $I_m \otimes T^nV$ is similar.

(2). An easy induction on n show that a basis of T^nV is given by the $e_{i_1} \otimes \dots e_{i_n}$, for $(i_1, \dots, i_n) \in \{1, \dots, d\}^n$. So $\dim_k(T^nV) = d^n$.

(3). I claim that a basis of S^nV is given by the $e_1^{n_1} \dots e_d^{n_d}$, with $n_1, \dots, n_d \in \mathbb{Z}_{\geq 0}$ such that $n_1 + \cdots + n_d = n$.

Indeed, this family is clearly generating by (2). (If $i_1, \dots, i_n$ are any elements of $\{1, \dots, d\}$, choose $\sigma \in \Sigma_n$ such that $i_{\sigma(1)} \leq \cdots \leq i_{\sigma(n)}$. Then, in S^nV, $e_1 \dots e_{i_n} = s_{i_{\sigma(1)}} \dots e_{i_{\sigma(n)}}$ is the form $e_1^{n_1} \dots e_d^{n_d}$.)

Showing that this family is free is easier after we know the result of (d), so let's assume we do. Let $(e_1^*, \dots, e_d^*) \in V^*$ be the dual basis of $(e_1, \dots, e_d)$. Let $n_1, \dots, n_d \in \mathbb{Z}_{\geq 0}$ such that $n_1 + \cdots + n_d = n$. For $1 \leq r \leq d$ and $n_1 + \cdots + n_r \leq i \leq n_1 + \cdots + n_{r+1} - 1$, let $f_i = e_i^*$. Define a multilinear map $f : V^n \to k$ by

$$f(v_1, \dots, v_n) = \sum_{\sigma \in \Sigma_n} \prod_{i=1}^{n} f_i(v_{\sigma(i)}).$$

Then f is obviously symmetric, so it gives a map $S^nV \to k$, that we will also denote by f. If $m_1, \dots, m_d \in \mathbb{Z}_{\geq 0}$ and $m_1 + \cdots + m_d = n$, we have

$$f(e_1^{m_1} \dots e_d^{m_d}) = \begin{cases} 1 & \text{if } n_i = m_i \ \forall i \\ 0 & \text{otherwise.} \end{cases}$$

This shows that the family given above is linearly independent.

(4). Let $f \in \mathrm{Hom}_k(T^nV, W)$. Then f factors through S^nV if and only if, for every $v_1, \dots, v_n \in V$ and $\sigma \in \Sigma_n$, $f(v_1, \dots, v_n) = f(v_{\sigma(1)}, \dots, v_{\sigma(n)})$. This is the same as saying that f is symmetric.

(5). If $v_1, \dots, v_n \in V$ and $X \in \mathfrak{g}$, then

$$X(v_1 \otimes \cdots \otimes v_n) = \sum_{i=1}^{n} v_1 \otimes \cdots \otimes v_{i-1} \otimes (Xv_i) \otimes v_{i+1} \otimes \cdots \otimes v_n.$$

So $X(v_1 \otimes \cdots \otimes v_n - v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)}) =$

$$\sum_{i=1}^{n}(v_1 \otimes \cdots \otimes v_{i-1} \otimes (Xv_i) \otimes v_{i+1} \otimes \cdots \otimes v_n - v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(i-1)} \otimes (Xv_{\sigma(i)}) \otimes v_{\sigma(i+1)} \otimes \cdots \otimes v_{\sigma(n)}) \in I_n.$$

□

VII.6.8 Symmetric algebra and polynomial functions

We use the notation of problem VII.6.7, and assume that k is an infinite field.

Let V be a finite-dimensional k-vector space. We say that a map $f : V \to k$ is *polynomial* if, for every basis $e_1, \ldots, e_n$ of V, there exists $P \in k[X_1, \ldots, X_n]$ such that $f(\lambda_1 e_1 + \cdots + \lambda_n e_n) = P(\lambda_1, \ldots, \lambda_n)$ for every $\lambda_1, \ldots, \lambda_n \in k$. We denote by $k[V]$ the algebra of polynomials functions from V to k. Let $V^* = \mathrm{Hom}(V, k)$.

Show that the map $\varphi : T^\bullet V^* \to \mathrm{Map}(V, k)$ that sends $u_1 \otimes \cdots \otimes u_n \in T^n V^*$ to the map $x \mapsto u_1(x) \ldots u_n(x)$ is well-defined, has image contained in $k[V]$, and induces a k-algebra isomorphism $S^\bullet V^* \xrightarrow{\sim} k[V]$.

Solution. The map φ is well-defined, because the map $(V^*)^n \to \mathrm{Map}(V, k)$ sending $(u_1, \ldots, u_n)$ to $x \mapsto u_1(x) \ldots u_n(x)$ is multilinear. For every $u \in V^*$, the element of $\mathrm{Map}(V, k)$ is a polynomial map. As the image of φ is the k-subalgebra generated by these elements, it is contained in $k[V]$. As $k[V]$, φ factors through a map $\psi : S^\bullet V^* \to k[V]$.

Now let's fix a basis $(e_1, \ldots, e_n)$, and let $(e_1^*, \ldots, e_n^*)$ be the dual basis. By question (3) of problem VII.6.7, the elements $(e_1^*)^{m_1} \ldots (e_n^*)^{m_n}$, with $m_1, \ldots, m_n \in \mathbb{N}$, form a basis of $S^\bullet V^*$. We denote by $\psi' : S^\bullet V^* \to k[X_1, \ldots, X_n]$ sending each $(e_1^*)^{m_1} \ldots (e_n^*)^{m_n}$ to $X_1^{m_1} \ldots X_n^{m_n}$; this is clearly an isomorphism of k-algebras. We denote by $\mathrm{ev} : k[X_1, \ldots, X_n] \to k[V]$ the map sending $P \in k[X_1, \ldots, X_n]$ to the function $\lambda_1 e_1 + \cdots + \lambda_n e_n) \longmapsto P(\lambda_1, \ldots, \lambda_n)$. This is a morphism of k-algebra, and it is surjective by definition of $k[V]$. Also, we have $\psi = \mathrm{ev} \circ \psi'$, because this two maps are k-algebra maps which are equal on the generators $e_1^*, \ldots, e_n^*$ of $S^\bullet V^*$. So we just have to show that ev is injective. (Note that we have not yet used the hypothesis on the cardinality of k. Now it will become important.)

We want to prove the following fact : If $P \in k[X_1, \ldots, X_n]$ is nonzero, then there exists $\lambda_1, \ldots, \lambda_n \in k$ such that $P(\lambda_1, \ldots, \lambda_n) \neq 0$. We do an induction on n. If $n = 1$, this just follows from the fact that a nonzero polynomial has finitely many roots (and that k is infinite). Suppose that $n \geq 2$. Let $P \in k[X_1, \ldots, X_n] - \{0\}$, and write $P = P_0 + P_1 X_n + \cdots + P_d X_n^d$, with $P_0, \ldots, P_d \in k[X_1, \ldots, X_{n-1}]$ and $P_d \neq 0$. By the induction hypothesis, there exists $\lambda_1, \ldots, \lambda_{n-1}$ such that $P_d(\lambda_1, \ldots, \lambda_{n-1}) \neq 0$. Then $Q := P_0(\lambda_1, \ldots, \lambda_{n-1}) + P_1(\lambda_1, \ldots, \lambda_{n-1}) X_n + \cdots + P_d(\lambda_1, \ldots, \lambda_{n-1}) X_n^d \in k[X_n]$ is nonzero, so, by the first step of the induction, there exists $\lambda_n \in k$ such that $Q(\lambda_n) = P(\lambda_1, \ldots, \lambda_n) \neq 0$.

□

VII.6.9 Some representations of $\mathfrak{sl}_2(k)$

By definition, the *standard representation* of $\mathfrak{g} := \mathfrak{sl}_2(k)$ is the inclusion $\mathfrak{g} \subset \mathfrak{gl}_2(k)$. It's a representation of $\mathfrak{g}$ on $V := k^2$. For every $n \geq 0$, we write $W_{n+1} = S^n V$ and consider the symmetric power representation $\mathfrak{g}$ on this space.

(1). Show that the three elements $e := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $f := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ and $h := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ form a basis of $\mathfrak{g}$, and calculate their Lie brackets.

(2). Find a basis of W_{n+1} and write the action of e, f and h in that basis.

(3). Show that W_{n+1} is an irreducible representation of $\mathfrak{g}$ if $\mathrm{char}(k) = 0$ or $n < \mathrm{char}(k)$.

(4). If $\mathrm{char}(k) > 0$ and $n = \mathrm{char}(k)$, show that W_n is not irreducible.

(5). (*) What happens if $n > \mathrm{char}(k)$?

Solution.

(1). It's obvious that (e, f, h) is a basis of $\mathfrak{g}$. We have $[e, f] = h$, $[h, e] = 2e$ and $[h, f] = -2f$.

(2). Let (v_1, v_2) be the standard basis of k^2. Then, by problem VII.6.7(3) a basis of W_{n+1} is $(v_1^n, v_1^{n-1}v_2, \ldots, v_2^n)$. We have $ev_1 = 0$, $ev_2 = v_1$, $fv_1 = v_2$, $fv_2 = 0$, $hv_1 = v_1$ and $hv_2 = -v_2$. So

$$e(v_1^i v_2^{n-i}) = (n-i)v_1^{i+1}v_2^{n-(i+1)},$$
$$f(v_1^i v_2^{n-i}) = iv_1^{i-1}v_2^{n-(i-1)}$$

and

$$h(v_1^i v_2^{n-i}) = (2i-n)v_1^i v_2^{n-i}.$$

(3). Let V be a nonzero $\mathfrak{g}$-invariant subspace of W_{n+1}. We want to show that $V = W_{n+1}$. Let $v \in V - \{0\}$, and write $v = \sum_{i=i_0}^n a_i v_1^i v_2^{n-i}$, with $a_i \in k$ and $a_{i_0} \neq 0$. Then

$$e^{n-i_0}v = (n-i_0)!a_{i+0}v_2^n.$$

By the assumption on $\mathrm{char}(k)$, $(n-i_0)!$ is invertible in k, so $v_2^n \in V$. For every $i \in \{0, \ldots, n\}$, $f^i v_2^n = i! v_1^i v_2^{n-i}$. As $i! \in k^\times$, $v_1^i v_2^{n-i} \in V$. So $V = W_{n+1}$.

(4). Suppose that $p = \mathrm{char}(k) > 0$. Then $ev_2^p = fv_2^p = hv_2^p = 0$, so kv_2^p is a nonzero subrepresentation of W_{p+1}, and W_{p+1} is not irreducible.

□

VII.6.10 Representations of $\mathfrak{sl}_2(k)$ in characteristic 0

Suppose that k is algebraically closed and $\mathrm{char}(k) = 0$, and let $\mathfrak{g} = \mathfrak{sl}_2(k)$. Let (W, u) be a representation of $\mathfrak{g}$ on a k-vector space. We use the notation of problem VII.6.9.

(1). For every $a \in k$, let $W(a) \subset W'(a)$ be the a-eigenspace and the generalized a-eigenspace of $u(h)$ on W (that is, $W'(a) = \bigcup_{n\geq 1} \mathrm{Ker}((u(h) - a\mathrm{id}_W)^n)$). Show that, for every $a \in k$, $u(e)$ sends $W(a)$ (resp. $W'(a)$) to $W(a+2)$ (resp. $W'(a+2)$) and $u(f)$ sends $W(a)$ (resp. $W'(a)$) to $W(a-2)$ (resp. $W'(a-2)$).

(2). Let $v \in W$ be such that $u(e)v = 0$, and set $v_k = u(f)^k v$ for every $k \geq 0$. Find (explicit) polynomials $P_{l,k}(t) \in \mathbb{Z}[t]$ (for $0 \leq l \leq k$) such that, if $0 \leq l \leq k$, $u(e)^l v_k = P_{l,k}(u(h))e_{k-l}$, and such that $\deg(P_{l,k}) = l$.

From now on, we assume that $\dim_k W < \infty$ et $W \neq 0$.

(3). For every $a \in k$, show that we have $u(e)^N(W'(a)) = u(f)^N(W'(a)) = 0$ for N big enough.

(4). For every $a \in k$, show that $u(h)$ is diagonalizable on $W'(a)$ (so that $W(a) = W'(a)$), and that $a \in \mathbb{Z}_{\geq 0}$ if $W(a) \neq 0$.

(5). Find an eigenvector v of $u(h)$ in W such that $u(e)v = 0$, and let $a \in \mathbb{Z}_{\geq 0}$ be the eigenvalue of v. For every $d \geq 0$, let $v_d = u(f)^d v$, and let V be the subspace of W generated by $(v_0, v_1, \dots)$. Show that V is a subrepresentation of W, that it is of dimension $a + 1$, and that it is isomorphic to the representation W_{a+1} of problem VII.6.9.

In particular, the W_{n+1} are (up to isomorphism) the only irreducible representations of $\mathfrak{g}$.

Solution.

(1). Let U be the universal enveloping algebra of $\mathfrak{g}$, and extend u to a k-algebra map $U \to \mathrm{End}_k(W)$. Note that, *in* U,

$$he = [h, e] + eh = 2e + eh = e(h + 2),$$

So, for every $n \in \mathbb{Z}_{\geq 0}$,

$$(h - (a + 2))^n e = e(h - a)^n.$$

Applying u gives

$$(u(h) - (a + 2)\mathrm{id}_V)^n u(e) = u(e)(u(h) - a)\mathrm{id}_V)^n,$$

and hence

$$u(e)\,\mathrm{Ker}(u(h) - a\mathrm{id}_V)^n \subset \mathrm{Ker}(u(h) - (a + 2)\mathrm{id}_V)^n.$$

This shows that $u(e)$ sends $W(a)$ to $W(a - 2)$ and $W'(a)$ to $W'(a - 2)$.

The proof for $u(f)$ is similar, starting with the fact that $hf = f(h - 2)$ in U.

(2). We can take $P_{0,k} = 1$.

Let's construct the $P_{1,k}$ by induction on k. If $k = 0$, then $u(e)v_k = u(e)v = 0$, so we can take $P_{1,0} = 0$. Suppose that we know that $P_{1,k}$ exists, with $k \geq 0$. In U, we have

$$ef^{k+1} = (ef)f^k = ([e, f] + fe)f^k = (h + fe)f^k.$$

Also, we have seen in the proof of (1) that $fh = (h + 2)f$, so, for every $P \in \mathbb{Z}[t]$, $fP(h) = P(h + 2)f$. Similarly, $eP(h) = P(h - 2)e$. Hence

$$u(e)v_{k+1} = u(e)u(f)^{k+1}v = (2u(h) + u(f)u(e))v_k.$$

Using the induction hypothesis gives

$$u(e)v_{k+1} = 2u(h)v_k + u(f)P_{1,k}(u(h))v_{k-1} = (u(h) + P_{1,k}(u(h) + 2))v_k$$

(with the convention that $v_{-1} = 0$). So we can take $P_{1,k+1} = t + P_{1,k}(t + 2)$. We see easily that this gives $P_{1,k}(t) = k(t + k - 1)$, for every $k \geq 1$.

Now fix k and let's construct $P_{l,k}$ by induction on l. We gave already done the cases $l = 0$ and $l = 1$, so let's assume that $1 \leq l \leq k - 1$ and that we have shown the existence of $P_{l,k}$. We have

$$u(e)^{l+1}v_k = u(e)P_{l,k}(u(h))v_{k-l} = P_{l,k}(u(h)-2)u(e)v_{k-l} = P_{l,k}(u(h)-2)P_{1,k-l}(u(h))v_{k-(l+1)}.$$

So we can take $P_{l+1,k}(t) = P_{l,k}(t - 2)P_{1,k-l}(t)$.

Unpacking the induction formula above gives

$$P_{l,k}(t) = P_{1,k-l+1}(t)P_{1,k-l+2}(t - 2) \dots P_{1,k-1}(t - 2(l - 1)),$$

ie

$$P_{l,k}(t) = (k - 1)(k - 2) \dots (k - l + 1)(t + k - l)(t + k - l - 1) \dots (t + k - l - (l - 1)).$$

This obviously has degree l.

(3). By (1), we know that $u(e)^N W'(a) \subset W'(a - 2N)$ and $u(f)^N W'(a) \subset W'(a + 2N)$. As $\dim_k W < \infty$, the endomorphism $u(h)$ of W has only finitely many eigenvalues, so there are only finitely many $b \in k$ such that $W'(b) \neq 0$. Hence, if N is big enough, $W'(a - 2N) = W'(a + 2N) = 0$, which proves the claim.

(4). We reason by induction on the smallest $i \in \mathbb{Z}_{\geq 1}$ such that $W'(a + 2i) = 0$.

If $i = 1$, then $W'(a + 2) = 0$, so $u(e)W'(a) = 0$ by (1). By (3), there exists $N \geq 0$ such that $u(f)^N W'(a) = 0$. by (2), for every $v \in W'(a)$,

$$0 = u(e)^N u(f)^N v = P_{N,N}(u(h))v.$$

So the minimal polynomial of $u(h)$ on $W'(a)$ divides $P_{N,N}(t)$. As $P_{N,N}(t)$ has simple roots, $u(h)$ is semisimple on $W'(a)$. As all roots of $P_{N,N}(t)$ are $0, 1, \dots, N-1$, $a \in \{0, \dots, N-1\}$, and in particular $a \in \mathbb{Z}$.

Now assume that we know the result for $i \geq 1$ and let's prove it for $i+1$. As $W'(a+2+2i) = 0$, we know by the induction hypothesis that $u(h)$ is semisimple on $W'(a+2)$. So, for every $v \in W'(a)$,

$$(a+2)u(e)v = u(h)u(e)v = u(e)(u(h)+2)v,$$

ie $u(e)(u(h)-a)v = 0$.

(5). Choose $a \in \mathbb{Z}$ such that $W(a) \neq 0$ and $W(a+2) = 0$, and let v be any nonzero element of $W(a)$. Then $u(e)v \in W(a+2)$, so $u(e)v = 0$. Let's show that the resulting V is a subrepresentation of W. First, V is obvioously stable by $u(e)$. Second, for every $d \in \mathbb{Z}_{\geq 0}$, $v_d \in W(a-2d)$, so v_d is an eigenvector of $u(h)$. So V is stable by $u(h)$. Finally, let $d \in \mathbb{Z}_{\geq 0}$. If $d = 0$, $u(e) = 0$, If $d \geq 1$, by (2), $u(e)v_d = P_{1,d}(u(h))v_{d-1} \in kv_{d-1} \subset V$ as v_{d-1} is an eigenvector of $u(h)$. So V is stable by $u(e)$, $u(f)$ and $u(h)$.

Let d be the biggest integer such that $v_d \neq 0$. Then $\dim_k V = d+1$, because $(v_0, \dots, v_d)$ is a basis of V (this family is generating, and it's free because it's made up of eigenvectors of $u(h)$ with pairwise different eigenvalues). Also, $u(f)^{d+1}v = v_{d+1} = 0$, so $P_{d+1,d+1}(u(h))v = 0$, so a is a root of $P_{d,d}(t)$, so $a \in \{0, \dots, d\}$. Suppose that $a \leq d-1$. Then $v_{a+1} \neq 0$, and we have

$$u(e)v_{a+1} = P_{1,a+1}(u(h))v_a = (a+1)(u(h)+a)v_a = 0,$$

because $v_a \in W(a-2a) = W(-a)$. Applying (2) again, we see that, if $N \geq 0$ is such that $u(f)^N = 0$, then

$$0 = u(e)^N u(f)^N v_d = P_{N,N}(u(h))v_d = P_{N,N}(-a)v_d.$$

But the roots of $P_{N,N}$ are $0, 1, \dots, N-1$, so $P_{N,N}(-a) \neq 0$, which gives a contradiction. So $a = d$.

We now consider the map $\varphi : V \to W_{d+1}$ sending $v_i \in V$ to $d(d-1)\dots(d-i+1)v_1^{d-1}v_2^i \in W_{d+1}$. (Sorry about the awful notation.) Using the fact that $u(f)v_i = v_{i+1}$, $u(h)v_i = (d-2i)v_i$ and $u(e)v_i = i(d-(i-1))v_{i-1}$ by (2) (with the convention $v_{-1} = 0$), we see that this is an isomorphism of representations of $\mathfrak{g}$.

□

VII.6.11 The Jacobson-Morozov theorem (for $\mathfrak{gl}_n(k)$)

We still assume that k is algebraically closed of characteristic 0. Let W be a finite-dimensional k-vector space, and let $N \in \mathrm{End}_k(W)$ be nilpotent. Show that there exists a unique semisimple representation $u : \mathfrak{sl}_2(k) \to \mathfrak{gl}(W)$ of $\mathfrak{sl}_2(k)$ on W such that $N = u(e)$.

(Actually, all finite-dimensional representations of $\mathfrak{sl}_2(k)$ are semisimple by corollary VI.8.4 of chapter VI, so the semisimplicity hypothesis is not necessary.)

Solution. Let $u : \mathfrak{sl}_2(k) \to \mathfrak{gl}(W)$ be a semisimple representation such that $N = u(e)$, and let $W = V_1 \oplus \cdots \oplus V_r$ be the decomposition of W into irreducible subrepresentations of $\mathfrak{sl}_2(k)$. Let $n_i = \dim_k(V_i)$. By problem VII.6.10(5), $V_i \simeq W_{n_i}$ as a representation of $\mathfrak{sl}_2(k)$, so $u(e)$ is a Jordan block of length $n_i - 1$ (ie maximal length) on V_i. In particular, $W = V_1 \oplus \cdots \oplus V_r$ is the decomposition of W given by the Jordan normal form of N, so it is determined by N. The representations of $\mathfrak{sl}_2(k)$ on the V_i are also uniquely determined, because $\mathfrak{sl}_2(k)$ has a unique irreducible representation of each dimension. This gives uniqueness.

Let's prove the existence of u. Let $W = V_1 \oplus \cdots \oplus V_r$ be the unique decomposition of W such that N stabilizes all the V_i and acts on each V_i by a Jordan block. Fix i, let $n_i = \dim V_i - 1$. Then we can find a basis $(x_0, \ldots, x_{n_i})$ of V_i such that $Nv_j = jv_{j+1}$ for $0 \leq j \leq n_i - 1$, and $Nv_{n_i} = 0$. Sending x_j to $v_1^{n_i - j} v_2^j$ gives an isomorphism of vector spaces $V_i \xrightarrow{\sim} W_{n_i+1}$ that sends N to the endormorphism of W_{n_i+1} induced by e. So we get an ismorphism of vector spaces between W and the $\bigoplus_{i=1}^r W_{n_i+1}$ that sends N to the endomorphism induces by e.

□

VII.6.12 Clebsch-Gordan decomposition

We use the notation of problem VII.6.9, and assume moreover that $k = \mathbb{C}$.

For every $n \in \mathbb{Z}_{\geq 0}$, we have defined an irreducible representation W_{n+1} of $\mathfrak{sl}_2(\mathbb{C})$. The character of such a representation is defined in proposition VI.14.4.4 of chapter VI.

(1). Calculate the character $\chi_{W_{n+1}}$ of W_{n+1}.

(2). For $n, m \in \mathbb{Z}_{\geq 0}$ write $W_{n+1} \otimes W_{m+1}$ as a direct sum of irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$. *(Hint : A finite-dimensional representation of $\mathfrak{sl}_2(\mathbb{C})$ is uniquely determined by its character. (Why ?))*

Solution.

(1). Denote by $\mathfrak{t}$ the space of diagonal matrices in $\mathfrak{s}_2(\mathbb{C})$, and let $e \in \mathfrak{t}^*$ be the map sending $\begin{pmatrix} X & 0 \\ 0 & -X \end{pmatrix}$ to X. (Then the roots of $\mathfrak{sl}_2(\mathbb{C})$ are $2e$ and $-2e$.) For every $\lambda \in \mathfrak{t}^*$, denote by c_λ the corresponding basis element in $\mathbb{Z}[\mathfrak{t}^*]$. By (2) of problem VII.6.9, the character of W_{n+1} is $\sum_{i=0}^n c_{(2i-n)e} = c_{-ne} + c_{(-n+2)e} + \cdots + c_{(n-2)e} + c_{ne}$.

(2). By theorem VI.10.3 of chapter VI, finite-dimensional representations of $\mathfrak{sl}_2(\mathbb{C})$ are uniquely determined by their character. The easiest way to figure out the decomposition into irreducibles of $W_{n+1} \otimes W_{m+1}$ is to calculate characters in a few examples and then to

try to extrapolate.

The correct formula is : for any $n, m \in \mathbb{N}$,

$$W_{n+1} \otimes W_{m+1} = W_{n+m+1} \oplus W_{n+m-1} \oplus W_{n+m-3} \oplus \cdots \oplus W_{|n-m|+1}.$$

This is easily checked on characters.

□

VII.6.13 Dual representation

Show that, if V is an irreducible representation of $\mathrm{SU}(2)$ on a finite-dimensional $\mathbb{C}$-vector space, then $V \simeq V^*$ as representations of $\mathrm{SU}(2)$.

Is the same true for irreducible representations of $\mathrm{U}(1)$? What about $\mathrm{SU}(3)$?

Solution. By theorem VI.11.5 of chapter VI (and the remark following it), we know that the complex irreducible finite-dimensional representations of $\mathrm{SU}(2)$ are the $W_{n+1} = \mathrm{Sym}^n \mathbb{C}^2$, where the representation on $\mathbb{C}^2$ is given by the inclusion $\mathrm{SU}(2) \subset \mathrm{GL}_2(\mathbb{C})$. Let T_c be the diagonal torus of $\mathrm{SU}(2)$. Then, if $t = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \in T_c$, we have

$$\chi_{W_{n+1}}(t) = \lambda^n + \lambda^{n-2} + \cdots + \lambda^{-n+2} + \lambda^{-n}$$

(see problem VII.6.12). By proposition II.1.1.11 of chapter II, we have, for every $t \in T_c$, $\chi_{W_{n+1}^*}(t) = \chi_{W_{n+1}}(t^{-1}) = \chi_{W_{n+1}}(t)$. As the character of a finite-dimensional representation determines the representation up to isomorphism (see theorem VI.10.3 of chapter VI), we get $W_{n+1}^* \simeq W_{n+1}$.

The analogous statement is false for $\mathrm{U}(1)$ and $\mathrm{SU}(3)$. For example, let V be the 1-dimensional representation of $U(1)$ given by the inclusion $\mathrm{U}(1) \subset \mathbb{C}^\times$. Then, for every $\lambda \in \mathrm{U}(1)$, $\chi_V(\lambda) = \lambda$ and $\chi_{V^*}(\lambda) = \lambda^{-1}$; so $\chi_V \neq \chi_{V^*}$. Similarly, let W be the 3-dimensional representation of $\mathrm{SU}(3)$ given by the inclusion $\mathrm{SU}(3) \subset \mathrm{GL}_3(\mathbb{C})$. Then, for $t = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \in \mathrm{SU}(3)$, $\chi_W(t) = \lambda_1 + \lambda_2 + \lambda_3$ and $\chi_{W^*}(t) = \lambda_1^{-1} + \lambda_2^{-1} + \lambda_3^{-1}$. For example, if we take $\lambda_1 = \lambda_2 = i$ and $\lambda_3 = -1$, then $\chi_W(t) = 2i - 1 \neq \chi_{W^*}(t) = -2i - 1$.

□

VII.6.14 Some representations of $\mathfrak{sl}_n(k)$

We want to generalize some of the results of problem VII.6.9.

By definition, the *standard representation* of $\mathfrak{g} := \mathfrak{sl}_n(k)$ is the inclusion $\mathfrak{g} \subset \mathfrak{gl}_n(k)$. It's a representation of $\mathfrak{g}$ on $V := k^n$. For every $d \geq 0$, we write $W_{d+1} = S^d V$ and consider the symmetric power representation $\mathfrak{g}$ on this space.

If k is a field of characteristic 0, show that all these representations are irreducible. What are their highest weights ?

Solution. Let $(e_1, \dots, e_n)$ be the standard basis of k^n. By problem VII.6.7(3), we know that the $e_1^{d_1} \dots e_n^{d_n}$, for $d_1, \dots, d_n \in \mathbb{N}$ such that $d_1 + \cdots + d_n = d$, form a basis of $S^d V$.

If $(d_1, \dots, d_n), (d'_1, \dots, d'_n) \in \mathbb{N}^n$, we write $(d_1, \dots, d_n) \preceq (d'_1, \dots, d'_n)$ if $d_1 + \cdots + d_i \leq d'_1 + \cdots + d'_i$ for $1 \leq i \leq n-1$ and $d_1 + \cdots + d_n = d'_1 + \cdots + d'_n$. (This is just the Bruhat order.)

For $1 \leq i \leq n-1$, we denote by X_i the matrix in $M_n(k)$ defined by

$$X_i e_j = \begin{cases} e_i & \text{if } j = i+1 \\ 0 & \text{otherwise} \end{cases}$$

(that is, X_i is the elementary matrix often denoted by $E_{i,i+1}$). Then $X_i \in \mathfrak{sl}_n(k)$ (because $\mathrm{Tr}(X_i) = 0$). If $d_1, \dots, d_n \in \mathbb{N}$, then

$$X_i(e_1^{d_1} \dots e_n^{d_n}) = \begin{cases} 0 & \text{if } d_{i+1} = 0 \\ d_{i+1} e_1^{d_1} \dots e_{i-1}^{d_{i-1}} e_i^{1+d_i} e_{i+1}^{-1+d_{i+1}} e_{i+2}^{d_{i+2}} \dots e_n^{d_n} & \text{otherwise.} \end{cases}$$

Let $d_1, \dots, d_n \in \mathbb{N}$. Let

$$X = X_{n-1}^{d_1+\cdots+d_{n-1}} \dots X_2^{d_1+d_2} X_1^{d_1},$$

where we take the produit in the universal enveloping algebra U of $\mathfrak{sl}_n(\mathbb{C})$ (which also acts on W_{d+1}). Then, for all $d'_1, \dots, d'_n \in \mathbb{N}$ such that $d = d'_1 + \cdots + d'_n$, we have

$$X(e_1^{d'_1} \dots e_n^{d'_n}) = \begin{cases} (d_1)!(d_1+d_2)! \dots (d_1 + \cdots + d_{n-1})! e_n^d & \text{if } d'_i = d_i \text{ for every } i \\ 0 & \text{if } (d_1, \dots, d_n) \not\preceq (d'_1, \dots, d'_n). \end{cases}$$

Now let's prove that W_{d+1} is irreducible. Let V be a nonzero subrepresentation of W_{d+1}, choose $v \in V - \{0\}$, write

$$v = \sum_{(d_1,\dots,d_n)} a_{d_1,\dots,d_n} e_1^{d_1} \dots e_n^{d_n},$$

where the sum is on the $(d_1, \dots, d_n) \in \mathbb{N}^n$ such that $d_1 + \cdots + d_n = d$, and choose $(d_1, \dots, d_n)$ maximal for the order $\preceq$ such that $a_{d_1,\dots,d_n} \neq 0$. By the calculation above, if $X = X_{n-1}^{d_1+\cdots+d_{n-1}} \dots X_2^{d_1+d_2} X_1^{d_1}$, then Xv is a nonzero multiple of e_1^d, so $e_1^d \in V$.

Now let $Y_i = {}^t X_i$, for $1 \leq i \leq n-1$. We have

$$Y_i e_j = \begin{cases} e_{i+1} & \text{if } j = i \\ 0 & \text{otherwise.} \end{cases}$$

Let $d_1, \dots, d_n \in \mathbb{N}$ be such that $d_1 + \cdots + d_n = d$, and let $Y = Y_{n-1}^{d_n} Y_{n-2}^{d_{n-1}+d_n} \dots Y_1^{d_2+\cdots+d_n}$. Then $Y e_1^d$ is a nonzero multiple of $e_1^{d_1} \dots e_n^{d_n}$, so $e_1^{d_1} \dots e_n^{d_n} \in V$. This shows that $V = W_{d+1}$.

Finally, let's find the weights of W_{d+1}. It is clear that each $e_1^{d_1} \dots e_n^{d_n}$ is an eigenvector for every the diagonal element $\mathrm{diag}(\lambda_1, \dots, \lambda_n)$ of $\mathfrak{sl}_n(\mathbb{C})$, with eigenvalue $d_1\lambda_1 + \cdots + d_n\lambda_n$. Let $\mathfrak{t}$ be the subspace of diagonal matrices in $\mathfrak{sl}_n(\mathbb{C})$. We have just seen that the weights of W_{d+1} are the characters of $\mathfrak{t}$ of the form $\mathrm{diag}(\lambda_1, \dots, \lambda_n) \longmapsto d_1\lambda_1 + \cdots + d_n\lambda_n$, with $d_1, \cdots + d_n \in \mathbb{N}$ and $d_1 + \cdots + d_n = d$, and that they all have multiplicity 1. The maximal weight for the Bruhat is the one corresponding to $(d_1, \dots, d_n) = (d, 0, \dots, 0)$, i.e., $\mathrm{diag}(\lambda_1, \dots, \lambda_n) \longmapsto d\lambda_1$, and the corresponding highest weight vector is e_1^d.

Note that, just as in the proof of proposition VI.9.2.2 of chapter VI, out strategy to show the irreducibility of W_{d+1} was to show that every nonzero subrepresentation contains a highest weight vector, and then that W_{d+1} is generated (as a $U(\mathfrak{sl}_n(\mathbb{C}))$-module) by a highest weight vector.

□

VII.6.15 A generating family for the universal enveloping algebra

Let $\mathfrak{g}$ be a k-Lie algebra (where k is a commutative ring), and let $(\iota, U(\mathfrak{g}))$ be its universal enveloping algebra. We assume that $\mathfrak{g}$ is finitely generated as a k-module[24] and choose a generating family $(X_1, \dots, X_r)$ of $\mathfrak{g}$.

Show that $U(\mathfrak{g})$ is generated as a k-module by the $\iota(X_1)^{n_1} \dots \iota(X_r)^{n_r}$, for $n_1, \dots, n_r \geq 0$.

Solution. The first thing to do is to show that the subspace generated by the $\iota(X_1)^{n_1} \dots \iota(X_r)^{n_r}$ is equal to the subalgebra generated by $\iota(\mathfrak{g})$.

We write $Y_i = \iota(X_i)$ and $U = U\mathfrak{g}$. For every $n \geq 0$, let U'_n be the k-subspace of U generated by the $Y_1^{n_1} \dots Y_r^{n_r}$ for $n_1, \dots, n_r \in \mathbb{Z}_{\geq 0}$ such that $n_1 + \cdots + n_r \leq n$, and U''_n be the k-subspace of U generated by the $Y_{r_1} \dots Y_{r_m}$, with $r_1, \dots, r_m \in \{1, \dots, r\}$ and $m \leq n$. Obviously, $U'_n \subset U''_n$. Let's show by induction on n that $U'_n = U''_n$. We have $U'_0 = U''_0 = k$ and $U'_1 = U''_1 = k \oplus \iota(\mathfrak{g})$, so take $n \geq 2$ and suppose that we know that $U'_{n-1} = U''_{n-1}$. Let $m \geq n$ and $r_1, \dots, r_m \in \{1, \dots, r\}$. We want to show that $Y_{r_1} \dots Y_{r_m} \in U'_n$. This follows from the induction hypothesis if $m < n$, so we may assume $n = m$. For every $i \in \{1, \dots, n-1\}$, $Y_{r_i} Y_{r_{i+1}} = [Y_{r_i}, Y_{r_{i+1}}] + Y_{r_{i+1}} Y_{r_i}$, so

$$Y_{i_1} \dots Y_{i_n} = Y_{i_1} \dots Y_{r_{i-1}} Y_{r_{i+1}} Y_{r_i} Y_{r_{i+2}} \dots Y_{r_n} \mod U''_{n-1} = U'_{n-1}.$$

As $\mathfrak{S}_n$ is generated by the transpositions $(i, i+1)$, $1 \leq i \leq n-1$, we see that, for every $\sigma \in \mathfrak{S}_n$,

$$Y_{i_1} \dots Y_{i_n} = Y_{i_{\sigma(1)}} \dots Y_{i_{\sigma(n)}} \mod U'_{n-1}.$$

[24] This is not really necessary, we just don't want to worry about the best way to order infinite sets.

But we can choose σ such that $i_{\sigma(1)} \leq \cdots \leq i_{\sigma(n)}$, and then $Y_{i_{\sigma(1)}} \dots Y_{i_{\sigma(n)}}$ will be in U'_n. So $Y_{i_1} \dots Y_{r_n} \in U'_n$.

Let $U' = \sum_{n \geq 0} U'_n$. We have seen that $U' = \sum_{n \geq 0} U''_n$, that is, that U' is the subalgebra of U generated by $\iota(\mathfrak{g})$. We want to show that $U' = U$. We use the universal property of the universal enveloping algebra. As ι sends $\mathfrak{g}$ to U', there exists a unique k-algebra map $\varphi : U \to U'$ such that $\varphi\iota = \iota$. Let ψ be the endomorphism of U that is the composition of $\varphi : U \to U'$ and of the inclusion $U' \subset U$. Then $\psi \circ \iota = \iota$, so, by the universal property of U, $\psi = \mathrm{id}_U$. This implies that $U' = U$.

□

VII.6.16 Universal enveloping algebra and differential operators (and a proof of the Poincaré-Birkhoff-Witt theorem for $\mathfrak{gl}_n(k)$ if $\mathrm{char}(k) = 0$)

In this problem, we assume that k is a field of characteristic 0, we fix a positive integer n, and we write $G = \mathrm{GL}_n$, seen as a linear algebraic as in problem VII.6.5. The goal of the problem is to give a description of the universal enveloping algebra of $\mathrm{Lie}(G) = \mathfrak{gl}_n(k)$. (We could make this work for any linear algebraic group if we assumed k algebraically closed, but we'll stick to GL_n for simplicity.)

Let

$$A = k[t_{ij}, 1 \leq i, j \leq n][\tfrac{1}{\det}].$$

[25]

We see A as an algebra of functions $G(k) = \mathrm{GL}_n(k) \to k$ by sending $f = P \det^k \in A$, with $P \in k[t_{ij}, 1 \leq i, j \leq n]$, to the function

$$\widetilde{f} : g = (g_{ij}) \longmapsto P(g_{ij}) \det(g)^k.$$

Because k is infinite, the function $\widetilde{f}$ uniquely determines the rational fraction f (you can assume this), so we will just identify them and write $f(g)$ for $\widetilde{f}$.

Now we will define some k-linear endomorphisms of A :

- For any $i, j \in \{1, \dots, n\}$, we have the endomorphism $\frac{\partial}{\partial t_{ij}}$ that sends f to $\frac{\partial f}{\partial t_{ij}}$.
- If $a \in A$, we write m_a for the endomorphism of A that sends f to af.
- If $g \in G(k)$, we write L_g for the endomorphism of A that sends f to the function $x \longmapsto f(gx)$ on $G(k)$. (It is very easy to check that this function is still in A.)

[25] This A is the k-algebra of regular functions on G, see problem VII.6.17).

Let $D(G)$ be the subalgebra of $\mathrm{End}(A)$ (= the k-algebra of k-linear endomorphisms of A) generated by the $\frac{\partial}{\partial t_{ij}}$ for $i, j \in \{1, \ldots, n\}$ and the m_a for $a \in A$. This is called the algebra of (algebraic) differential operators on G.

(1). Show that, for every element D of $D(G)$, we can write

$$D = \sum_{(k_{11},\ldots,k_{nn}) \in \mathbb{N}^{\{1,\ldots,n\}\times\{1,\ldots,n\}}} a_{k_{11},\ldots,k_{nn}} \frac{\partial^{k_{11}+\cdots+k_{nn}}}{(\partial t_{11})^{k_{11}} \ldots (\partial t_{nn})^{k_{nn}}},$$

where $(a_{k_{11},\ldots,k_{nn}})$ is a uniquely determined family $a_{k_{11},\ldots,k_{nn}}$ of elements of A, indexed by $\mathbb{N}^{\{1,\ldots,n\}\times\{1,\ldots,n\}}$, that has almost all (= all but a finite number) of its elements equal to 0.

(2). For every $d \in \mathbb{N}$, let $D_d(G)$ be the subspace of $D(G)$ of operators of the form

$$D = \sum_{(k_{11},\ldots,k_{nn}) \in \mathbb{N}^{\{1,\ldots,n\}\times\{1,\ldots,n\}} | k_{11}+\ldots k_{nn} \leq d} a_{k_{11},\ldots,k_{nn}} \frac{\partial^{k_{11}+\cdots+k_{nn}}}{(\partial t_{11})^{k_{11}} \ldots (\partial t_{nn})^{k_{nn}}}.$$

We call elements of $D_d(G)$ differential operators of order $\leq d$.

Show that $D_d(G)D_{d'}(G) \subset D_{d+d'}(G)$ for all $d, d' \in \mathbb{N}$. [26]

(3). Let

$$D_{\mathrm{inv}}(G) = \{D \in D(G) | \forall g \in G(k),\ L_g \circ D = D \circ L_g\}.$$

We call elements of $D_{\mathrm{inv}}(G)$ invariant differential operators on G.

Show that $D_{\mathrm{inv}}(G)$ is a subalgebra of $D(G)$ containing the unit element, and that the linear transformation $\varphi : D_{\mathrm{inv}}(G) \to A^* := \mathrm{Hom}_k(A, k)$ sending D to $f \longmapsto D(f)(1)$ is injective.

(4). Let k' be the k-algebra $k \oplus k\varepsilon$, with ε^2 (this is called the k-algebra of dual numbers), and let $\mathfrak{g} = \mathrm{Lie}(G) = \mathfrak{gl}_n(k)$. If $f \in A$, $g \in G(k)$ and $X = (x_{ij}) \in \mathfrak{gl}$, show that

$$f(g + \varepsilon X) = f(g) + \varepsilon \sum_{1 \leq i,j \leq n} x_{ij} \frac{\partial f}{\partial t_{ij}}(g).$$

(5). If $X \in \mathfrak{g}$, $f \in A$ and $g \in G(k)$, set

$$\widetilde{X}(f)(g) = \frac{1}{\varepsilon}(f(g(1 + \varepsilon X)) - f(g)).$$

Show that this makes sense, that $\widetilde{X}(f) \in A$ (hence $\widetilde{X} \in \mathrm{End}(A)$) and that $\widetilde{X} \in D_{\mathrm{inv}}(G)$.

(6). Show that $\widetilde{X}$ is a derivation[27] of A for every $X \in \mathfrak{g}$.

[26] Hence $D(G)$, with the filtration given by the $D_d(G)$, is a filtered k-algebra.
[27] See problem VII.6.4(2).

(7). Show that the map $X \longmapsto \widetilde{X}$ from $\mathfrak{g}$ to $D_{\text{inv}}(G)$ is a Lie algbera map (i.e. that it sends $[X,Y]$ to $\widetilde{X} \circ \widetilde{Y} - \widetilde{Y} \circ \widetilde{X}$ for all $X, Y \in \mathfrak{g}$).

(8). Let $(\iota, U\mathfrak{g})$ be the universal enveloping algebra of $\mathfrak{g}$, and let $\alpha : U\mathfrak{g} \to D_{\text{inv}}(G)$ be the unique k-algebra map such that $\alpha(\iota(X)) = \widetilde{X}$ for every $X \in \mathfrak{g}$. [28] Show that α is injective.

More precisely, let $(E_{ij})_{(i,j)\in\{1,\dots,n\}^2}$ be the canonical basis of $\mathfrak{g}$. We have seen in problem VII.6.15 that $U\mathfrak{g}$ is generated as a vector space by the products $\iota(E_{11})^{d_{11}} \dots \iota(E_{nn})^{d_{nn}}$, with $d_{11}, \dots, d_{nn} \in \mathbb{Z}_{\geq 0}$, where we use the lexicographic order on $\{1, \dots, n\}^2$ in the products above. Show that the images of these elements by α, i.e. the $\widetilde{E}_{11}^{d_{11}} \dots \widetilde{E}_{nn}^{d_{nn}}$, form a linearly independent family in $D(G)$.

Hint : Let's denote by $D_{=d}(G)$ the A-submodule of $D(G)$ (freely) generated by the $\prod_{i,j} \delta_{ij}^{d_{ij}}$ such that $\sum_{i,j} d_{ij} = d$. Show that $D(G) = \bigoplus_{d\geq 0} D_{=d}(G)$, [29] and denote by o_d the projection $D(G) \to D_{=d}(G)$. Then if we have a relation among the $\widetilde{E}_{11}^{d_{11}} \dots \widetilde{E}_{nn}^{d_{nn}}$, apply $\varphi \circ o_d$ to it (for a well-chosen value of d), and evaluate this on appropriate elements of A.

(9). Show that α is surjective.

Hint : Remember the filtration $(U_d\mathfrak{g})$ of $U\mathfrak{g}$ defined in the solution of problem VII.6.15, compare it to the filtration $(D_d(G))$ and look at what happens on the quotients of these filtrations.

Note that question (8) implies that the family of generators $\iota(E_{11})^{d_{11}} \dots \iota(E_{nn})^{d_{nn}}$, with $d_{11}, \dots, d_{nn} \in \mathbb{Z}_{\geq 0}$, of $U\mathfrak{g}$ is actually a basis. This result is actually true for any Lie algebra over a commutative ring k that is free as a k-module, and is called the Poincaré-Birkhoff-Witt theorem. [30]

Solution.

(1). First note that $a \longmapsto m_a$ is an injective k-algebra map from A to $\mathrm{End}(A)$. We use it to identify A to a subalgebra of $\mathrm{End}(A)$. For every i, j, write $\partial_{ij} = \frac{\partial}{\partial t_{ij}} \in \mathrm{End}(A)$. Note that $\partial_{ij}\partial_{i'j'} = \partial_{i'j'}\partial_{ij}$, for every $i, i', j, j' \in \{1, \dots, n\}$.

For every $d \in \mathbb{Z}_{\geq 0}$, let D_d be the A-submodule of $\mathrm{End}(A)$ generated by the $\delta_{11}^{k_{11}} \dots \delta_{nn}^{k_{nn}}$, with $k_{ij} \geq 0$ and $\sum_{i,j} k_{ij} \leq d$, and let D'_d be the A-submodule generated by the $\delta_{i_1,j_1} m_{a_1} \delta_{i_2,j_2} m_{a_2} \dots \delta_{i_e,j_e} m_{a_e}$, for $i_1, j_1, \dots, i_e, j_e \in \{1, \dots, n\}$, $a_1, \dots, a_e \in A$ and $e \leq d$. I claim that $D_d = D'_d$. (This will prove the existence of the $a_{k_{11},\dots,k_{nn}}$ in the question, because clearly $D(G) = \sum_{d\geq 0} D'_d$.)

Let's prove the claim by induction on d. For $d = 0$, $D_d = D'_0 = A$. Choose $d \geq 1$, and suppose the claim known for $d - 1$. Let $e \leq d$ and $i_1, j_1, \dots, i_e, j_e \in \{1, \dots, n\}$,

[28] Note that the existence of the injective map $X \longmapsto \widetilde{X}$ from $\mathfrak{g}$ to $D_{\text{inv}}(G)$ forces ι to be injective.
[29] Warning : this not make $D(G)$ a graded algebra, because $D_{=d}(G)D_{=d'}(G) \not\subset D_{=d+d'}(G)$ in general.
[30] See theorem 4.3 of chapter III of part I of Serre's book [31].

$a_1, \dots, a_e \in A$. We want to show that $D := \delta_{i_1,j_1} m_{a_1} \delta_{i_2,j_2} m_{a_2} \dots \delta_{i_e,j_e} m_{a_e} \in D_d$. This follows from the induction hypothesis if $e < d$, so we may assume $e = d$. By the product rule, $\delta_{i_1,j_1} m_{a_1} = m_{\delta_{i_1,j_1}(a)} + m_a \delta_{i_1,j_1}$, so $D \in AD'_{d-1} + A\delta_{i_1,j_1} D'_{d-1}$. As $D'_{d-1} = D_{d-1}$ by the induction hypothesis, this implies that $D \in D_d$.

Now we prove the uniqueness of the $a_{k_{11},\dots,k_{nn}}$ in the question. This is the same as saying that the family $(\delta_{11}^{k_{11}} \dots \delta_{nn}^{k_{nn}})_{k_{11},\dots,k_{nn} \in \mathbb{Z}_{\geq 0}}$ of $D(G)$ is free over A. So suppose that we have

$$D := \sum_{k_{11},\dots,k_{nn} \in \mathbb{Z}_{\geq 0}} a_{k_{11},\dots,k_{nn}} \delta_{11}^{k_{11}} \dots \delta_{nn}^{k_{nn}} = 0,$$

where $a_{k_{11},\dots,k_{nn}} \in A$ and both the sum has only a finite number of nonzero terms. If the coefficients are not all 0, choose $k_{11}, \dots, k_{nn} \in \mathbb{Z}_{\geq 0}$ such that $a_{k_{11},\dots,k_{nn}} \neq 0$ and $a_{l_{11},\dots,l_{nn}} = 0$ if there exists $(i,j) \in \{1,\dots,n\}^2$ such that $l_{ij} > k_{ij}$. Let $f = \prod_{i,j \in \{1,\dots,n\}} t_{ij}^{k_{ij}} \in A$. Then

$$0 = D(f) = \left(\prod_{i,j} k_{ij}! \right) a_{k_{11},\dots,k_{nn}},$$

hence $a_{k_{11},\dots,k_{nn}} = 0$, a contradiction.

(2). Obviously, $D_d(G)$ is the subspace D_d defined in the answer of (1). We showed in (1) that $D_d = D'_d$. As the inclusion $D'_{d_1} D'_{d_2} \subset D'_{d_1+d_2}$ is obvious from the definition of the D'_d, this gives the result.

(3). The unit element (which is id_A) is obviously in $D_{\mathrm{inv}}(G)$, and it's clear that $D_{\mathrm{inv}}(G)$ is a k-subspace of $D(G)$. Let $D, D' \in D_{\mathrm{inv}}(G)$. For every $g \in G(k)$,

$$L_g \circ (D \circ D') = D \circ L_g \circ D' = (D \circ D' \circ L_g),$$

so $D \circ D' \in D_{\mathrm{inv}}(G)$. This show that $D_{\mathrm{inv}}(G)$ is a subalgebra of $D(G)$.

To show that φ is injective, let's take $D \in \operatorname{Ker} \varphi$ and try to show that $D = 0$. Take $f \in A$ and try to calculate $f' := D(f)$. Let $g \in G(k)$. Then, because D is left invariant,

$$f'(g) = (L_g(f'))(1) = (L_g \circ D)(f)(1) = (D \circ L_g)(f)(1) = D(L_g(f))(1) = \varphi(D)(L_g(f)).$$

As $D \in \operatorname{Ker} \varphi$, this is equal to 0. So $f'(g) = 0$ for every $g \in G(k)$, and hence $f' = 0$.[31] We've showed that $D(f) = 0$ for every $f \in A$, which means that $D = 0$.

(4). By a stroke of luck (the fact that $\mathfrak{g} = M_n(K)$ is stable by left multiplication by $G(k)$), we have $g + \varepsilon X \in G(k')$, so it makes sense to apply f to it. (In general, you should use $g(1 + \varepsilon X)$ instead, which is what we want for (5) anyway.)

Let $f_1, f_2 \in A$, and let $f = f_1 f_2$. First we suppose that f_1 and f_2 satisfy the conclusion and we show that f does. Indeed,

[31] Technically, we've proved that $\widetilde{f'} = 0$, but this implies $f' = 0$ and we are explicitely allowed to use this fact.

$$f(g+\varepsilon X) = f_1(g+\varepsilon X)f_2(g+\varepsilon X) =$$

$$\left(f_1(g) + \varepsilon \sum_{i,j} x_{i,j}\delta_{ij}(f_1)(g)\right)\left(f_2(g) + \varepsilon \sum_{i,j} x_{i,j}\delta_{ij}(f_2)(g)\right)$$

The result then follows from the product rule and the fact that $\varepsilon^2 = 0$. Now assume that f_2 and f satisfy the conclusion and that $f_2 \in A^\times$, and we want to show that f_1 satisfies the conclusion. We have

$$f(g+\varepsilon X) = f(g) + \varepsilon \sum_{i,j} x_{i,j}\delta_{ij}(f)(g) =$$

$$f_1(g+\varepsilon X)f_2(g+\varepsilon X) = f_1(g+\varepsilon X)\left(f_2(g) + \varepsilon \sum_{i,j} x_{i,j}\delta_{ij}(f_2)(g)\right)$$

Using the fact that $f_2(g) \neq 0$ and that $\varepsilon^2 = 0$, we get

$$f_1(g+\varepsilon X) = \left(f(g) + \varepsilon \sum_{i,j} x_{i,j}\delta_{ij}(f)(g)\right) f_2(g)^{-1}\left(1 - \varepsilon \sum_{i,j} x_{ij} f_2(g)^{-1}\delta_{ij}(f_2)(g)\right) =$$

$$f_1(g) + \varepsilon \sum_{i,j} x_{i,j}(f_2(g)^{-1}\delta_{ij}(f)(g) - f_2(g)^{-2}\delta_{ij}(f_2)(g)) =$$

$$f_1(g) + \varepsilon \sum_{i,j} x_{i,j}\delta_{ij}(f)(g).$$

As the conclusion of (4) is clearly true for every degree 1 monomial $t_{ij} \in A$, it's true for every element of $k[t_{ij}] \subset A$ by the first calculation above, and in particular for the powers of $\det$. But then the second calculation show that this conclusion is also true for the functions $f(\det)^{-d}$, for $f \in k[t_{ij}]$ and $d \geq 0$, so we get it for every element of A.

(5). By (4), we have

$$f(g(1+\varepsilon X)) - f(g) = \varepsilon \sum_{ij} (gX)_{ij}\delta_{ij}(f)(g),$$

so the definition of $\widetilde{X}(f)(g)$ makes sense, and we get

$$\widetilde{X}(f)(g) = \sum_{i,j} (gX)_{ij}\delta_{ij}(f)(g).$$

For every $(i,j) \in \{1,\dots,n\}$, let $a_{ij} \in A$ be the function that sends $g \in G(k)$ to the (i,j) entry of gX. (This is clearly a polynomial function of the entries of g, it's even linear.) Then the calculation above shows that $\widetilde{X}$ is equal to the differential operator $\sum_{i,j} m_{a_{ij}}\delta_{ij}$. (In particular, $\widetilde{X}$ sends A to A.)

We still have to show that $\widetilde{X}$ is left invariant. Let $g_1, g_2 \in G(k)$ and $f \in A$. Then

$$\widetilde{X}(L_{g_1}(f))(g_2) = \frac{1}{\varepsilon}(f(g_1 g_2(1+\varepsilon X)) - f(g_1 g_2))$$

and

$$(L_{g_1}\widetilde{X}(f))(g_2) = \widetilde{X}(f)(g_1 g_2) = \frac{1}{\varepsilon}(f(g_1 g_2(1+\varepsilon X) - f(g_1 g_2)).$$

So $\widetilde{X} L_{g_1} = L_{g_1}\widetilde{X}$.

(6). In the proof of (5), we have written $\widetilde{X}$ as a sum $\sum_{ij} m_{a_{ij}}\delta_{ij}$, with the a_{ij} in A. As each δ_{ij} is a derivation of A (that's the product rule), this easily implies that $\widetilde{X}$ is a derivation of A.

(7). Let $X, Y \in \mathfrak{g}$, and let $Z = [X, Y]$. We want to compare $D := \widetilde{X} \circ \widetilde{Y} - \widetilde{Y}\widetilde{X}$ and $\widetilde{Z}$.

Let's first show that D is a derivation. Let $f_1, f_2 \in A$. Then

$$D(f_1 f_2) = \widetilde{X}(\widetilde{Y}(f_1 f_2)) - \widetilde{Y}(\widetilde{X}(f_1 f_2)) = \widetilde{X}(\widetilde{Y}(f_1)f_2 + f_1\widetilde{Y}(f_2)) - \widetilde{Y}(\widetilde{X}(f_1)f_2 + f_1\widetilde{X}(f_2)) =$$

$$\widetilde{X}(\widetilde{Y}(f_1))f_2 + \widetilde{Y}(f_1)\widetilde{X}(f_2) + \widetilde{X}(f_1)\widetilde{Y}(f_2) + f_1\widetilde{X}(\widetilde{Y}(f_2))$$
$$-\widetilde{Y}(\widetilde{X}(f_1))f_2 - \widetilde{X}(f_1)\widetilde{Y}(f_2) - \widetilde{Y}(f_1)\widetilde{X}(f_2) - f_1\widetilde{Y}(\widetilde{X}(f_2)) =$$
$$D(f_1)f_2 + f_1 D(f_2).$$

As both D and $\widetilde{Z}$ are derivations, they are determined by their action on the subspace of linear functions in A. Indeed, as a difference of derivations is a derivation, it suffices to show that, if $\delta \in \mathrm{End}(A)$ is a derivation and $\delta(f) = 0$ for every linear function f, then $\delta = 0$. First, the condition and the Leibniz rule imply that $\delta(f) = 0$ for every $f \in k[t_{ij}]$. To finish the proof, we just have to show that if $f_1 \in A$, $f_2 \in A^\times$ and $\delta(f_2) = \delta(f_1 f_2) = 0$, then $\delta(f_1) = 0$. But this follows from the fact $\delta(f_1) = f_2^{-1}(\delta(f_1 f_2) - f_1\delta(f_2))$. (Note the similarity with the proof of (4).)

So let's show that $\widetilde{Z}$ and D are equal on linear functions. Let $f \in k[t_{ji}]$ be linear (ie homogeneous of degree 1). Then applying the definition in (5) gives

$$\widetilde{X} f(g) = f(gX),$$

hence

$$\widetilde{Y}(\widetilde{X} f)(g) = f(gXY).$$

Similary,

$$\widetilde{X}(\widetilde{Y} f)(g) = f(gYX),$$

so

$$(Df)(g) = f(g(XY - YX)) = \widetilde{Z} f(g).$$

(8). First let's note that, thanks to (1), if we denote by $D_{=d}(G)$ the A-submodule of $D(G)$ (freely) generated by the $\prod_{i,j}\delta_{ij}^{d_{ij}}$ such that $\sum_{i,j} d_{ij} = d$, then $D(G) = \bigoplus_{d\geq 0} D_{=d}(G)$. Let's denote by o_d the projection $D(G) \to D_{=d}(G)$, that is, the operator that returns the order d part of a differential operator.

Let $(E_{ij})_{(i,j)\in\{1,\dots,n\}^2}$ be the canonical basis of $\mathfrak{g}$. Note that

$$\widetilde{E}_{ij} = \sum_{r=1}^{n} m_{t_{jr}}\delta_{ir}.$$

Let's prove that the $\widetilde{E}_{11}^{d_{11}}\dots\widetilde{E}_{nn}^{d_{nn}}$, form a linearly independent family in $D(G)$. This will show that α is injective.

Consider a relation

$$D := \sum_{(d_{11},\dots,d_{nn})\in\mathbb{Z}_{\geq 0}^{n^2}} \alpha_{d_{11},\dots,d_{nn}}\widetilde{E}_{11}^{d_{11}}\dots\widetilde{E}_{nn}^{d_{nn}} = 0,$$

where the $\alpha_{d_{11},\dots,d_{nn}}$ are in k and almost all of them are 0. Then we must have $o_d(D) = 0$ for every $d \in \mathbb{Z}_{\geq 0}$. Let d be the biggest integer such that there exist a n^2-uple $(d_{11},\dots,d_{nn})$ with $d = \sum d_{ij}$ and $\alpha_{d_{11},\dots,d_{nn}} \neq 0$, and let's calculate $o_d(D)$. This will be simpler than D for two reasons. First, only the products $\prod_{i,j}\widetilde{E}_{ij}^{d_{ij}}$ with $\sum_{i,j} d_{ij} = d$ will contribute. Second, thanks to the equality $\delta_{ij}m_a = m_{\delta_{ij}a} + m_a\delta_{ij}$ for every $a \in A$, when we calculate the image by o_d of the products $\prod_{i,j}\widetilde{E}_{ij}^{d_{ij}}$, we can pretend that the δ_{ij} commute with all the operators of the form m_a. Now we want to calculate $o_d(D)(f)(1)$ for $f \in A$. The evaluation at 1 introduces a third simplification : in the calculation of $o_d(\prod_{i,j}\widetilde{E}_{ij}^{d_{ij}})(f)(1)$, all the $m_{t_{ij}}$ (which we can move to the left by the second simplification) will go to 0 unless $i = j$, and to 1 if $i = j$. Taking into account the formula for $\widetilde{E}_{ij}$ above, we finally get

$$o_d(\prod_{i,j}\widetilde{E}_{ij}^{d_{ij}})(f)(1) = (\prod_{ij}\delta_{ij}^{d_{ij}}f)(1)$$

for every $f \in A$, hence

$$o_d(D)(f)(1) = (\sum_{d_{11}+\dots+d_{nn}=d} \alpha_{d_{11},\dots,d_{nn}}\prod_{ij}\delta_{ij}^{d_{ij}})(f)(1).$$

For every n^2-uple $(d_{11},\dots,d_{nn})$ such that $\sum_{i,j} d_{ij} = d$, applying the formula above to $f = \prod_{i,j} t_{ij}^{d_{ij}}$ gives $\alpha_{d_{11},\dots,d_{nn}} = 0$. But this contradicts the choice of d.

(9). Remember the filtration U_d of $U\mathfrak{g}$ defined in the solution of problem VII.6.15. (we use the basis (E_{ij}) of $\mathfrak{g}$ to defined it). The map $X \longmapsto \widetilde{X}$ clearly sends U_d to $D_d(G)$ (because $\widetilde{E}_{ij} \in D_1(G)$), and in (8) we proved that this induces an injection $\alpha_d : U_d/U_{d-1} \to D_d(G)/D_{d-1}(G)$ for every $d \geq 0$, with $U_{-1} = 0$,

$D_{-1}(G) = 0$. (Because the map $U_d \to U_d/U_{d-1} \to D_d(G)/D_{d-1}(G)$ is the same as $U_d \to D_d(G) \stackrel{o_d}{\to} D_{=d}(G) \simeq D_d(G)/D_{d-1}(G)$.) Also by (8), the generating family of $U\mathfrak{g}$ defined in problem VII.6.15 is a basis (because its image in $D(G)$ is free). So we get a basis of U_d/U_{d-1} given by the images of the $\iota(E_{11})^{d_{11}} \dots \iota(E_{nn})^{d_{nn}}$ for $d_{11} + \dots + d_{nn}$, and we see that

$$\dim_k(U_d/U_{d-1}) = N_d := |\{(d_{11}, \dots, d_{nn}) \in \mathbb{Z}_{\geq 0}^{n^2} | d_{11} + \dots d_{nn} = d\}|.$$

For every $d \geq 0$, let $P_d \subset k[t_{ij}]$ be the subspace of polynomials of degree $\leq d$ and $\varphi_d : D_d(G) \to \mathrm{Hom}(P_d, k)$ be the map $D \longmapsto (f \longmapsto D(f)(1))$. I claim that φ_d is injective on $D_d(G) \cap D_{\mathrm{inv}}(G)$. Indeed, let $D \in D_d(G) \cap D_{\mathrm{inv}}(G)$. Reasoning as in (3) and using that $L_g(P_d) = P_d$ for every $g \in G(k)$, we see that $D(f) = 0$ for every $f \in P_d$. Write $D = \sum_{(d_{11},\dots,d_{nn})} a_{d_{11},\dots,d_{nn}} \prod_{i,j} \delta_{ij}^{d_{ij}}$, where $a_{d_{11},\dots,d_{nn}} = 0$ for $\sum_{i,j} d_{ij} > d$. Suppose $D \neq 0$, and choose $(d_{11}, \dots, d_{nn})$ such that $a_{d_{11},\dots,d_{nn}} \neq 0$ and $\sum_{i,j} d_{ij}$ is maximal for this property. Applying D to $\prod_{i,j} t_{ij}^{d_{ij}} \in P_d$, we get $a_{d_{11},\dots,d_{nn}} = 0$, a contradiction. So we've proved the claim, and in particular we get that $\dim_k(D_d(G) \cap D_{\mathrm{inv}}(G)) \leq \dim_k(P_d) = \dim_k(U_d)$.

To finish the proof that $\alpha : U\mathfrak{g} \to D_{\mathrm{inv}}(G)$ is an isomorphism, it suffices to show that $\alpha_d : U_d/U_{d-1} \to (D_d(G) \cap D_{\mathrm{inv}}(G))/(D_{d-1}(G) \cap D_{\mathrm{inv}}(G))$ is an isomorphism for every $d \geq 1$. We prove this by induction on d. For $d = 0$, the result is obvious as the source and target of α_0 are both k. Suppose that $d \geq 1$, and the result is known for $d-1$. In particular, we get $\dim_k(U_{d-1}) = \dim_k(D_{d-1}(G) \cap D_{\mathrm{inv}}(G))$, so

$$\dim_k((D_d(G) \cap D_{\mathrm{inv}}(G))/(D_{d-1}(G) \cap D_{\mathrm{inv}}(G))) \leq \dim_k(P_d) - \dim_k(U_{d-1}) = N_d.$$

As the source of α_d is of dimension N_d and α_d is injective, this shows that α_d is bijective.

□

VII.6.17 Regular functions on an algebraic group

We use the notation of problem VII.6.5.

Let $G \subset \mathrm{GL}_n$ be a linear algebraic group over k, given as the set of zeroes of a family $(P_\alpha)_{\alpha \in I}$ of polynomials. We set $\det = \det((X_{ij})_{1 \leq i,j \leq n}) \in k[X_{ij}, 1 \leq i, j \leq n]$, and $R = k[X_{ij}, 1 \leq i, j \leq n][\frac{1}{\det}]$.

The *ring of regular functions on* G is by definition the ring $R_G = R/(P_\alpha, \alpha \in I)$. It is a k-algebra.

On the other hand, the set R'_G of regular functions from G to k is defined as follows : an element of R'_G is the data, for every k-algebra A, of a map (of sets) $f_A : G(A) \to A$ such that, for every map of k-algebras $u : A \to B$, if we denote by $G(u) : G(A) \to G(B)$ the obvious

map (given by applying u to the coordinates), then $f_B \circ G(u) = u \circ f_A$, or in other words, the following diagram commutes :

$$\begin{array}{ccc} G(A) & \xrightarrow{f_A} & A \\ {\scriptstyle G(u)}\downarrow & & \downarrow{\scriptstyle u} \\ G(B) & \xrightarrow[f_B]{} & B \end{array}$$

(Note the similarity with the definition of morphisms of algebraic groups in the next problem.)

The set R'_G has an obvious structure of k-algebra : If (f_A), (f'_A) are two elements of R'_G and $\lambda \in k$, we set $(f_A) + (f'_A) = (f_A + f'_A)$, $(f_A)(f'_A) = (f_A f'_A)$ and $\lambda(f_A) = (\lambda f_A)$.

(1). If $P \in k[X_{ij}, 1 \leq i, j \leq n]$, it defines for every k-algebra A a map $G(A) \to A$ by sending $g = (g_{ij})$ to $P(g_{ij})$. Use this to get a map of k-algebras $\varphi : R_G \to R'_G$.

(2). Show that φ is an isomorphism. *(Hint : If $(f_A) \in R'_G$, find a preimage of (f_A) by applying f_{R_G} to a well-chosen element of $G(R_G)$.)*

Solution.

(1). If $P \in k[X_{ij}]$ and A is a k-algebra, let $\varphi(P)_A : G(A) \to A$ be the map $(g_{ij}) \longmapsto P(g_{ij})$. Then $P \longmapsto \varphi(P)_A$ is a morphism of k-algebras $k[X_{ij}] \to Maps(G(A), A)$ (where the algebra structure on the right hand side is given by pointwise addition and multiplication), and $\varphi(det)_A$ send $G(A)$ to $A^\times$, so $P \longmapsto \varphi(P)_A$ extends to a k-algebra map $R \to Maps(G(A), A)$. This map sends all the P_α to 0, so it defines a k-algebra map $R_G \to Maps(G(A), A)$, that we'll still call $f \longmapsto \varphi(f)_A$. It's now clear that the family $(\varphi(f)_A)$ (as A varies) defines an element of R'_G for every $P \in R_G$, and that this is a map of k-algebras $\varphi : R_G \to R'_G$.

(2). Let's try to construct an inverse $\psi : R'_G \to R_G$ of φ. Let g_0 be the element of $(X_{ij})_{1 \leq i,j \leq n}$ of $M_n(k[X_{ij}])$. Then g_0 is not in $\mathrm{GL}_n(k[X_{ji}])$ because its determinant $\det$ is not invertible, but g_0 is in $\mathrm{GL}_n(R)$, because in R, $\det$ is invertible (by construction of R). Of course, g has no reason to be its $G(R)$, but its image g in $\mathrm{GL}_n(R_G)$ is in $G(R_G)$, because $P_\alpha(g) = P_\alpha(X_{ij}) = 0$ in R_G.

Now if $f = (f_A)$ be an element of R'_G, we set $\psi(f) = f_{R_G}(g) \in R_G$. This is obviously a morphism of k-algebras. Let's show that $\psi \circ \varphi = \mathrm{id}_{R_G}$. This is almost tautological. Let $P \in R_G$, then $\psi(\varphi(P)) = P(X_{ij})$, that is, $\psi(\varphi(P)) = P$.

Let's show that $\psi \circ \varphi = \mathrm{id}_{R'_G}$. This is a bit less tautological but not very hard. Let $f = (f_A) \in R'_G$, and set $P = \psi(f)$ and $f' = \varphi(P)$. Let A be a k-algebra, and let $h = (h_{ij}) \in G(A)$. Define a k-algebra map $k[X_{ij}] \to A$ by sending X_{ij} to h_{ij}. Because $\det(h) \in A^\times$, this map extends to a k-algebra map $R \to A$. Because $P_\alpha(h) = 0$ for every $\alpha \in I$, it further goes to quotient and defines a k-algebra map $u : R_G \to A$. Now applying the compatibility property of f, we see that $u \circ f_{R_G} = f_A \circ G(u)$. Applying this to the element $g \in G(R_G)$ defined above and noting that $G(u)(g) = h$ (by the very definition of

u) gives

$$f_A(h) = u(f_{R_G}(g)) = u(P) = P(h_{ij}) = f'_A(h).$$

□

VII.6.18 Differentiating morphisms of algebraic groups

We use the notation of problem VII.6.5.

Let k be a commutative ring, and let $G \subset \mathrm{GL}_n$ and $H \subset \mathrm{GL}_m$ be two linear algebraic groups over k. A *morphism of algebraic groups* $\rho : G \to H$ is the data, for every k-algebra A, of a morphism of groups $\rho_A : G(A) \to H(A)$ such that, for every map of k-algebras $u : A \to B$, if we denote by $G(u) : G(A) \to G(B)$ and $H(u) : H(A) \to H(B)$ the obvious maps (given by applying u to the coordinates), then $\rho_B \circ G(u) = H(u) \circ \rho_A$, or in other words, the following diagram commutes :

$$\begin{array}{ccc} G(A) & \xrightarrow{\rho_A} & H(A) \\ {\scriptstyle G(u)}\downarrow & & \downarrow{\scriptstyle H(u)} \\ G(B) & \xrightarrow[\rho_B]{} & H(B) \end{array}$$

(1). If $\rho : G \to H$ is a morphism of algebraic groups, we define $d\rho : \mathrm{Lie}\, G \to \mathrm{Lie}\, H$ by

$$d\rho(X) = \frac{1}{\varepsilon}(\rho(1 + \varepsilon X) - \rho(1)).$$

Show that this is well-defined and a morphism of Lie algebras over k.

(2). If $k = \mathbb{C}$ and $\rho : G \to H$ is a morphism of algebraic groups over k, show that $\rho_{\mathbb{C}} : G(\mathbb{C}) \to H(\mathbb{C})$ is a continuous morphism of groups, and that the map $d\rho : \mathrm{Lie}\, G \to \mathrm{Lie}\, H$ defined in question (1) and the map $d\rho_{\mathbb{C}} : \mathrm{Lie}(G(\mathbb{C})) \to \mathrm{Lie}(H(\mathbb{C}))$ defined in theorem VI.5.2 of chapter VI agree on $\mathrm{Lie}(G(\mathbb{C}))$. (Note that the question makes sense by (4) of problem VII.6.5.)

(3). If $\rho : \mathrm{SL}_n \to H$ is a morphism of algebraic groups over a field k of characteristic 0, show that $d\rho$ uniquely determines ρ. [32]

Solution.

(1). Remember that the Lie algebra of G is by definition the set of $g - I_n$, where g is in the kernel of the map $G(k[\varepsilon]) \to G(k)$ coming from the k-algebra map $k[\varepsilon] \to k$, $\varepsilon \longmapsto 0$. By

[32] We do not really need the first group to be SL_n, but we do need it to be connected, and I haven't defined what this means for algebraic groups.

definition of a morphism of algebraic groups, we have a commutative diagram

$$\begin{array}{ccc} G(k[\varepsilon]) & \xrightarrow{\rho_{k[\varepsilon]}} & H(k[\varepsilon]) \\ \downarrow & & \downarrow \\ G(k) & \xrightarrow[\rho_k]{} & H(k) \end{array}$$

so $\rho_{k[\varepsilon]}$ induces a map $\mathrm{Ker}(G(k[\varepsilon]) \to G(k)) \to \mathrm{Ker}(H(k[\varepsilon]) \to H(k))$, i.e., a map $d\rho : \mathrm{Lie}(G) \to \mathrm{Lie}(H)$. Let's see that this map is given by the formula of the problem. Let $X \in \mathrm{Lie}(G)$. Then, by the definition of $d\rho$ we just gave,

$$1 + \varepsilon d\rho(X) = \rho(1 + \varepsilon X),$$

which is exactly what we wanted to prove.

Now let's show that $d\rho$ is a morphism of k-Lie algebras. Let $X_1, X_2 \in \mathrm{Lie}(G)$ and $\lambda \in k$. We'll write ρ instead of $\rho_{k[\varepsilon]}$ in what follows. First, as $\rho : G(k[\varepsilon]) \to H(k[\varepsilon])$ is a morphism of groups, we have

$$1 + \varepsilon d\rho(X_1 + X_2) = \rho(1 + \varepsilon(X_1 + X_2)) = \rho((1 + \varepsilon X_1)(1 + \varepsilon X_2))$$

$$= (1 + \varepsilon d\rho(X_1))(1 + \varepsilon d\rho(X_2)) = 1 + \varepsilon(d\rho(X_1) + d\rho(X_2)),$$

i.e., $d\rho(X_1 + X_2) = d\rho(X_1) + d\rho(X_2)$.

Let $u : k[\varepsilon] \to k[\varepsilon]$ be the k-algebra map sending ε to $\lambda\varepsilon$. Then we have a commutative diagram

$$\begin{array}{ccc} G(k[\varepsilon]) & \xrightarrow{\rho} & H(k[\varepsilon]) \\ {\scriptstyle G(u)}\downarrow & & \downarrow{\scriptstyle H(u)} \\ G(k[\varepsilon]) & \xrightarrow[\rho]{} & H(k[\varepsilon]) \end{array}$$

so

$$1 + \varepsilon d\rho(\lambda X_1) = \rho(1 + k\varepsilon X_1) = \rho \circ G(u)(1 + \varepsilon X_1) = H(u) \circ \rho(1 + \varepsilon X_1)$$

$$= H(u)(1 + \varepsilon d\rho(X_1)) = 1 + \varepsilon\lambda d\rho(X_1),$$

i.e., $d\rho(\lambda X_1) = \lambda d\rho(X_1)$.

Finally, we consider $k' = k[\varepsilon] \otimes_k k[\varepsilon] = k[\varepsilon_1, \varepsilon_2]$, where $\varepsilon_1 = \varepsilon \otimes 1$ and $\varepsilon_2 = 1 \otimes \varepsilon$. We have maps $u_1, u_2, u : k[\varepsilon] \to k'$ defined by $u_1(\varepsilon) = \varepsilon_1$, $u_2(\varepsilon) = \varepsilon_2$, and $u(\varepsilon) = \varepsilon_1\varepsilon_2 = \varepsilon \otimes \varepsilon$. Let $g_i = G(u_i)(1 + \varepsilon X_i)$ and $h_i = H(u_i)(1 + \varepsilon d\rho(X_i))$, $i = 1, 2$. We have seen in the solution of problem VII.6.5 that $g_1 g_2 g_1^{-1} g_2^{-1} = G(u)(1 + \varepsilon[X_1, X_2])$ and $h_1 h_2 h_1^{-1} h_2^{-1} = H(u)(1 + \varepsilon[d\rho(X_1), d\rho(X_2)])$. As $\rho_{k'} \circ G(u) = H(u) \circ \rho$, we get

$$H(u)(1 + \varepsilon d\rho([X_1, X_2])) = H(u)(\rho(1 + \varepsilon[X_1, X_2])) = H(u)(1 + \varepsilon[d\rho(X_1), d\rho(X_2)]),$$

hence, thanks to the injectivity of u, $d\rho([X_1, X_2]) = [d\rho(X_1), d\rho(X_2)]$.

(2). The fact that $\rho_{\mathbb{C}}$ is a morphism of groups follows immediately from the definition of a morphism of algebraic groups.

For every $i, j \in \{1, \dots, m\}$ and every commutative ring A, let $t_{ij,A} : M_m(A) \to A$ be the function giving the (i, j)-th entry. Then the family $(t_{ij,A} \circ \rho_A)$, for A varying over all $\mathbb{C}$-algebras, defines an element of R'_G in the notation of problem VII.6.17, and so, by (2) of that same problem, it comes from an element ρ_{ij} of $\mathbb{C}[X_{rs}, 1 \leq r, s \leq n][\det^{-1}]$. In other words, if $g \in G(\mathbb{C})$, all the entries of $\rho_{\mathbb{C}}(g)$ are given by polynomials (independent on g) in the entries of g and in $\det(g)$. This implies that $\rho_{\mathbb{C}}$ is continuous.

For every commutative ring k, every rational function $f \in k(t_1, \dots, t_p)$ and every $a = (a_1, \dots, a_p) \in k^p$ such that all the $\frac{\partial f}{\partial t_p}(a)$ are defined, we set define a linear form $df(a) : k^p \to k$ by

$$df(a)(x_1, \dots, x_p) = \sum_{s=1}^{p} x_s \frac{\partial f}{\partial t_s}(a_1, \dots, a_p).$$

This is similar to what we did in the solution of (3) of problem VII.6.5, and, just as in this solution, we see that, for every $X \in \operatorname{Lie} G$, we have

$$\rho(I_n + \varepsilon X) = I_m + \varepsilon(d\rho_{ij}(1)(X))_{1 \leq i,j \leq m}.$$

So the map $d\rho : \operatorname{Lie} G \to \operatorname{Lie} H$ is given by $d\rho(X) = (d\rho_{ij}(1)(X))_{1 \leq i,j \leq m}$.

On the other hand, if $X \in \operatorname{Lie}(G(\mathbb{C}))$, then

$$d\rho_{\mathbb{C}}(X) = \frac{d}{dt}\rho_{\mathbb{C}}(e^{tX})|_{t=0} = (\frac{d}{dt}\rho_{ij}(e^{tX})|_{t=0})_{1 \leq i,j \leq m},$$

which gives the same result.

(3). Without loss of generality, we may assume that $H = \mathrm{GL}_m$. So let $\rho_1, \rho_2 : \mathrm{SL}_n \to \mathrm{GL}_m$ be two morphisms of algebraic groups over k such that $d\rho_1 = d\rho_2$. We want to prove that $\rho_1 = \rho_2$.

We have seen in (2) that the entries of ρ_1 and ρ_2 are given by polynomials in the coordinates of $M_n(k)$ and in $\det$. Let k' be the smallest subfield of k containing all the coefficients of these polynomials, then k' is of finite transcendence degree over $\mathbb{Q}$, so there exists an injective $\mathbb{Q}$-algebra map $k' \to \mathbb{C}$. All the objects appearing in the problem (SL_n, GL_m, ρ_1, ρ_2) are defined over k', and the equality of the differentials also hold if we consider ρ_1 and ρ_2 as morphisms of algebraic groups over k'. (By the explicit formula for the differential in (2), for example.) So we may assume that k is a subfield of $\mathbb{C}$ that is of finite transcendence degree over $\mathbb{Q}$.

Let A be a k-algebra, let $g = (a_{ij}) \in \mathrm{SL}_n(A)$. We want to prove that $\rho_{1,A}(g) = \rho_{2,A}(g)$. As g is also in $\mathrm{SL}_n(A')$, where A' is the k-subalgebra of A generated by the a_{ij}, we may assume that $A = A'$. In the polynomial algebra $B = k[t_{ij}, 1 \leq i, j \leq n]$, consider the

element $D = \det((t_{ij})_{1\leq i,j\leq n})$. Let $B' = B[t]/(t^n D - 1)$. The k-algebra map map $B \to A$ sending each t_{ij} to a_{ij} sends D to $\det(g) = 1$, so we may extend it to a map $u : B' \to A$ by sending t to 1. Let $g' = (t^{-1}t_{ij})_{1\leq i,j\leq n} \in M_n(B')$, then $\det(g') = t^{-n}D = 1$, so $g' \in \mathrm{SL}_n(B')$. Also, the image of g' by $\mathrm{SL}_n(u) : \mathrm{SL}_n(B') \to \mathrm{SL}_n(A)$ is g, so it suffices to show that $\rho_{1,B'}(g') = \rho_{2,B'}(g')$. Choose a family $(z_{ij})_{1\leq i,j\leq n}$ of elements of $\mathbb{C}$ that are algebraically independent over k, and define a k-algebra map $v : B' \to \mathbb{C}$ by sending each t_{ij} to z_{ij}, and sending t to a primitive nth root of $\det((z_{ij})_{1\leq i,j\leq n})$. This is injective,[33] so it suffices to show that $\rho_{1,\mathbb{C}}(g'') = \rho_{2,\mathbb{C}}(g'')$, where $g'' = \mathrm{SL}_n(v)(g') \in \mathrm{SL}_n(\mathbb{C})$. But we know that $\rho_{1,\mathbb{C}} = \rho_{2,\mathbb{C}}$ by (2) (which says that $d\rho_{1,\mathbb{C}} = d\rho_{2,\mathbb{C}}$), remark VI.5.4 of chapter VI and problem VII.5.4(2).

□

VII.6.19 Semisimple representations of the Lie algebra $\mathfrak{gl}_n(\mathbb{C})$

If $\mathfrak{g}$ is a Lie algebra over a commutative ring k, the *center* of $\mathfrak{g}$ is by definition $\{X \in \mathfrak{g} | \forall Y \in \mathfrak{g},\ [X, Y] = 0\}$.

(1). If k is a commutative ring, calculate the center $\mathfrak{z}$ of $\mathfrak{gl}_n(k)$. If n is invertible in k, show that $\mathfrak{gl}_n(n) = \mathfrak{z} \times \mathfrak{sl}_n(k)$ as Lie algebras.

From now, we take $k = \mathbb{C}$, and we write $\mathfrak{g} = \mathfrak{gl}_n(\mathbb{C})$.

(2). Give an example of a non semisimple-representation finite-dimensional representation of $\mathfrak{g}$.

(3). If $u : \mathfrak{g} \to \mathfrak{gl}(V)$ is an irreducible finite-dimensional representation of $\mathfrak{g}$, show that $u_{|\mathfrak{sl}_n(\mathbb{C})}$ is still irreducible and that $u(\mathfrak{z})$ is contained in the subalgebra $k \cdot \mathrm{id}_V \subset \mathfrak{gl}(V)$.

(4). Let $\Lambda_{\mathfrak{g}}^+ = \{(a_1, \ldots, a_n) \in \mathbb{C}^n | a_i - a_{i+1} \in \mathbb{Z}_{\geq 0} \text{ for } 1 \leq i \leq n-1\}$. We extend the Bruhat order of definition VI.11.1 of chapter VI to $\mathbb{C}^n$ in the following way : if $(a_1, \ldots, a_n), (b_1, \ldots, b_n) \in \mathbb{C}^n$, we say that $(a_1, \ldots, a_n) \preceq (b_1, \ldots, b_n)$ if and only if $b_1 - a_1 \in \mathbb{Z}_{\geq 0}$, $(b_1 + b_2) - (a_1 + a_2) \in \mathbb{Z}_{\geq 0}$,…,$(b_1 + \cdots + b_{n-1}) - (a_1 + \cdots + b_{n-1}) \in \mathbb{Z}_{\geq 0}$, and $b_1 + \cdots + b_n = a_1 + \cdots + a_n$. [34]

Show that, if $(a_1, \ldots, a_n) \in \Lambda_{\mathfrak{g}}^+$, there exists $b \in \mathbb{C}$ such that $(a_1 + b, \ldots, a_n + b) \in \mathbb{Z}^n$, and that $\overline{\lambda} := (a_1 + b, \ldots, a_n + b)$ is in Λ^+ (see definition VI.11.2 of chapter VI) and independent of the choice of b, and the map $\lambda \longmapsto \overline{\lambda}$ just defined respects the Bruhat order.

(5). Construct a bijection $\lambda \longmapsto W_\lambda$ between $\Lambda_{\mathfrak{g}}^+$ and isomorphism classes of irreducible representations of $\mathfrak{g}$ such that :

[33] Because D is a linear combination of monomials where all the t_{ij} have exponent at most 1, so it cannot be a rth power in $k(t_{ij}, 1 \leq i, j \leq n)$ if $r \geq 2$.

[34] Note that this a lift of the Bruhat order of definition VI.14.4.1 of chapter VI.

(a) For every $\lambda \in \Lambda_{\mathfrak{g}}^+$, the action of $\mathfrak{sl}_n(\mathbb{C})$ on W_λ makes W_λ the irreducible representation of $\mathfrak{sl}_n(\mathbb{C})$ of highest weight $\overline{\lambda}$ as in theorem VI.11.5 of chapter VI.

(b) $W_{(1,\dots,1)} = \mathrm{Tr}$.

(c) For all $\lambda, \mu \in \Lambda_{\mathfrak{g}}^+$, we have

$$W_\lambda \otimes W_\mu = W_{\lambda+\mu} \oplus \bigoplus_{\nu \prec \lambda+\mu} W_\nu^{\oplus c_\nu}.$$

(6). Give the irreducible representations of $\mathfrak{g}$ corresponding to the elements $(1, 0, \dots, 0)$, $(1, 1, 0, \dots, 0)$,..., $(1, \dots, 1)$ of $\Lambda_{\mathfrak{g}}^+$ and to the elements $(m, 0, \dots, 0)$, $m \in \mathbb{Z}_{\geq 0}$.

(7). If $\lambda \in \Lambda_{\mathfrak{g}}^+ \cap \mathbb{Z}^n$, show that W_λ comes (by differentiation, as in theorem VI.5.2 of chapter VI) from a continuous representation of $\mathrm{GL}_d(\mathbb{C})$ on W_λ, and give a formula for the character of this representation on diagonal matrices in $\mathrm{GL}_d(\mathbb{C})$.

Solution.

(1). Let's show that $\mathfrak{z} = kI_n$. It is clear that every multiple of I_n is central. Conversely, let $A = (a_{ij}) \in \mathfrak{z}$. Denote by E_{ij}, $1 \leq i, j \leq n$, the elementary matrices in $M_n(k)$ (E_{ij} has (i, j)-th equal to 1, and all other entries equal to 0). Then, for all i, j,

$$0 = [A, E_{ij}] = \sum_{k=1}^{n} a_{ki} E_{kj} - \sum_{l=1}^{n} a_{jl} E_{il}.$$

This gives $a_{ij} = 0$ if $i \neq j$, and $a_{ii} = a_{jj}$ for all $i, j \in \{1, \dots, n\}$, i.e., $A \in kI_n$.

We have a k-linear map $\varphi : \mathfrak{z} \times \mathfrak{sl}_n(k) \to \mathfrak{gl}_n(k)$, $(X, Y) \longmapsto X + Y$. It is a morphism of Lie algebras because, if $X, X' \in \mathfrak{z}$, then $[X, Y] = [X', Y] = 0$ for every $Y \in \mathfrak{gl}_n(k)$, so, for $Y, Y' \in \mathfrak{sl}_n(k)$, we get

$$[\varphi(X, Y), \varphi(X', Y')] = [X + Y, X' + Y'] = [Y, Y'] = \varphi([X, X'], [Y, Y']).$$

Also note that, if $A = X + Y$ with $Y \in \mathfrak{sl}_n(k)$ and $X = xI_n \in \mathfrak{z}$, then $\mathrm{Tr}(A) = \mathrm{Tr}(X) + \mathrm{Tr}(Y) = \mathrm{Tr}(X) = nx$. So, if n is invertible in k, the morphism $\psi : \mathfrak{gl}_n(k) \to \mathfrak{z} \times \mathfrak{sl}_n(k)$, $A \longmapsto (\frac{1}{n}\mathrm{Tr}(A)I_n, A - \frac{1}{n}\mathrm{Tr}(A)I_n)$, is an inverse of φ.

(2). By (1), we have $\mathfrak{g} \simeq \mathbb{C} \times \mathfrak{sl}_n(\mathbb{C})$. So we can take any non-semisimple representation of $\mathbb{C}$ and compose it with the first projection to get the desired representation of $\mathfrak{g}$. For example, the map $u : a \longmapsto \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}$ is a representation of the Lie algebra $\mathbb{C}$ on $\mathbb{C}^2$, but it is not semisimple, because its only subrepresentations are 0, $\mathbb{C} \times \{0\}$ and $\mathbb{C}^2$.

(3). For every $X \in \mathfrak{z}$, $u(X)$ commutes with all the elements of $u(\mathfrak{g})$, so it is a $\mathfrak{g}$-equivariant map from V to itself. As V is an irreducible representation of $\mathfrak{g}$, i.e. a simple $U(\mathfrak{g})$-module,

its $\mathfrak{g}$-equivariant endomorphisms must form a division algebra by Schur's lemma (theorem I.1.4.1 of chapter I), but this division algebra is finite-dimensional over its subalgebra $\mathbb{C}\mathrm{id}_V$ (because it's a subspace of $\mathrm{End}_{\mathbb{C}}(V)$), hence equal to $\mathbb{C}\mathrm{id}_V$ by problem VII.1.3(1).

As $u(\mathfrak{z}) \subset \mathbb{C}\mathrm{id}_V$, every subspace of V is invariant by $\mathfrak{z}$. But $\mathfrak{g} = \mathfrak{z} \oplus \mathfrak{sl}_n(\mathbb{C})$, so $\mathfrak{g}$ and $\mathfrak{sl}_n(\mathbb{C})$ have the same invariant subspaces in V, i.e., only 0 and V.

(4). Let $\lambda = (a_1, \dots, a_n) \in \Lambda^+_{\mathfrak{g}}$. For $i \in \{1, \dots, n-1\}$, we have $a_i - a_n = (a_i - a_{i+1}) + \cdots + (a_{n-1} - a_n) \in \mathbb{Z}$, so we can take $b = -a_n$. Also, $(a_i - b) - (a_{i+1} - b) = a_i - a_{i+1} \in \mathbb{Z}_{\geq 0}$ for every $i \in \{1, \dots, n-1\}$, so $\overline{\lambda} = (a_1 - a_n, \dots, a_n - a_n) \mod (1, \dots, 1)$ is indeed in Λ^+. It is clear that the map $\lambda \longmapsto \overline{\lambda}$ respects the Bruhat order.

We still need to prove the following fact : If $b, b' \in \mathbb{C}$ are such that $(a_1 + b, \dots, a_n + b), (a_1 + b', \dots, a_n + b') \in \mathbb{Z}^n$, then $(a_1 + b, \dots, a_n + b) - (a_1 + b', \dots, a_n + b') \in (1, \dots, 1)\mathbb{Z}$, i.e., $b - b' \in \mathbb{Z}$. But this is obvious, because $b - b' = (a_1 - b') - (a_1 - b)$.

(5). By (1) and (3), every irreducible representation of $\mathfrak{g}$ is of the form $X \longmapsto a\mathrm{Tr}(X)u(X - \frac{1}{n}\mathrm{Tr}(X))$, where $a \in \mathbb{C}$ is a $\mathbb{C}$-linear map and u is an irreducible representation of $\mathfrak{sl}_n(\mathbb{C})$. Conversely, every representation of this form is clearly irreducible.

Let $\lambda = (a_1, \dots, a_n) \in \Lambda^+_{\mathfrak{g}}$, let $\overline{\lambda} \in \Lambda^+$ be as in (4), and let $W_{\overline{\lambda}}$ be the corresponding irreducible representation of $\mathfrak{sl}_n(\mathbb{C})$. We set $W_\lambda = W_{\overline{\lambda}}$ as $\mathbb{C}$-vector space. We make $\mathfrak{sl}_n(\mathbb{C})$ act on W_λ via its action on $W_{\overline{\lambda}}$, and $\mathfrak{z}$ act on W_λ by $\frac{1}{n}(a_1 + \cdots + a_n)\mathrm{Tr}$.

This construction satisfies conditions (1) and (2). To show that it satisfies (3) and that it gives the desired bijection, by the description of irreducible representations of $\mathfrak{g}$ given above (and remark VI.11.7 of chapter VI), it suffices to show that the map $\varphi : \Lambda^+_{\mathfrak{g}} \to \mathbb{C} \times \Lambda^+$ sending $\lambda = (a_1, \dots, a_n)$ to $(\frac{1}{n}(a_1 + \cdots + a_n), \overline{\lambda})$ is bijective and sends the Bruhat order on $\Lambda^+_{\mathfrak{g}}$ to the order $\leq$ on $\mathbb{C} \times \Lambda^+$ given by : $(a, \lambda) \leq (a', \lambda')$ if and only $a = a'$ and $\lambda \preceq \lambda'$.

The statement about the orders is an immediate consequence of the definitions. Let's show that φ is injective. Let $\lambda = (a_1, \dots, a_n)$, $\lambda' = (a'_1, \dots, a'_n) \in \Lambda^+_{\mathfrak{g}}$ such that $\varphi(\lambda) = \varphi(\lambda')$. Then (by definition of the map $\lambda \longmapsto \overline{\lambda}$), there exists $b \in \mathbb{C}$ such that $a'_i = a_i + b$ for every i. But $a_1 + \cdots + a_n = a'_1 + \cdots + a'_n$, so $b = 0$ and $\lambda = \lambda'$. Let's show that φ is surjective. Choose an element of Λ^+, that we write $\lambda \mod (1, \dots, 1)$, for some $\lambda = (a_1, \dots, a_n) \in \mathbb{Z}^n$, and let $a \in \mathbb{C}$. Let $b = a - \frac{1}{n}(a_1 + \cdots + a_n)$, and $\mu = (a_1 + b, \dots, a_n + b)$. Then $\overline{\mu} = \overline{\lambda}$ and $\frac{1}{n}(b_1 + \cdots + b_n) = b + \frac{1}{n}(a_1 + \cdots + a_n) = a$, so $\varphi(\mu) = (a, \overline{\lambda})$.

(6). For $d \in \{1, \dots, n\}$, we write $\varpi_d = (\underbrace{1, \dots, 1}_{d}, 0, \dots, 0) \in \Lambda^+_{\mathfrak{g}}$. Let's show that $W_{\varpi_d} = \Lambda^d \mathbb{C}^n$, where $\mathfrak{g}$ acts on $\mathbb{C}^n$ in the obvious way (i.e. via the identification

$\mathfrak{g} = M_n(\mathbb{C})$).

We just need to check that the two representations agree on $\mathfrak{sl}_n(\mathbb{C})$ and on $\mathbb{C}I_n$. For $\mathfrak{sl}_n(\mathbb{C})$ and $1 \leq d \leq n-1$, this follows from condition (a) in (5). Let's calculate the action of $\mathbb{C}I_n$ on $\Lambda^d\mathbb{C}^n$ (and identify $\Lambda^n\mathbb{C}^n$). Denote by $(e_1, \ldots, e_n)$ the canonical basis of $\mathbb{C}^n$. Then, by proposition VI.9.1.4 of chapter VI, we have a basis of $\Lambda^d\mathbb{C}^n$ given by the $e_{i_1} \wedge \cdots \wedge e_{i_d}$, with $1 \leq i_1 < \cdots < i_d \leq n$. In particular, $\Lambda^n\mathbb{C}^n$ is 1-dimensional, with basis $e_1 \wedge \cdots \wedge e_n$. Let $A = (a_{ij}) \in \mathfrak{g}$. Then, if $1 \leq i_1 < \cdots < i_d \leq n$,

$$A(e_{i_1} \wedge \cdots \wedge e_{i_d}) = \sum_{r=1}^{d} e_{i_1} \wedge \cdots \wedge e_{i_{r-1}} \wedge (Ae_{i_r}) \wedge e_{i_{r+1}} \wedge \cdots \wedge e_{i_d} = \sum_{i=1}^{n} a_{i_r,i_r} e_{i_1} \wedge \cdots \wedge e_{i_d}.$$

If $d = n$, the action of A mutliplies the unique basis element by $a_{11} + \cdots + a_{nn}$, so the representation of $\mathfrak{g}$ on $\Lambda^n\mathbb{C}^n$ is indeed the one given by Tr. On the other hand, for any d, if $A \in \mathbb{C}I_n$ and $a = a_{11}$, then A acts as $da = \frac{d}{n}\mathrm{Tr}(A)$ on $\Lambda^d\mathbb{C}^n$, which is what we wanted to prove.

Let $\lambda = (m, 0, \ldots, 0)$, with $m \in \mathbb{Z}_{\geq 0}$, and let's show that $W_\lambda = S^m\mathbb{C}^n$, where $\mathfrak{g}$ acts on $\mathbb{C}^n$ in the obvious way as before. By problem VII.6.14, we already know that $S^m\mathbb{C}^n$ is the irreducible representation of $\mathfrak{sl}_n(\mathbb{C})$ with highest weight $\overline{\lambda}$, so we just need to check that $\mathbb{C}I_n$ acts in the correct way, i.e. by $\frac{m}{n}\mathrm{Tr}$. Let $(e_1, \ldots, e_n)$ be the canonical basis of $\mathbb{C}^n$. By problem VII.6.7(3), a basis of $S^m\mathbb{C}^n$ is given by the elements $e_1^{d_1} \ldots e_n^{d_n}$, with $d_1, \ldots, d_n \in \mathbb{Z}_{\geq 0}$ and $d_1 + \ldots d_n = m$. If $A = aI_n$, then

$$A(e_1^{d_1} \ldots e_n^{d_n}) = \sum_{i=1}^{n} ad_i(e_1^{d_1} \ldots e_n^{d_n}) = \tfrac{d_1+\ldots d_n}{n}\mathrm{Tr}(A)(e_1^{d_1} \ldots e_n^{d_n}).$$

This gives the desired conclusion.

(7). We use a strategy similar to the one in the proof of theorem VI.11.5 of chapter VI to construct the desired representation of $\mathrm{GL}_n(\mathbb{C})$. First, a convention : Remember that $\Lambda^n\mathbb{C}^n$ is the 1-dimensional representation of $\mathfrak{g}$ given by Tr. For any $a \in \mathbb{C}$, we denote by $(\Lambda^n\mathbb{C}^n)^{\otimes a}$ the 1-dimensional representation of $\mathfrak{g}$ given by $a\mathrm{Tr}$.

Let $\lambda = (a_1, \ldots, a_n) \in \Lambda_{\mathfrak{g}}^+ \cap \mathbb{Z}^n$. Set $d_n = a_n$ and $d_i = a_i - a_{i+1}$ for $1 \leq i \leq n-1$. Then $\lambda = d_1\varpi_1 + \cdots + d_n\varpi_n$, where the ϖ_i are as in (6). So, by condition (c) in (5), W_λ is an irreducible subrepresentation of $V := \bigotimes_{i=1}^n (\Lambda^d\mathbb{C}^n)^{\otimes a_i}$ (this makes sense because $d_1, \ldots, d_{n-1} \in \mathbb{Z}_{\geq 0}$) (it even has multiplicity 1 in V). Now, note that the action of $\mathfrak{g}$ on V comes from an action of $\mathrm{GL}_n(\mathbb{C})$. Indeed, this is true for the obvious action of $\mathfrak{g}$ on $\mathbb{C}^n$, so it's true for all the $\Lambda^d\mathbb{C}^n$, so it's also true for $(\Lambda^{d_i}\mathbb{C}^\times)^{\otimes d_i}$ if $1 \leq i \leq n-1$, because then $d_i \geq 0$. For the last factor, note that the action of $\mathrm{GL}_n(\mathbb{C})$ on $\Lambda^n\mathbb{C}^n$ is just multiplication by the determinant (by definition of the determinant), so, as long as $d_n \in \mathbb{Z}$ (which is true by the hypothesis that $\lambda \in \mathbb{Z}^n$), we can talk about the representation $\det^{d_n}$ of $\mathrm{GL}_n(\mathbb{C})$, whose corresponding representation of $\mathfrak{g}$ is $d_n\mathrm{Tr}$, i.e., $(\Lambda^n\mathbb{C}^n)^{\otimes d_n}$. So, we got an action of $\mathrm{GL}_n(\mathbb{C})$ on V whose differential is the action of $\mathfrak{g}$. But then, by remark VI.8.3 of chapter

VI, a subspace of V is invariant under the action of $\mathfrak{g}$ if and only if it is invariant it is invariant under the action of $\mathrm{GL}_n(\mathbb{C})$. In particular, the subspace W_λ of V is invariant under the action of $\mathrm{GL}_n(\mathbb{C})$, and it gives the desired representation of $\mathrm{GL}_n(\mathbb{C})$.

Come back for a moment to the case where λ is just an element of $\Lambda^+_{\mathfrak{g}}$. First we want to find the character of W_λ as a representation of $\mathfrak{g}$.

Let $\mathfrak{t} = \mathbb{C}^n$ be the space of diagonal matrices in $\mathfrak{g}$, we identify its dual $\mathfrak{t}^*$ to $\mathbb{C}^n$ by using the dual of the canonical basis, which we call $(e_1^*, \dots, e_n^*)$. First we define the ring where our calculations will take place (see definition VI.14.4.2 of chapter VI) : For every $\lambda \in \mathfrak{t}^*$, let

$$C_\lambda = \{\mu \in \mathfrak{t}^* | \mu \preccurlyeq \lambda\}.$$

We define A to be the set of formal sums $\sum_{\lambda \in \mathfrak{t}^*} a_\lambda c_\lambda$, where $a_\lambda \in \mathbb{Z}$, such that there exists $\lambda_1, \dots, \lambda_r \in \mathfrak{t}^*$ such that, if $\lambda \notin C_{\lambda_1} \cup \dots \cup C_{\lambda_r}$, then $a_\lambda = 0$. This contains the group algebra $\mathbb{Z}[\mathfrak{t}^*]$ (where we denote the basis element of $\mathbb{Z}[\mathfrak{t}^*]$ corresponding to λ by c_λ, as in the case of $\mathfrak{sl}_n(\mathbb{C})$). We define the multiplication on A by

$$(\sum_{\lambda \in \mathfrak{t}^*} a_\lambda c_\lambda)(\sum_{\lambda \in \mathfrak{t}^*} b_\lambda c_\lambda) = \sum_{\lambda \in \mathfrak{t}^*} (\sum_{\mu_1 + \mu_2} a_{\mu_1} b_{\mu_2}) c_\lambda.$$

It is easy to check that the sums defining the coefficients on the right-hand side are finite, so this makes sense (and extends the multiplication on $\mathbb{Z}[\mathfrak{t}^*]$).

Let V be a finite-dimensional representation of $\mathfrak{g}$. For $\lambda \in \mathfrak{t}^*$, we set

$$V(\lambda) = \{v \in V | \forall X \in \mathfrak{t},\ X \cdot v = \lambda(X) v\}.$$

(See definition VI.12.1.1 of chapter VI.) The character of V is

$$\chi_V = \sum_{\lambda \in \mathfrak{t}^*} \dim(V(\lambda)) c_\lambda \in A.$$

We set $\Phi^+ = \{e_i^* - e_j^*, i < j\} \subset \mathfrak{t}^*$ and $\rho = \frac{1}{2} \sum_{\alpha \in \Phi^*} \alpha$. (These all lift the similar-named objects of sections VI.12.2 and VI.13 of chapter VI, by the explicit formulas given there). Also, make $W := \mathfrak{S}_n$ act on $\mathfrak{t}^* = \mathbb{C}^n$ in the usual way. Fix $\lambda \in \Lambda^+_{\mathfrak{g}}$, and let

$$\chi_\lambda = \sum_{\sigma \in W} \mathrm{sgn}(\sigma) c_{\sigma(\lambda+\rho)-\rho} \prod_{\alpha \in \Phi^+} (1 + c_{-\alpha} + c_{-2\alpha} + \dots) \in A.$$

Let's show that this is character of W_λ. Write $\lambda = (\lambda_1, \dots, \lambda_n)$ and $\chi_\lambda = \sum_{\mu \in \mathfrak{t}^*} a_\mu c_\mu$, with $a_\mu \in \mathbb{Z}$. We want to show that $a_\mu = \dim(W_\lambda(\mu))$, for every $\mu \in \mathfrak{t}^*$.

By definition of χ_λ, a_μ is 0 unless μ is of the form $\sigma(\lambda + \rho) - \rho - \sum_{\alpha \in \Phi^+} n_\alpha \alpha$, with $\sigma \in W$ and $n_\alpha \in \mathbb{Z}_{\geq 0}$. In particular, if $\mu = (\mu_1, \dots, \mu_n)$ is such that $a_\mu \neq 0$, then $\mu_1 + \dots + \mu_n = \lambda_1 + \dots + \lambda_n$.

By the construction of W_λ in (5), we know that $\mathbb{C}I_n$ acts on W_λ by multiplication by $\frac{\lambda_1+\cdots+\lambda_n}{n}\mathrm{Tr}$. So, for every $\mu \in \mathfrak{t}^*$, $W_\lambda(\mu)$ is also equal to

$$\{v \in W_\lambda | \forall X \in \mathfrak{t}',\ X \cdot v = \mu(X)v\},$$

where $\mathfrak{t}' = \mathfrak{sl}_n(\mathbb{C}) \cap \mathfrak{t}$, i.e., to the $\mu_{|\mathfrak{t}'}$-weight space for W_λ seen as a representation of $\mathfrak{sl}_n(\mathbb{C})$. The dimension of this weight space is given by the Weyl character formula, i.e., theorem VI.13.2 of chapter VI. Using example VI.14.4.3 of the same chapter, we see that, for every $\mu' \in (\mathfrak{t}')^*$, the coefficient of μ' in the character of the representation of $\mathfrak{sl}_n(\mathbb{C})$ on W_λ is the sum of the a_μ over all the extensions μ of μ' to a character of $\mathfrak{t}^*$. But there is at most one such extension μ such that $a_\mu \neq 0$, because $a_\mu \neq 0$ implies that the sum of the coefficients of μ is equal to $\lambda_1 + \cdots + \lambda_n$ by the observation above. So, for every $\mu \in \mathfrak{t}^*$ such that $a_\mu \neq 0$, we get that a_μ is equal the coefficient of $\mu' = \mu_{|\mathfrak{t}'}$ in the character of the representation of $\mathfrak{sl}_n(\mathbb{C})$ on W_λ, i.e., to $\dim W_\lambda(\mu') = \dim W_\lambda(\mu)$.

Now assume that λ is also in $\mathbb{Z}^n$. Then we have seen that there is a representation of $\mathrm{GL}_n(\mathbb{C})$ on W_λ inducing the representation of $\mathfrak{g}$. Let T be the subgroup of diagonal matrices in $\mathrm{GL}_n(\mathbb{C})$, we have $T = (\mathbb{C}^\times)^n$ and $\mathrm{Lie}(T) = \mathfrak{t}$. If $\mu = (\mu_1, \ldots, \mu_n) \in \mathbb{Z}^n$, we denote by e^μ the character of T given by $(z_1, \ldots, z_n) \longmapsto z_1^{\mu_1} \ldots z_n^{\mu_n}$. Note that $d(e^\mu) : \mathrm{Lie}(T) = \mathfrak{t} \to \mathbb{C}$ is just μ, see as an element of $\mathfrak{t}^*$. Let $\mu \in \mathfrak{t}^*$ such that $a_\mu \neq 0$. Then μ is of the form $\sigma(\lambda + \rho) - \rho - \sum_{\alpha \in \Phi^+} n_\alpha \alpha$, with $\sigma \in W$ and $n_\alpha \in \mathbb{Z}_{\geq 0}$, and in particular $\mu \in \mathbb{Z}^n$, so e^μ makes sense. By remark VI.8.3 of chapter VI, the weight space $W_\lambda(\mu)$ is stable by the action of T. As T is connected, the action of T on $W_\lambda(\mu)$ is determined by the action of $\mathfrak{t}$ (see remark VI.5.4 of chapter VI), and so it has to be given by multiplication by the character e^μ. As $W_\lambda = \bigoplus_{\mu \in \mathfrak{t}^*} W_\lambda(\mu)$, this implies that W_λ, as a representation of T, is isomorphic to $\bigoplus_{\mu \in \mathfrak{t}^*} a_\mu e^\mu$, where "$a_\mu e^\mu$" means "the direct sum of a_μ copies of the 1-dimensional representation e^μ".

Let $D \in \mathbb{C}[x_1, \ldots, x_n]$ be the polynomial defined by

$$D = \prod_{1 \leq i < j \leq n} (x_i - x_j) = x_1^{n-1} x_2^{n-2} \ldots x_{n-1} \prod_{1 \leq i < j \leq n} (1 - x_i^{-1} x_j).$$

Let $N_\lambda \in \mathbb{C}[x_1, \ldots, x_n]$ be the polynomial defined by

$$N_\lambda = \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma) \prod_{i=1}^n x_{\sigma(i)}^{\lambda_i + n - i}.$$

For every $\mu = (\mu_1, \ldots, \mu_n) \in \mathbb{Z}^n$, write $x^\mu = x_1^{\mu_1} \ldots x_n^{\mu_n}$. If $\sigma \in W$, we have

$$\sigma(\lambda + \rho) - \rho = (\lambda_{\sigma^{-1}(1)} - \sigma^{-1}(1) + 1, \ldots, \lambda_{\sigma^{-1}(n)} - \sigma^{-1}(n) + n),$$

hence

$$\prod_{i=1}^n x_{\sigma(i)}^{\lambda_i + n - i} = \prod_{i=1}^n x_i^{\lambda_{\sigma^{-1}(i)} - \sigma^{-1}(i) + n} = (x_1^{n-1} x_2^{n-2} \ldots x_{n-1}) x^{\sigma(\lambda+\rho)-\rho}.$$

So we get that

$$\frac{N_\lambda}{D} = \frac{\sum_{\sigma\in W}\operatorname{sgn}(\sigma)x^{\sigma(\lambda+\rho)-\rho}}{\prod_{1\leq i<j\leq n}(1-x_i^{-1}x_j)}.$$

Using the formula for χ_{W_λ} calculated above (and the explicit description of Φ^+), we finally get that the trace of an element $(z_1,\dots,z_n)$ of $(\mathbb{C}^\times)^n = T$ on W_λ is given by $\frac{N_\lambda(z_1,\dots,z_n)}{D(z_1,\dots,z_n)}$.

□

VII.6.20 Differential of a tensor product of representations

Let $\mathfrak{g}$ is the Lie algebra of a closed subgroup G of $\mathrm{GL}_n(\mathbb{C})$, let $\rho_1 : G \to \mathrm{GL}(V_1)$ and $\rho_2 : G \to \mathrm{GL}(V_2)$ be continuous finite-dimensional representations of G on complex vector spaces, and let $\rho_3 : G \to \mathrm{GL}(V_1 \otimes_{\mathbb{C}} V_2)$ and $\rho_4 : G \to \mathrm{GL}(\mathrm{Hom}_{\mathbb{C}}(V_1, V_2))$ be the tensor product and Hom representations.

Show that, for every $X \in \mathfrak{g}$,

$$d\rho_3(X) = d\rho_1(X) \otimes \mathrm{id}_{V_2} + \mathrm{id}_{V_1} \otimes d\rho_2(X)$$

and

$$d\rho_4(X)(f) = d\rho_2(X) \circ f - f \circ d\rho_1(X)$$

if $f \in \mathrm{Hom}_{\mathbb{C}}(V_1, V_2)$.

Solution. Let $v : \mathbb{R} \to V_1$, $w : \mathbb{R} \to V_2$ and $f : \mathbb{R} \to \mathrm{Hom}_{\mathbb{C}}(V_1, V_2)$ be three derivable functions. We want to prove the following formulas :

$$\frac{d}{dt}(v(t) \otimes w(t)) = v'(t) \otimes w(t) + v(t) \otimes w'(t)$$

and

$$\frac{d}{dt}(f(t)(v(t))) = f'(t)(v(t)) + f(t)(v'(t)).$$

The formulas of the problem will follows immediately from this and from the formula for the differential of a representation in (i) of theorem VI.5.2 of chapter VI.

Let $t \in \mathbb{R}$. For the first formula, we want to calculate the limit of $\frac{1}{h}(v(t+h) \otimes w(t+h) - v(t) \otimes w(t))$ as h goes to 0. We write $v(t+h) = v(t) + hv'(t) + h\varepsilon(h)$ and $w(t+h) = w(t) + hw'(t) + h\eta(h)$, where $\varepsilon : \mathbb{R} \to V_1$ and $\eta : \mathbb{R} \to V_2$ are functions that tend to 0 as $h \to 0$. Then

$$v(t+h)\otimes w(t+h) = v(t)\otimes w(t) + h(v'(t)\otimes w(t) + v(t)\otimes w'(t)) + h^2(v'(t)\otimes w'(t) + \varepsilon(h)\eta(h))$$

$$+h\eta(h)(v(t) + hv'(t)) + h\varepsilon(h)(w(t) + hw'(t)),$$

so

$$\lim_{h\to 0}\left(\frac{v(t+h)\otimes w(t+h)-v(t)\otimes w(t)}{h}-v'(t)\otimes w(t)-v(t)\otimes w'(t)\right)=0.$$

For the second formula, write $f(t+h)=f(t)+hf'(t)+hg(h)$, where $g:\mathbb{R}\to \mathrm{Hom}_{\mathbb{C}}(V_1,V_2)$ is a function that tends to 0 as $h\to 0$. Then

$$f(t+h)(v(t+h))=f(t)(v(t))+h(f'(t)(v(t))+f(t)(v'(t)))+h^2(f'(t)(v'(t)+\varepsilon(h)))$$

$$+h^2g(h)(\varepsilon(h))+h^2f'(t)(\varepsilon(h)))+h(g(h)(v(t))+f(t)(\varepsilon(h)),$$

so

$$\lim_{h\to 0}\left(\frac{f(t+h)(v(t+h))-f(t)(v(t))}{h}-f'(t)(v(t))-f(t)(v'(t))\right)=0.$$

□

VII.6.21 Casimir element

Show that the Casimir element of definition VI.14.2.1 of chapter VI is in the center of the universal enveloping algebra of $\mathfrak{sl}_n(\mathbb{C})$.

Solution. We use the notation of section VI.14.2 of chapter VI. By section VI.12.2 of the same chapter, if the E_{ij} are the elementary matrices in $M_n(\mathbb{C})$, then the Casimir element is given by

$$c=\frac{1}{2}\sum_{i=1}^{n-1}H_i^2+\sum_{i\neq j}E_{ij}E_{ji},$$

where $H_i=E_{ii}-E_{i+1,i+1}$ and all the products must be taken in the universal envelopping algebra of $\mathfrak{sl}_n(\mathbb{C})$ and *not* in $M_n(\mathbb{C})$. It suffices to show that c commutes with every element of a basis of $\mathfrak{sl}_n(\mathbb{C})$.

Instead of doing a direct calculation, I'll show you a particular instance of the general method of proof. Consider the bilinear form B on $\mathfrak{sl}_n(\mathbb{C})$ given by $B(X,Y)=\mathrm{Tr}(XY)$. This is a non-degenerate form, and a positive multiple of what is called in general the *Killing form*. [35] Let's denote by $(A_i)_{1\leq i\leq n^2-1}$ the basis $(H_1,\dots,H_{n-1},E_{ij},i\neq j)$ of $\mathfrak{sl}_n(\mathbb{C})$, for some arbitrary order on the set $\{(i,j)|i\neq j\}$. Then the dual basis (A_i^*) for the form B is $(\frac{1}{2}H_1,\dots,\frac{1}{2}H_{n-1},E_{ji},i\neq j)$. Note that the Casimir element is given by $c=\sum_{j=1}^{n^2-1}A_jA_j^*$, and we want to show that it commutes with every A_i. The key observation is that, for all $X,Y,Z\in\mathfrak{sl}_n(\mathbb{C})$, we have

$$B([X,Y],Z)+B(X,[Y,Z]).$$

[35] The Killing form is defined in general as the bilinear form $(X,Y)\longmapsto \mathrm{Tr}((\mathrm{ad}\,X)(\mathrm{ad}\,Y))$. In the case of $\mathfrak{sl}_n(\mathbb{C})$, it is relatively easy to check that $\mathrm{Tr}((\mathrm{ad}\,X)(\mathrm{ad}\,Y))=2n\mathrm{Tr}(XY)$.

(This follows directly from the definition of B.) Let $i \in \{1, \ldots, n^2-1\}$. By the observation above, the adjoint of $\operatorname{ad} A_i$ is $-\operatorname{ad} A_i$, so, if we write

$$[A_i, A_j] = \sum_{k=1}^{n^2-1} c_{ijk} A_k,$$

then

$$[A_i, A_j^*] = \sum_{k=1}^{n^2-1} c_{ikj} A_k^*.$$

Using this and the fact that $[X, YZ] = [X, Y]Z + Y[X, Z]$ for all $X, Y, Z \in U(\mathfrak{sl}_n(\mathbb{C}))$ (again an easy calculation), we get

$$[A_i, c] = \sum_{j=1}^{n^2-1} [A_i, A_j A_j^*] = \sum_{j=1}^{n^2-1} ([A_i, A_j]A_j^* + A_j[A_i, A_j^*]) = \sum_{j=1}^{n^2-1} \sum_{k=1}^{n^2-1} (c_{ijk} A_k A_j^* - c_{ikj} A_j A_k^*),$$

and the last sum is clearly 0 (separate the two terms, switch j and k in the second one).

□

VII.7 Exercises involving several chapters

VII.7.1 The algebraic Peter-Weyl theorem (chapters V and VI)

In this problem, $G = \mathrm{SU}(n)$, $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$ and we also use the algebraic group SL_n (over $\mathbb{C}$) defined in problem VII.6.5. We identify $\mathfrak{g}$ to $\mathrm{Lie}(G) \otimes_{\mathbb{R}} \mathbb{C}$ as in the proof of corollary VI.8.4 of chapter VI. So, by this corollary, we have a $1-1$ correspondence between continuous finite-dimensional representations of G (over $\mathbb{C}$) and finite-dimensional representations of $\mathfrak{g}$.

The Peter-Weyl theorem (theorem V.5.2 of chapter V) gives an injective map with dense image

$$\iota : \bigoplus_{(\rho, V_\rho) \in \widehat{G}} \mathrm{End}(V_\rho) \to L^2(G).$$

Remember that $\widehat{G}$ is the set of isomorphism classes of continuous irreducible representations of G on finite-dimensional $\mathbb{C}$-vector spaces. If $\rho : G \to \mathrm{GL}(V_\rho)$ is such a representation and $u \in \mathrm{End}(V_\rho)$, then $\iota(u)$ is by definition the function $g \longmapsto \mathrm{Tr}(\rho(g)^{-1} \circ u)$.

The goal of this problem is to describe the image of ι.

We use the definitions of problem VII.6.18 (morphisms of algebraic groups and their differentials), and we also make the following definitions : A *representation* of SL_n is a morphism of

algebraic groups (over $\mathbb{C}$) $\rho : \mathrm{SL}_n \to \mathrm{GL}_m$. We say that two representations $\rho : \mathrm{SL}_n \to \mathrm{GL}_m$ and $\rho' : \mathrm{SL}_n \to \mathrm{GL}_{m'}$ are *equivalent* if $m = m'$ and there exists $g \in \mathrm{GL}_m(\mathbb{C})$ such that $\rho' = g\rho g^{-1}$ (i.e. for every $\mathbb{C}$-algebra A, for every $x \in \mathrm{SL}_n(A)$, $\rho'_A(x) = g\rho_A(x)g^{-1}$). We say that a representation $\rho : \mathrm{SL}_n \to \mathrm{GL}_m$ is *irreducible* if the only $\mathbb{C}$-subspaces of $\mathbb{C}^m$ stable by SL_n are 0 and $\mathbb{C}^m$. (We say that a subspace V of $\mathbb{C}^m$ is stable by SL_n if for every $\mathbb{C}$-algebra A and for every $x \in \mathrm{SL}_n(A)$, $\rho_A(x)(V \otimes_{\mathbb{C}} A) \subset V \otimes_{\mathbb{C}} A$). Finally, we denote by $\widehat{\mathrm{SL}_n}$ the set of equivalence classes of irreducible representations of SL_n.

(1). If $\rho : \mathrm{SL}_n \to \mathrm{GL}_m$ is a representation, then the restriction of $\rho_{\mathbb{C}} : \mathrm{SL}_n(\mathbb{C}) \to \mathrm{GL}_m(\mathbb{C})$ to $\mathrm{SU}(n) \subset \mathrm{SL}_n(\mathbb{C})$ is a morphism of groups $\rho_c : \mathrm{SU}(n) \to \mathrm{GL}_m(\mathbb{C})$. Show that this induces a bijection $\widehat{\mathrm{SL}_n} \xrightarrow{\sim} \widehat{G}$.

(2). Remember the definition of the ring of regular functions R_{SL_n} in problem VII.6.17. If $f \in R_{\mathrm{SL}_n}$, then we get a "polynomial" map $f_{\mathbb{C}} : \mathrm{SL}_n(\mathbb{C}) \to \mathbb{C}$, and we can restrict it to $G = \mathrm{SU}(n) \subset \mathrm{SL}_n(\mathbb{C})$. Show that the resulting map $r : R_{\mathrm{SL}_n} \to L^2(G)$ is injective.

(3). Show that $r(R_{\mathrm{SL}_n})$ contains the image of ι.

(4). If $f \in R_{\mathrm{SL}_n}$, show that there exist $f_1, \ldots, f_r, h_1, \ldots, h_r \in R_{\mathrm{SL}_n}$ such that, for every $\mathbb{C}$-algebra A and every $x, y \in \mathrm{SL}_n(A)$,

$$f(xy) = \sum_{i=1}^{r} f_i(x)h_i(y).$$

(You may admit the (easy) fact that the ring of regular functions on $\mathrm{SL}_n \times \mathrm{SL}_n$ is $R_{\mathrm{SL}_n} \otimes_{\mathbb{C}} R_{\mathrm{SL}_n}$.)

(5). For every $f \in R_{\mathrm{SL}_n}$, show that the subrepresentation of $G \times G$ generated by $r(f)$ in $L^2(G)$ is finite-dimensional.

(6). Show that $r(R_{\mathrm{SL}_n})$ is equal to the image of ι.

In other words, the image of ι is the subalgebra of polynomials functions on $\mathrm{SU}(n)$. This is also a general fact; for example, it will be true for all connected compact subgroups of $\mathrm{GL}_n(\mathbb{C})$, with the appropriate changes.

Solution.

(1). We have in problem VII.6.18 that, if $\rho : \mathrm{SL}_n \to \mathrm{GL}_m$ is a representation, then $\rho_{\mathbb{C}} : \mathrm{SL}_n(\mathbb{C}) \to \mathrm{GL}_m(\mathbb{C})$ is a continuous morphism of groups, whose differential $d\rho_{\mathbb{C}} : \mathfrak{sl}_n(\mathbb{C}) \to \mathfrak{gl}_m(\mathbb{C})$ is equal to the differential $d\rho$ defined algebraically, and that $d\rho$ (or $d\rho_{\mathbb{C}}$) uniquely determines ρ. Also, by corollary VI.8.4 of chapter VI, $d\rho_{\mathbb{C}}$ is uniquely determined by its restriction to $\mathfrak{su}(n)$, which is also equal to $d\rho_c$. So the map $\rho \longmapsto \rho_c$ induces an injection from the set of morphisms of algebraic groups $\mathrm{SL}_n \to \mathrm{GL}_m$ to the set of continuous morphisms of groups $\mathrm{SU}(n) \to \mathrm{GL}_m(\mathbb{C})$.

Let $\rho : \mathrm{SL}_n \to \mathrm{GL}_m$ be a morphism of algebraic groups, let $\rho_c : \mathrm{SU}(n) \to \mathrm{GL}_m(\mathbb{C})$ be the corresponding continuous morphism of groups. Obviously, if the subspace V of $\mathbb{C}^m$ is

stable by SL_n, it is stable by $\mathrm{SU}(n)$. Conversely, let V be a subspace of $\mathbb{C}^m$, and suppose that it is stable by $\mathrm{SU}(n)$. Then, by remark VI.8.3 and corollary VI.8.4 of chapter VI, V is also stable by $\mathfrak{su}(n)$, hence by $\mathfrak{sl}(n)$, hence by $\mathrm{SL}_n(\mathbb{C})$. We want to show that it is stable by SL_n. Let A be a $\mathbb{C}$-algebra, let $g \in \mathrm{SL}_n(A)$, and let $x \in V \otimes_{\mathbb{C}} A$. As in the solution of problem VII.6.18(3), we can find a subfield k of $\mathbb{C}$ of finite transcendence degree over $\mathbb{Q}$ and a k-subspace V_0 of k^m such that ρ makes sense as a morphism of algebraic groups over k and $V = V_0 \otimes_k \mathbb{C}$, and we can find morphisms of k-algebras $u : B' \to A$ and $v : B' \to \mathbb{C}$, and elements $g' \in \mathrm{SL}_n(B)$ and $x' \in V_0 \otimes_k B'$ such that $g = \mathrm{SL}_n(u)(g')$, $x = (\mathrm{id}_{V_0} \otimes_k u)(x')$ and v is injective. Let $g'' = \mathrm{SL}_n(v)(g') \in \mathrm{SL}_n(\mathbb{C})$ and $x'' = (\mathrm{id}_{V_0} \otimes v)(x') \in V_0 \otimes_k \mathbb{C} = V$. Then $\rho_{\mathbb{C}}(g'')(x'') \in V$, so $\rho_{B'}(x') \in V_0 \otimes_k B'$, and finally $\rho_A(g)(x) \in V_0 \otimes_k A = V \otimes_{\mathbb{C}} A$.

By the results proved in the above paragraph, ρ is irreducible if and only if ρ_c is irreducible. Also, the equivalence relations on both sides are given by conjugating by elements of $\mathrm{GL}_m(\mathbb{C})$. So we have shown that the construction $\rho \longmapsto \rho_c$ induces an injective map $\widehat{\mathrm{SL}_n} \to \widehat{\mathrm{SU}(n)}$. It remains to show that this map is surjective, i.e., that every irreducible representation of $\mathrm{SU}(n)$ (on a finite-dimensional $\mathbb{C}$-vector space) comes from a representation of the algebraic group SL_n. For this, we use theorem VI.11.5 of chapter VI, and we also use its notation. Let ρ_c be an irreducible representation of $\mathrm{SU}(n)$. By this theorem, it is given by a $\lambda \in \Lambda^+$. Writing $\lambda = d_1\varpi_1 + \cdots + d_{n-1}\varpi_{n-1}$ as in the proof of the theorem, we can realize ρ_c on some subspace W of the representation $V := \bigotimes_{i=1}^{n-1} (\Lambda^i \mathbb{C}^n)^{\otimes d_i}$. Now, observe that the standard representation of $\mathrm{SU}(n)$ on $\mathbb{C}^n$ (i.e. the one coming from the inclusion $\mathrm{SU}(n) \subset \mathrm{GL}_n(\mathbb{C})$) comes from a representation of SL_n (given by the inclusion $\mathrm{SL}_n \subset \mathrm{GL}_n$), as do its exterior powers and any tensor product of these (because the construction of the exterior power representations, and of tensor products of representations, make sense over any ring of coefficients), so the representation of $\mathrm{SU}(n)$ on V comes from a representation of SL_n. But we saw above that $\mathrm{SU}(n)$ and SL_n have the same invariant subspaces in any representation, so W is also invariant by SL_n, hence gives a representation ρ of SL_n such that $\rho_c = \rho_{\mathbb{C}|\mathrm{SU}(n)}$. We have already shown that this ρ has to be irreducible, so we are done.

(2). The map $r : R_{\mathrm{SL}_n} \to L^2(G)$ is obviously a map of $\mathbb{C}$-algebras, so we have to show that its kernel is trivial. Let $f \in R_{\mathrm{SL}_n}$ be such that $f_{\mathbb{C}|\mathrm{SU}(n)} = 0$. We want to show that $f = 0$. By problem VII.6.17, it suffices to show that, for every $\mathbb{C}$-algebra A and every $g \in \mathrm{SL}_n(\mathbb{C})$, $f(g) = 0$. Using the same trick as in the proof of (3) of problem VII.6.18, we see that it actually suffices to show that $f(g) = 0$ for every $g \in \mathrm{SL}_n(\mathbb{C})$. Let $g \in \mathrm{SL}_n(\mathbb{C})$. The polar decomposition for matrices says that we can write $g = su$, with s Hermitian positive definite and $u \in \mathrm{SU}(n)$.[36] As s is Hermitian positive definite, there exists $h \in \mathrm{SU}(n)$ such that hsh^* is diagonal with real positive eigenvalues $\lambda_1, \ldots, \lambda_n$. Let A be the diagonal

[36]These s and u are actually uniquely determined by g. Compare with problem VII.5.4(11)(c)(iii). If you've never the polar decomposition, here is how to prove existence : The matrix g^*g is Hermitian definite positive, so it can be diagonalized in an orthogonal basis and has real positive eigenvalues, so we can find a Hermitian definite positive matrix s (take it to be diagonalizable in the same basis as g^*g) such that $s^2 = g^*g$. As $\det(s)^2 = |\det(g)|^2 = 1$ and $\det(s) \in \mathbb{R}_{>0}$, s is in $\mathrm{SL}_n(\mathbb{C})$. Now if $u = gs^{-1}$, then $u^*u = s^{-1}g^*gs^{-1} = 1$ so $u \in \mathrm{U}(n)$, and $\det(u) = \det(g)\det(s)^{-1} = 1$ so $u \in \mathrm{SU}(n)$.

matrix with eigenvalues $\log\lambda_1, \ldots, \log\lambda_n$, then $\mathrm{Tr}(A) = \log\det(s) = 0$. Now consider the function $F : \mathbb{C} \to \mathbb{C}$ defined by $h(t) = f(h^* e^{tA} hu)$. This is well-defined, because $e^{tA} \in \mathrm{SL}_n(\mathbb{C})$ for every $t \in \mathbb{C}$. It is also a holomorphic function, because the matrix exponential is defined by an absolutely convergent power series, and f is a polynomial function. If $t \in i\mathbb{R}$, then $e^{tA} \in \mathrm{SU}(n)$, so $h^* e^{tA} hu \in \mathrm{SU}(n)$, so $F(t) = 0$ by the hypothesis on f. By the identity theorem for holomorphic functions, F is identically 0. Now note that, if $t = 1$, then $h^* e^{tA} hu = su = g$. So $f(g) = 0$.

(3). Remember that ι is the direct sum of the maps ι_ρ defined just before theorem V.5.2 of chapter V. So we have to show that the image of r contains the image of all the ι_ρ. Let $\rho : \mathrm{SU}(n) \to \mathrm{GL}(V_\rho)$ be an irreducible representation. We choose an isomorphism $V_\rho \simeq \mathbb{C}^m$. By (1) we have a morphism of algebraic groups $\rho' : \mathrm{SL}_n \to \mathrm{GL}_m$ such that $\rho = \rho'_{\mathbb{C}|\mathrm{SU}(n)}$. Now remember that ι_ρ is the function $M_m(\mathbb{C}) \to L^2(\mathrm{SU}(n))$ sending $u \in M_m(\mathbb{C})$ to $g \longmapsto \mathrm{Tr}(\rho(g)^{-1} \circ u)$. If $u \in M_m(\mathbb{C})$, then, for every $\mathbb{C}$-algebra A, the function $\mathrm{SL}_n(A) \to A$, $g \longmapsto \mathrm{Tr}(\rho'_A(g)^{-1} \circ u)$, makes sense, and this gives an element f of R_{SL_n} such that $r(f) = \iota_\rho(u)$.

(4). Let $f \in R_{\mathrm{SL}_n}$. We define a function $F \in R_{\mathrm{SL}_n \times \mathrm{SL}_n}$ [37] in the following way : For every $\mathbb{C}$-algebra A, for every $g_1, g_2 \in \mathrm{SL}_n(A)$, $F(g_1, g_2) = f(g_1 g_2)$. Note that the map $R_{\mathrm{SL}_n} \otimes_{\mathbb{C}} R_{\mathrm{SL}_n} \to R_{\mathrm{SL}_n \times \mathrm{SL}_n}$ sends $h_1 \otimes h_2$ to the regular function given on $\mathrm{SL}_n(A) \times \mathrm{SL}_n(A$ by $(g_1, g_2) \longmapsto f_1(g_1) f_2(g_2)$, for any $\mathbb{C}$-algebra A.

By the fact that we admitted,[38] we can write $F = \sum_{i=1}^r f_i \otimes h_i$, with $f_1, \ldots, f_r, h_1, \ldots, h_r \in R_{\mathrm{SL}_n}$. This immediately gives the conclusion.

(5). Let $f \in R_{\mathrm{SL}_n}$. By (4), we can find element $h_1, \ldots, h_r, h'_1, \ldots, h'_r, h''_1, \ldots, h''_r \in R_{\mathrm{SL}_n}$ such that, for every $\mathbb{C}$-algebra A and every $x, y, z \in \mathrm{SL}_n(A)$,

$$f(xyz) = \sum_{i=1}^{r} h_1(x) h'_1(y) h''_1(z).$$

By definition of the action of $G \times G$ on $L^2(G)$, the subrepresentation of $G \times G$ generated by $r(f) \in L^2(G)$ is the span of all the functions $L_{x^{-1}} R_y r(f) : G \to \mathbb{C}$, $g \longmapsto f(x^{-1} g y)$, for $x, y \in \mathrm{G}$. By the formula above, this is contained in $\mathrm{Span}(r(h'_1), \ldots, r(h'_r))$, so it is finite-dimensional.

(6). We have seen in the proof of theorem V.2 of chapter V that, if we make G act on $L^2(G)$ by the left regular action, then every finite-dimensional G-representation of $L^2(G)$ is contained in $\mathrm{Im}(\iota)$. So by (5), $\mathrm{Im}(r) \subset \mathrm{Im}(\iota)$. But we have seen in (4) that $\mathrm{Im}(\iota) \subset \mathrm{Im}(r)$, so finally $\mathrm{Im}(\iota) = \mathrm{Im}(r)$.

□

[37] Note that $\mathrm{SL}_n \times \mathrm{SL}_n$ is an algebraic subgroup of GL_{2n} : just take the matrices in GL_{2n} that have two diagonal blocks, both of determinant 1.

[38] And that is very easy to prove using the description of regular functions as polynomials on the group, i.e., the first description in problem VII.6.17.

VII.7.2 Polarization

This is actually just a lemma for the next problem, problem VII.7.3.

Let V and W be finite-dimensional $\mathbb{C}$-vector spaces and $f : V^d \to W$ be a symmetric d-linear form. If $f(x, \ldots, x) = 0$ for every $x \in V$, show that $f = 0$.

Hint : You can approach this problem in at least two ways. If you are good with algebraic manipulations, you can find (and prove) the formula giving f from the function $V \to W$, $x \longmapsto f(x, \ldots, x)$. (This will actually work for modules over any ring where $d!$ is invertible, not just vector spaces over $\mathbb{C}$). Or you could use representation theory : first prove that the dth symmetric power of the standard representation of $\mathfrak{sl}(V)$ is irreducible, then find a way to apply this to the question.

Solution. Write $D(x) = f(x, \ldots, x)$.

Let's take the first hint and suppose that V and W are k-modules, where k is a commutative ring. We will show that, for every $x_1, \ldots, x_d \in V$,

$$d! f(x_1, \ldots, x_d) = \sum_{\varnothing \neq S \subset \{1,\ldots,d\}} (-1)^{d-|S|} D\left(\sum_{i \in S} x_i\right).$$

If $d!$ is invertible in k, this clealry implies the result.

Let $S \subset \{1, \ldots, d\}$ be nonempty. Then

$$D\left(\sum_{i \in S} x_i\right) = \sum_{i_1,\ldots,i_d \in S} f(x_{i_1}, \ldots, x_{i_d}) = \sum_{\substack{i_1 \leq \cdots \leq i_d \\ i_1,\ldots,i_d \in S}} N(i_1, \ldots, i_d) f(x_{i_1}, \ldots, x_{i_d}),$$

where

$$N(i_1, \ldots, i_d) = |\{(i_{\sigma(1)}, \ldots, i_{\sigma(d)}) \in \mathbb{Z}^d, \sigma \in \mathfrak{S}_d\}|.$$

Note that $N(i_1, \ldots, i_d)$ only depends on $i_1, \ldots, i_d$, not on S. So we get

$$\sum_{\varnothing \neq S \subset \{1,\ldots,d\}} (-1)^{d-|S|} D\left(\sum_{i \in S} x_i\right) = \sum_{\varnothing \neq S \subset \{1,\ldots,d\}} (-1)^{d-|S|} \sum_{\substack{i_1 \leq \cdots \leq i_d \\ i_1,\ldots,i_d \in S}} N(i_1, \ldots, i_d) f(x_{i_1}, \ldots, x_{i_d}),$$

which by the fact that $N(i_1, \ldots, i_d) f(x_{i_1}, \ldots, x_{i_d})$ doesn't depend on S is equal to

$$\sum_{1 \leq i_1 \leq \cdots \leq i_d \leq d} N(i_1, \ldots, i_d) f(x_{i_1}, \ldots, x_{i_d}) \sum_{\{i_1,\ldots,i_d\} \subset S \subset \{1,\ldots,d\}} (-1)^{d-|S|}.$$

Suppose that $\{i_1, \ldots, i_d\} \subsetneq \{1, \ldots, d\}$, and let $T = \{1, \ldots, d\} - \{i_1, \ldots, i_d\}$. Then

$$\sum_{\{i_1,\ldots,i_d\} \subset S \subset \{1,\ldots,d\}} (-1)^{d-|S|} = (-1)^{|T|} \sum_{S' \subset T} (-1)^{|S'|} = (1-1)^{|T|} = 0.$$

So only the term with $i_1 = 1, \ldots, i_d = d$ survives in the sum above. Note also that $N(1, \ldots, d) = d!$. Finally, we get

$$\sum_{\varnothing \neq S \subset \{1,\ldots,d\}} (-1)^{d-|S|} D\left(\sum_{i \in S} x_i\right) = d! f(x_1, \ldots, x_d),$$

as desired.

Suppose that we wanted to apply the second hint. This time we take V and W to be finite-dimensional vector spaces over $\mathbb{C}$. We may asumme that $V = \mathbb{C}^n$. By problem VII.6.14, the representation of $\mathfrak{sl}_n(\mathbb{C})$ on the symmetric power $S^d(V)$ is irreducible for every $d \geq 0$. By remark VI.8.3 of chapter VI, the representation of $\mathrm{SU}(n)$ on that same symmetric power (induced by the standard representation on $\mathbb{C}^n$) is also irreducible.

Let U be the subspace of $S^d(V)$ spanned by all the elements of the form $x \otimes \cdots \otimes x$, $x \in V$. This space is nonzero, and it is clearly stable by $\mathrm{SU}(n)$. As $S^d(V)$ is an irreducible representation of $\mathrm{SU}(n)$, we get that $U = S^d(V)$.

The fact that f is symmetric says that f factors through a linear map $S^d(V) \to W$, that we will still call f. (See problem VII.6.7(4).) The hypothesis says that $f_{|U} = 0$. But we have just seen that $U = S^d(V)$, so $f_{|U} = 0$ implies that $f = 0$.

□

VII.7.3 Pseudo-characters (chapters I and II)

Historical remarks : Pseudo-characters were first introduced by Wiles ([36]) and Taylor ([35]) to study the deformation rings of representations of absolute Galois groups of number fields. The theory was then developed more systematically in papers of Nyssen ([22]) and Rouquier ([24]). The original definition of pseudo-characters of degree d does not work well if $d!$ is not invertible in the coefficient ring. In his article [7], Chenevier introduced a refinement of pseudo-characters, called *determinants*, that have the expected properties in all characteristics.

The exposition here follows section 2 of Bellaiche's notes [2], with some help from Dotsenko's notes [9].

In this problem, k is a commutative ring and R is a (not necessarily commutative) k-algebra.

A *central function* on R is a k-linear map $f : R \to k$ such that $f(xy) = f(yx)$ for every $x, y \in R$. If $f : R \to k$ is a central function and r is a positive integer, we define a function $S_r(f) : R^{\otimes r} \to k$ in the following way : For every $\sigma \in \mathfrak{S}_r$, let $\sigma = c_1 \ldots c_m$ be its decomposition into cycles with disjoint supports, write $c_i = (a_{i1} \ldots a_{in_i})$, and define $f_\sigma : R^{\otimes r} \to k$ by

$$f_\sigma(x_1 \otimes \cdots \otimes x_r) = \prod_{i=1}^{r} f(x_{a_{i1}} \ldots x_{a_{in_i}}).$$

Then $S_r(f)$ is given by

$$S_r(f) = \sum_{\sigma \in \mathfrak{S}_r} \mathrm{sgn}(\sigma) f_\sigma.$$

Let $d \in \mathbb{Z}_{\geq 1}$. We say that a central function $f : R \to k$ is a *pseudo-character of degree* d if $S_d(f)$ is not identically 0 and $S_{d+1}(f) = 0$. (Pseudo-characters are also often called pseudo-representations.)

The idea is that a pseudo-character of degree d looks like the character (i.e. the trace) of a representation of R on a free k-module of dimension d. In this problem, we will make the previous statement more precise.

VII.7.3.1 First properties and an example

(1). Show that the definition of $S_r(f)$ above doesn't depend on the choices and makes sense.

(2). Show that, for every $r \geq 1$, the r-linear map $S_r(f) : R^{\otimes r} \to k$ is symmetric.

(3). Show that $S_1(f) = f$ and that, for every $r \geq 1$,

$S_{r+1}(f)(x_1 \otimes \cdots \otimes x_{r+1}) =$

$$f(x_{r+1}) S_r(f)(x_1 \otimes \cdots \otimes x_r) - \sum_{i=1}^{r} S_r(f)(x_1 \otimes \cdots \otimes x_{i-1} \otimes (x_i x_{r+1}) \otimes x_{i+1} \otimes \cdots \otimes x_r).$$

(4). Show that a pseudo-character of degree 1 is the same as a nonzero k-linear map that respects multiplication.

(5). Suppose that k is local and that $d!$ is invertible in k. If $f : R \to k$ is a pseudo-character of degree d, show that $f(1) = d$. *(Hint : Use the relation between $S_{r+1}(x_1, \ldots, x_r, 1)$ and $S_r(x_1, \ldots, x_r)$ to calculate $S_{d+1}(1, \ldots, 1)$.)*

(6). Suppose that $R = M_r(k)$ and that $f : R \to k$ is a k-linear central function, and let $d = f(1)$.

Show that r divides d (in k) and that $f = \frac{d}{r}\mathrm{Tr}$.

(7). Let $\mathbb{H}$ be the $\mathbb{R}$-algebra of quaternions (see problem VII.1.6). Consider the function $f : \mathbb{H} \to \mathbb{R}$ given by $f(a + bi + cj + dk) = 2a$, for all $a, b, c, d \in \mathbb{R}$. Show that f is a pseudo-character of degree 2, but that there is no representation $u : \mathbb{H} \to M_2(\mathbb{R})$ such that $f = \mathrm{Tr} \circ u$.

Solution.

(1). First, the formula giving f_σ is clearly linear in each x_i, so it does define a function on $R^{\otimes r}$. We have to check that the definition of f_σ doesn't depend on the choices. It doesn't

depend on the ordering we choose on the cycles because k is commutative, and it doesn't depend on the way we write each cycles because f is a central function (so, for every $x_1, \dots, x_r \in R$ and $i \in \{1, \dots, r\}$, $f(x_1 \dots x_r) = f(x_{i+1} \dots x_r x_1 \dots x_i)$).

(2). Let $\tau \in \mathfrak{S}_r$. Then, for every $\sigma \in \mathfrak{S}_r$, if $\sigma = c_1 \dots c_\ell$ is its decomposition of cycles with disjoint supports and $c_i = (r_{i,1} \dots r_{i,j_i})$, the decomposition in cycles of $\tau\sigma\tau^{-1}$ is $\tau\sigma\tau^{-1} = d_1 \dots d_\ell$, with $d_i = (\tau(r_{i,1}), \dots, \tau(r_{i,j_i}))$. So

$$f_\sigma(x_{\tau(1)} \otimes \cdots \otimes x_{\tau(r)}) = f_{\tau\sigma\tau^{-1}}(x_1 \otimes \cdots \otimes x_r).$$

As $\sigma \longmapsto \tau^{-1}\sigma\tau$ is an automorphism of $\mathfrak{S}_r$ that preserves sgn, we get that

$$S_r(f)(x_{\tau(1)} \otimes \cdots \otimes x_{\tau(r)}) = S_r(f)(x_1 \otimes \cdots \otimes x_r).$$

(3). The first equality is obvious.

Let $r \geq 1$. For every $i \in \{1, \dots, r+1\}$, let $C_i = \{\sigma \in \mathfrak{S}_{r+1} \,|\sigma(i) = r+1\}$. If $i = r+1$, C_{r+1} is a subgroup that canonically identities to $\mathfrak{S}_r$, and we have

$$\sum_{\sigma \in C_{r+1}} \mathrm{sgn}(\sigma) f_\sigma(x_1 \otimes \cdots \otimes x_{r+1}) = f(x_{r+1}) \sum_{\sigma \in \mathfrak{S}_r} \mathrm{sgn}(\sigma) f_\sigma(x_1 \otimes \cdots \otimes x_r).$$

If $1 \leq i \leq r$, then we have a bijection $C_i \xrightarrow{\sim} \mathfrak{S}_r$ sending $\sigma \in C_i$ to the element τ of $\mathfrak{S}_r$ defined by $\tau(j) = \sigma(j)$ if $j \neq i$, and $\tau(i) = \sigma^2(i) = \sigma(r+1)$. We have $\mathrm{sgn}(\tau) = -\mathrm{sgn}(\sigma)$, and

$$f_\sigma(x_1 \otimes \cdots \otimes x_{r+1}) = f_\tau(x_1 \otimes \cdots \otimes (x_i x_{r+1}) \otimes \cdots \otimes x_r).$$

This proves the equality of the question.

(4). Let $f : R \to k$ be a pseudo-character of degree 1. Then f is k-linear, and $S_1(f) = f \neq 0$. Also, for every $x_1, x_2 \in R$,

$$S_2(f)(x_1 \otimes x_r) = f(x_2)f(x_1) - f(x_1 x_2) = 0,$$

so f is multiplicative.

Conversely, a nonzero multiplicative k-linear map $f : R \to k$ is clearly a pseudo-character of degree 1.

(5). If $r \geq 1$ and $x_1, \dots, x_r \in R$, then, by question (2),

$$S_{r+1}(f)(x_1 \otimes \cdots \otimes x_r \otimes 1) = (f(1) - r) S_r(f)(x_1 \otimes \cdots \otimes x_r).$$

So, by an easy induction, for every $r \geq 1$,

$$S_r(f)(1 \otimes \cdots \otimes 1) = \prod_{i=0}^{r-1} (f(1) - i).$$

As f is a pseudo-character of degree d, $S_{d+1}(f) = 0$, and so $f(1)(f(1)-1)\dots(f(1)-d) = 0$, and there is a $i \in \{0,\dots,d\}$ such that $f(1)-i$ is in the maximal ideal of k. For every $j \in \{0,\dots,d\} - \{i\}$, $i-j$ is invertible in k (because $d!$ is invertible in k), so $f(1)-j$ is not in the maximal ideal of k, and hence $f(1)-j \in k^\times$. This shows that actually $f(1)-i = 0$, i.e. $f(1) = i$. Now suppose that $i \neq d$, then $f(1)-d \in k^\times$, and, so, for every $x_1,\dots,x_d \in R$,

$$0 = (f(1)-d)^{-1} S_{d+1}(f)(x_1 \otimes \cdots \otimes x_d \otimes 1) = S_d(f)(x_1 \otimes \cdots \otimes x_d),$$

contradicting the fact that $S_d(f) \neq 0$. So $i = d$, i.e. $f(1) = d$.

(6). Write E_{ij} for the matrix that has (i,j)-entry equal to 1 and all its other entries equal to 0. Then $E_{ij}E_{ji} = E_{ii}$ and $E_{ji}E_{ij} = E_{jj}$ for every i,j, so $f(E_{11}) = \cdots = f(E_{dd})$. In particular,

$$f(1) = f(E_{11} + \cdots + E_{rr}) = r f(E_{11}),$$

so r divides d and $f(E_{11}) = \frac{d}{r}$.

On the other hand, if $i \neq j$, then $E_{ij} = E_{ii}E_{ij}$ and $E_{ij}E_{ii} = 0$, so $f(E_{ij}) = f(0) = 0$. If $A = (a_{ij}) \in M_r(k)$, then $A = \sum_{i,j=1}^r a_{ij}E_{ij}$, so we get

$$f(A) = \sum_{i,j=1}^r a_{ij} f(E_{ij}) = f(E_{11})(a_1 + \cdots + a_r) = \tfrac{d}{r}\mathrm{Tr}(A).$$

(7). Remember from problem VII.1.6 that we have an isomorphism of $\mathbb{C}$-algebras $u : \mathbb{H} \otimes_{\mathbb{R}} \mathbb{C} \xrightarrow{\sim} M_2(\mathbb{C})$, and note that f is simply the function $x \longmapsto \mathrm{Re}(\mathrm{Tr}(u(x \otimes 1)))$. Now the fact that f is a pseudo-character of degree 2 follows from the next part of the problem.

Of course, it's also possible to solve this question by a direct calculation.

□

VII.7.3.2 A character is a pseudo-character

Let $u : R \to M_d(k)$ be a k-algebra map (i.e. a representation of R on the k-module k^d, or a left R-module structure on k^d).

The goal of this question is to show that the central map $f := \mathrm{Tr} \circ u : R \to k$ is a pseudo-character of degree $\leq d$, and that its degree is exactly d if k is local and $d!$ is invertible in k.

(1). Show that we may assume that $R = M_d(k)$ and $u = \mathrm{id}$ (and hence $f = \mathrm{Tr}$).

(2). Show that we may assume that $k = \mathbb{C}$.

(3). Let r be a positive integer. We make the group $\mathfrak{S}_r$ act on $(\mathbb{C}^d)^{\otimes r}$ by permuting the factors (i.e. $\sigma(v_1 \otimes \cdots \otimes v_r) = v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(r)}$). We make $M_d(\mathbb{C})$ act on $(\mathbb{C}^d)^{\otimes r}$ by the usual tensor power action (i.e. $x(v_1 \otimes \cdots \otimes v_r) = (xv_1) \otimes \cdots \otimes (xv_r)$).[39] Show that, for every $x \in M_d(\mathbb{C})$,

$$S_r(f_0)(x \otimes \cdots \otimes x) = \sum_{\sigma \in \mathfrak{S}_r} \mathrm{sgn}(\sigma) \mathrm{Tr}(x\sigma, (\mathbb{C}^d)^{\otimes r}),$$

where $f_0 : M_d(\mathbb{C}) \to \mathbb{C}$ is the trace.

(4). If $r \geq d+1$, show that the endomorphism $\sum_{\sigma \in \mathfrak{S}_r} \mathrm{sgn}(\sigma)\sigma$ of $(\mathbb{C}^d)^{\otimes r}$ is zero.

(5). Finish the proof that f is a pseudo-character of degree $\leq d$.

(6). If moreover k is local and $d!$ is invertible in k, show that the pseudo-character f is of degree d.

Solution.

(1). As u is a map of k-algebras, we have $S_r(\mathrm{Tr} \circ u) = S_r(\mathrm{Tr}) \circ u$ for every $r \geq 1$. So, if we know that Tr is a pseudo-character of degree $\leq d$, this will imply immediately that f is also a pseudo-character of degree $\leq d$. Suppose that k is local and that $d!$ is invertible in k, and let $d' \leq d$ be the degree of f. By question (5) of the previous part, we have $d' = f(1) = \mathrm{Tr}(u(1)) = \mathrm{Tr}(1) = d$. If $d' \neq d$, then $d' - d$ is invertible in k (because $d!$ is), which is impossible. So $d' = d$.

(2). Suppose that we know that $\mathrm{Tr} : M_d(k) \to k$ is a pseudo-character of degree d when $k = \mathbb{C}$.

Now let k be any commutative ring, let $r \geq 1$, and let $A^{(1)} = (a_{ij}^{(1)}), \ldots, A^{(r)} = (a_{ij}^{(r)}) \in M_d(k)$. Consider the polynomials ring $k' = \mathbb{Z}[X_{ij}^{(s)}, 1 \leq s \leq r, 1 \leq i, j \leq d]$, and let $B^{(s)} = (X_{ij}^{(s)}) \in M_d(k')$, for $1 \leq s \leq r$. We have a morphism of rings $\varphi : k' \to k$ sending each $X_{ij}^{(s)}$ to $a_{ij}^{(s)}$, and the corresponding morphism $\varphi : M_d(k') \to M_d(k)$ sends $B^{(s)}$ to $A^{(s)}$. Choosing rd^2 algebraically independent elements in $\mathbb{C}$, we also get an injective morphism of rings $\psi : k' \to \mathbb{C}$, and we still use ψ to denote the morphism $M_d(k') \to M_d(\mathbb{C})$. Because φ and ψ are morphisms of rings, we have

$$\varphi(S_r(\mathrm{Tr})(B^{(1)}, \ldots, B^{(r)})) = S_r(\mathrm{Tr})(A^{(1)}, \ldots, A^{(r)})$$

and

$$\psi(S_r(\mathrm{Tr})(B^{(1)}, \ldots, B^{(r)})) = S_r(\mathrm{Tr})(\psi(B^{(1)}), \ldots, \psi(B^{(r)})).$$

If $r = d+1$, this gives $\psi(S_{d+1}(\mathrm{Tr})(B^{(1)}, \ldots, B^{(d+1)})) = 0$ by the hypothesis. As ψ is injective, we get $S_{d+1}(\mathrm{Tr})(B^{(1)}, \ldots, B^{(d+1)}) = 0$, and applying φ gives $S_{d+1}(\mathrm{Tr})(A^{(1)}, \ldots, A^{(d+1)}) = 0$. So Tr is a pseudo-character of degree $\leq d$.

[39] Note that we think of $M_d(\mathbb{C})$ as an associative algebra and not as a Lie algebra here, so the action of $M_d(\mathbb{C})$ on tensor powers of $\mathbb{C}^d$ is given by the usual "diagonal" action.

Suppose that k is local and $d!$ is invertible in k, and let $d' \leq d$ be the degree of Tr. Then $d' = \mathrm{Tr}(1) = d$ (by question (4) of the previous part), and thius gives $d = d'$ as in question (1).

(3). Both sides of the equality that we are trying to prove are continuous in x, and they don't change if we replace x by gxg^{-1}, with $g \in \mathrm{GL}_n(\mathbb{C})$. As diagonalizable matrices are dense in $M_n(\mathbb{C})$, it suffices to prove the equality for x a diagonal matrix, say $x = \mathrm{diag}(x_1, \ldots, x_d)$.

Let $(e_1, \ldots, e_d)$ be the canonical basis of $\mathbb{C}^d$. Then a basis of $(\mathbb{C}^d)^{\otimes r}$ is given by $(e_{i_1} \otimes \cdots \otimes e_{i_r})_{1 \leq i_1, \ldots, i_r \leq d}$. Suppose that σ is a r-cycle. Then

$$(x\sigma)(e_{i_1} \otimes \cdots \otimes e_{i_r}) = x_{\sigma^{-1}(i_1)} \ldots x_{\sigma^{-1}(i_r)} (e_{\sigma^{-1}(i_1)} \otimes \cdots \otimes e_{\sigma^{-1}(i_r)}),$$

and this is propotional to $e_{i_1} \otimes \cdots \otimes e_{i_r}$ if and only if $i_1 = \cdots = i_r$. So

$$\mathrm{Tr}(x\sigma, (\mathbb{C}^d)^{\otimes r}) = \sum_{i=1}^{d} x_i^r = \mathrm{Tr}(x^r).$$

Now if σ is any element of $\mathfrak{S}_r$, let $\sigma = c_1 \ldots c_\ell$ be its decomposition into cycles with disjoint supports $I_1, \ldots, I_\ell \subset \{1, \ldots, r\}$. Then we have (by the formula for the trace of a tensor product of maps, see the proof of proposition II.1.1.3 in chapter II)

$$\mathrm{Tr}(x\sigma, (\mathbb{C}^d)^{\otimes r}) = \prod_{i=1}^{\ell} \mathrm{Tr}(xc_i, (\mathbb{C}^d)^{\otimes I_i}.$$

By the previous calculation, this is equal to $\prod_{i=1}^{\ell} \mathrm{Tr}(x^{|I_i|})$, which is exactly $(f_0)_\sigma(x \otimes \cdots \otimes x)$.

(4). As before, let $(e_1, \ldots, e_d)$ be the canonical basis of $\mathbb{C}^d$. We get a basis of $(\mathbb{C}^d)^{\otimes r}$ by taking the $e_{i_1} \otimes \cdots \otimes e_{i_r}$, for all $i_1, \ldots, i_r \in \{1, \ldots, d\}$. So let $i_1, \ldots, i_r \in \{1, \ldots, d\}$. As $r \geq d + 1$, there exists $s, t \in \{1, \ldots, r\}$ distinct such that $i_s = i_t$. Let τ be the transposition (st). Then we have $\mathfrak{S}_r = S \sqcup \tau S$, for S a set of representatives of the quotient $\{1, \tau\} \setminus \mathfrak{S}_r$. Moreover, for every $\sigma \in \mathfrak{S}_r$, $\tau\sigma(e_{i_1} \otimes \cdots \otimes e_{i_r}) = \sigma(e_{i_1} \otimes \cdots \otimes e_{i_r})$ and $\mathrm{sgn}(\tau\sigma) = -\mathrm{sgn}(\sigma)$. So

$$\sum_{\sigma \in \mathfrak{S}_r} \mathrm{sgn}(\sigma)\sigma(e_{i_1} \otimes \cdots \otimes e_{i_r}) = \sum_{\sigma \in S} \mathrm{sgn}(\sigma)\sigma(e_{i_1} \otimes \cdots \otimes e_{i_r}) + \sum_{\sigma \in \tau S} \mathrm{sgn}(\tau\sigma)\tau\sigma(e_{i_1} \otimes \cdots \otimes e_{i_r}) = 0.$$

(5). Let $r \geq 1$. We make $M_d(\mathbb{C})^r$ act on $(\mathbb{C}^d)^{\otimes r}$ in the following way : $(x_1, \ldots, x_r)(v_1 \otimes \cdots \otimes v_r) = (x_1 v_1) \otimes \cdots \otimes (x_r v_r)$. Then the map $M_d(\mathbb{C})^r \to \mathbb{C}$,

$$(x_1, \ldots, x_r) \longmapsto S_r(f_0)(x_1 \otimes \cdots \otimes x_r) - \sum_{\sigma \in \mathfrak{S}_r} \mathrm{sgn}(\sigma) \mathrm{Tr}((x_1, \ldots, x_r)\sigma, (\mathbb{C}^d)^{\otimes r})$$

is r-linear and symmetric (because both of its summands are; for the second summand, this is proved as in question (2) of teh previous part). By question (3), this map is zero an all the $x \otimes \cdots \otimes x$, $x \in M_d(\mathbb{C})$. By problem VII.7.2, it is therefore identically 0, i.e., for all $x_1, \ldots, x_r \in M_d(\mathbb{C})$,

$$S_r(f_0)(x_1 \otimes \cdots \otimes x_r) = \sum_{\sigma \in \mathfrak{S}_r} \mathrm{sgn}(\sigma) \mathrm{Tr}((x_1, \ldots, x_r)\sigma, (\mathbb{C}^d)^{\otimes r}).$$

By question (4), this show that $S_{d+1}(f_0) = 0$.

(6). We already accidentally proved this in the answer of question (2).

□

VII.7.3.3 Characteristic polynomial of a pseudo-character

In this question, we suppose that $d!$ is invertible in k.

(1). (Newton's identities) Show that there exists unique polynomials $a_0, \ldots, a_{d-1}$ in $\mathbb{Z}[\frac{1}{d!}][t_1, \ldots, t_d]$ such that, for every $\alpha_1, \ldots, \alpha_d \in \mathbb{C}$,

$$t^d + a_{d-1}(s_1, \ldots, s_d)t^{d-1} + \cdots + a_1(s_1, \ldots, s_d) + a_0(s_1, \ldots, s_d) = (t - \alpha_1) \ldots (t - \alpha_d),$$

where $s_r = \alpha_1^r + \cdots + \alpha_d^r$ for $1 \leq r \leq d$.

If $f : R \to k$ is a central function, we define the characteristic polynomial of f at $x \in R$ to be the polynomial

$$P_{x,f}(t) = t^d + a_{d-1}(s_1, \ldots, s_d)t^{d-1} + \cdots + a_1(s_1, \ldots, s_d)t + a_0(s_1, \ldots, s_d) \in k[t],$$

where $s_r = f(x^r)$ for $1 \leq r \leq d$.

(2). If $f = \mathrm{Tr} \circ u$ with $u : R \to M_d(k)$ a k-algebra map, show that $P_{x,f}(t)$ is the characteristic polynomial of $u(x)$ for every $x \in R$.

(3). Let $f : R \to k$ be a central function, and let $x \in R$. We set

$$Q_{x,f}(t) = \sum_{\sigma \in \mathfrak{S}_{d+1}} \mathrm{sgn}(\sigma) f(x^{|c_1|}) \ldots f(x^{|c_{l-1}|}) t^{|c_l|-1},$$

where $\sigma = c_1 \ldots c_l$ is the decomposition of σ into cycles with disjoint supports such that $d+1$ is in the support of c_l, and $|c_i|$ is the length of the cycle c_i (and $t^0 = 1$ by convention).

Show that

$$Q_{x,f}(t) = (-1)^d d! P_{x,f}(t).$$

Hint : It's possible to prove this by direct computation. You can also use the following trick to reduce to the case where $R = M_d(k)$ and $f = \mathrm{Tr}$. [40] First, notice that $(-1)^d d! P_{x,f}(t)$ and $Q_{x,f}(t)$ are the evaluations at $s_r = f(x^r)$ of polynomials P and Q in the indeterminates $t, s_1, \ldots, s_d$. Then show that it's enough to prove that the evaluations of P and Q at $s_r = \alpha_1^r + \cdots + \alpha_d^r$, $1 \leq r \leq d$, are equal for all $\alpha_1, \ldots, \alpha_d \in \mathbb{C}$.

(4). If $f : R \to k$ is a central function, show that

$$S_{d+1}(f)(x, \ldots, x, y) = (-1)^d d! f(P_{x,f}(x)y)$$

for all $x, y \in R$.

If $f : R \to k$ is a central function, its *kernel* is defined by

$$\mathrm{Ker}(f) = \{x \in R | \forall y \in R,\ f(xy) = 0\}.$$

We say that f is *faithful* if $\mathrm{Ker}(f) = 0$.

(5). Show that $\mathrm{Ker}(f)$ is a two-sided ideal of R for every central function $f : R \to k$.

(6). (Cayley-Hamilton theorem) If $f : R \to k$ is a faithful pseudo-character of degree d, show that $P_{x,f}(x) = 0$ for every $x \in R$.

Solution.

(1). We work in the polynomial ring $\mathbb{Z}[\alpha_1, \ldots, \alpha_d]$ (the α_i are indeterminates) and see $s_1, \ldots, s_d$ as elements of this ring. We also set $s_0 = \sum_{i=1}^d \alpha_i^0 = d$. For every $r \geq 0$, let

$$\sigma_r = \sum_{\substack{S \subset \{1,\ldots,d\} \\ |S|=r}} \prod_{i \in S} \alpha_i \in \mathbb{Z}[\alpha_1, \ldots, \alpha_d].$$

These are the elementary symmetric polynomials. Note that $\sigma_0 = 1$ and $\sigma_r = 0$ for $r > d$. We have

$$f(t) := \prod_{i=1}^d (t - \alpha_i) = \sum_{r=0}^d (-1)^r \sigma_r t^{d-r}$$

in $\mathbb{Z}[\alpha_1, \ldots, \alpha_d][t]$. This gives

$$f'(t) = \sum_{i=1}^d \frac{f(t)}{t - \alpha_i} = \sum_{r=0}^{d-1} (-1)^r (d-r) \sigma_r t^{d-r-1}.$$

In the ring of Laurent formal power series in $\frac{1}{t}$ with coefficients in $\mathbb{Z}[\alpha_1, \ldots, \alpha_d]$, we have

$$\frac{f'(t)}{f(t)} = \sum_{i=1}^d \frac{1}{t - \alpha_i} = \sum_{i=1}^d \sum_{r \geq 0} \frac{\alpha_i^r}{t^{r+1}} = \sum_{r \geq 0} \frac{s_r}{t^{r+1}},$$

[40] But don't forget to treat this case ! It is not totally trivial.

hence

$$f'(t) = f(t)\frac{f'(t)}{f(t)} = \sum_{i=0}^{d}\sum_{r\geq 0}(-1)^i\sigma_i s_r t^{d-i-r-1}.$$

By equating the coefficients of t^{d-k-1} in the two expressions for $f'(t)$, we get, for $0 \leq k \leq d-1$,

$$(-1)^k(d-k)\sigma_k = \sum_{r=0}^{k}(-1)^{k-r}\sigma_{k-r}s_r = (-1)^k d\sigma_k + \sum_{r=1}^{k}(-1)^{k-r}\sigma_{k-r}s_r,$$

hence

$$k\sigma_k = -\sum_{r=1}^{k}(-1)^{k-r}\sigma_{k-r}s_r.$$

Using this and the fact that $\sigma_1 = s_1$, an easy induction on k show that $\sigma_k \in \mathbb{Z}[\frac{1}{d!}][s_1, \ldots, s_d]$ for every $k \in \{1, \ldots, d-1\}$. Moreover, applying this to $d+1$ instead of d and then setting $\alpha_{d+1} = 0$, we also get $d\sigma_d = -\sum_{r=1}^{d}(-1)^{d-r}\sigma_{d-r}s_r$, hence $\sigma_d \in \mathbb{Z}[\frac{1}{d!}][s_1, \ldots, s_d]$. In particular, thanks to the formula $\prod_{i=1}^{d}(t-\alpha_i) = \sum_{r=0}^{d}(-1)^r\sigma_r t^{d-r}$, we get the existence of the polynomials $a_0, \ldots, a_{d-1}$.

Let's show the uniqueness of $a_0, \ldots, a_{d-1}$. So suppose that we have another family $b_0, \ldots, b_{d-1}$ of polynomials satisfying the same properties. Then, for all $\alpha_1, \ldots, \alpha_d \in \mathbb{C}$ and for $0 \leq r \leq d-1$,

$$a_r(s_1(\alpha_1, \ldots, \alpha_d), \ldots, s_d(\alpha_1, \ldots, \alpha_d)) = b_r(s_1(\alpha_1, \ldots, \alpha_d), \ldots, s_d(\alpha_1, \ldots, \alpha_d)).$$

As $\mathbb{C}$ is infinite, two polynomials in $\mathbb{Z}[X_1, \ldots, X_d]$ are equal if and only if they take the same value on every $(x_1, \ldots, x_d) \in \mathbb{C}^d$. So it suffices to prove that, for every $(x_1, \ldots, x_d) \in \mathbb{C}^d$, there exists $(\alpha_1, \ldots, \alpha_d) \in \mathbb{C}^d$ such that $x_r = \sum_{i=1}^{r}\alpha_i^r$ for $1 \leq r \leq d$. Define a family $y_1, \ldots, y_d \in \mathbb{C}$ inductively by $y_1 = x_1$ and

$$ky_k = -\sum_{r=1}^{k}(-1)^{k-r}y_{k-r}x_r$$

for $2 \leq k \leq d$. Let $\alpha_1, \ldots, \alpha_d \in \mathbb{C}$ be the roots of the polynomials $t^d + \sum_{i=1}^{d}(-1)^i y_i t^{d-i} \in \mathbb{C}[t]$. Then we have $y_r = \sigma_r(\alpha_1, \ldots, \alpha_d)$ for $1 \leq r \leq d$. Hence the relation above between the x_k and y_k gives that $x_r = s_r(\alpha_1, \ldots, \alpha_d)$ for $1 \leq r \leq d$, which is what we wanted.

(2). It suffices to prove the following statement : For every $A \in M_d(k)$, if $s_r = \mathrm{Tr}(A^r)$ for $1 \leq r \leq d$, then the characteristic polynomial of A is equal to $t^d + a_{d-1}(s_1, \ldots, s_d)t^{d-1} + \cdots + a_1(s_1, \ldots, s_d)t + a_0(s_1, \ldots, s_d)$. (Then we apply this to $A = u(x)$, for $x \in R$.)

Write $A = (a_{ij})$, and consider the ring $k' = \mathbb{Z}[X_{ij}]$ and $B = (X_{ij}) \in M_d(k')$. We have a map of rings $\varphi : k' \to k$ sending each X_{ij} to a_{ij}, and $\varphi(B) = A$. So $\varphi(\det(tI_d - B)) = \det(tI_d - A)$ and $\varphi(\mathrm{Tr}(B^r)) = \mathrm{Tr}(A^r)$ for every $r \geq 0$, and so it suffices to prove the statement in the case $k = k'$ and $A = B$. Also, choosing d^2 algebraically independent elements of $\mathbb{C}$ gives an injective map of rings $k' \to \mathbb{C}$, and so we may assume that $k' = \mathbb{C}$. Let $\alpha_1, \ldots, \alpha_d \in \mathbb{C}$ be the eignevalues of B, we have $\mathrm{Tr}(B^r) = s_r(\alpha_1, \ldots, \alpha_d)$ for $1 \leq r \leq d$ and

$$\det(tI_d - B) = (t - \alpha_1) \ldots (t - \alpha_d),$$

so the desired equality follows directly from the definition of the polynomials $a_0, \ldots, a_{d-1}$.

(3). Note that, if $x, y \in R$, we have

$$f(Q_{x,f}(x)y) = \sum_{\sigma \in \mathfrak{S}_{d+1}} \mathrm{sgn}(\sigma) f(x^{|c_1|}) \ldots f(x^{|c_{l-1}|}) f(x^{|c_l|-1}y),$$

with $\sigma = c_1 \ldots c_l$ as before. By definition of $S_{d+1}(f)$, this is equal to $S_{d+1}(f)(x, \ldots, x, y)$. We will use this identity later in the solution of (3).

Consider the following polynomials in $\mathbb{Z}[s_1, \ldots, s_d, t]$ (here $s_1, \ldots, s_d$ are seen as indeterminates) :

$$P = t^d + \sum_{i=0}^{d-1} a_i(s_1, \ldots, s_d) t^{i-1}$$

and

$$Q = \sum_{\sigma \in \mathfrak{S}_{d+1}} \mathrm{sgn}(\sigma) s_{|c_1|} \ldots s_{|c_{l-1}|} t^{|c_l|-1},$$

where $\sigma = c_1 \ldots c_l$ is the decomposition of σ into cycles with disjoint supports such that $d + 1$ is in the support of c_l. Let $x \in R$. Then $P_{x,f}(t)$ (resp. $Q_{x,f}(t)$) is obtained by evaluating P (resp. Q) at $s_r = f(x^r)$, $1 \leq r \leq d$. So, to prove the statement, we just need to show that $P = Q$ in $\mathbb{Z}[s_1, \ldots, s_d, t]$. As $\mathbb{C}$ is infinite, we just need to show that the evaluations of P and Q at every element $(s_1, \ldots, s_d)$ of $\mathbb{C}^d$ are equal. Let $s_1, \ldots, s_d \in \mathbb{C}$. We have seen in the solution of question (1) that there exist $\alpha_1, \ldots, \alpha_d \in \mathbb{C}$ such that $s_r = \alpha_1^r + \cdots + \alpha_d^r$ for $1 \leq r \leq d$. Let $A \in M_d(\mathbb{C})$ be a matrix with eigenvalues $\alpha_1, \ldots, \alpha_d$. Then $\mathrm{Tr}(A^r) = \alpha_1^r + \cdots + \alpha_d^r$ for every $r \geq 0$, so $P(s_1, \ldots, s_d, t) = P_{A,\mathrm{Tr}}(t)$ and $Q(s_1, \ldots, s_d, t) = P_{A,\mathrm{Tr}}(t)$. So to show that $P = Q$, it suffices to show the statement of question (3) in the case $R = M_d(\mathbb{C})$, $f = \mathrm{Tr}$.

Let's suppose that $R = M_d(\mathbb{C})$ and $f = \mathrm{Tr}$. As both sides of the equality to prove are continuous in $x \in M_d(\mathbb{C})$ and don't change if we replace x by gxg^{-1}, for $g \in \mathrm{GL}_n(\mathbb{C})$, we may assume that $x \in M_d(\mathbb{C})$ is a diagonalizable matrix with pairwise distinct eigenvalues. Remember that we saw at the beginning that, for every $y \in M_d(\mathbb{C})$,

$$\mathrm{Tr}(Q_{x,\mathrm{Tr}}(x)y) = S_{d+1}(\mathrm{Tr})(x, \ldots, x, y).$$

As Tr is a pseudo-character of degree $\leq d$ by the previous part, this is equal to 0 for every $y \in M_d(\mathbb{C})$, and so $Q_{x,\mathrm{Tr}}(x) = 0$. This means that the minimal polynomial of x divides $Q_{x,\mathrm{Tr}}$. As x has pairwise distinct eigenvalues, its minimal polynomial is its characteristic polynomial, i.e. $P_{x,\mathrm{Tr}}$ (by question (2)). Now note that $\deg(Q_{x,\mathrm{Tr}}) \leq d$ and that the degree d part of $Q_{x,\mathrm{Tr}}$ is given by

$$\sum_{\substack{\sigma \in \mathfrak{S}_{d+1} \\ \sigma \text{ is a } (d+1)-\text{cycle}}} \mathrm{sgn}(\sigma) t^d.$$

There are $d!$ $(d+1)$-cycles in $\mathfrak{S}_{d+1}$, and they all have signature $(-1)^d$, so the leading term of $Q_{x,\mathrm{Tr}}$ is $(-1)^d d! t^d$. As $P_{x,\mathrm{Tr}}$ is monic of degree d and divides $Q_{x,\mathrm{Tr}}$, we finally get $Q_{x,\mathrm{Tr}} = (-1)^d d! P_{x,\mathrm{Tr}}$.

(4). We have seen at the beginning of the solution of question (3) that, for all $x, y \in R$,

$$f(Q_{x,f}(x)y) = S_{d+1}(f)(x, \ldots, x, y).$$

But question (3) gives $Q_{x,f}(x) = (-1)^d d! P_{x,f}(x)$, so we get the desired equality immediately.

(5). It is clear on the definition that $\mathrm{Ker}(f)$ is a right ideal of R. Let's show that it is also a left ideal. This is also very easy. If $x \in \mathrm{Ker}(f)$ and $a \in R$, then, for every $y \in R$, $f((ax)y) = f(xya) = 0$ because f is a central function. So $ax \in \mathrm{Ker}(f)$.

(6). Let $x \in R$. By question (4), we have

$$0 = S_{d+1}(f)(x, \ldots, x, y) = (-1)^d d! f(P_{x,f}(x)y)$$

for every $y \in R$. As $d!$ is invertible in k, this implies that $f(P_{x,f}(x)y) = 0$ for every $y \in R$, i.e. that $P_{x,f}(x) \in \mathrm{Ker}(f)$. As f is faithful, this gives $P_{x,f}(x) = 0$.

□

VII.7.3.4 Pseudo-characters over an algebraically closed field

In this question, we suppose that k is an algebraically closed field where $d!$ is invertible and that the k-algebra R is finite-dimensional. Let $f : R \to k$ be a pseudo-character of degree d. The goal is to show that f is the trace of an actual representation. [41]

(1). Show that we may assume that $\mathrm{Ker}(f) = 0$.

From now on, we assume that f is faithful, i.e. that $\mathrm{Ker}(f) = 0$. Remember that $\mathrm{rad}(R)$ is the Jacobson radical of R.

[41] The conclusion is actually true without the hypothesis on R, but with a slightly more difficult proof. See corollary 4.4 of Rouquier's paper [24].

(2). Show that every element of $\mathrm{rad}(R)$ is nilpotent. *(Hint : use the Cayley-Hamilton theorem of the previous part.)*

(3). Show that $\mathrm{rad}(R) = 0$. *(Hint : Show that $f(x) = 0$ for every $x \in \mathrm{rad}(R)$.)*

(4). Conclude.

Solution.

(1). We know that $\mathrm{Ker}(f)$ is an ideal of R by question (5) of the previous part. Let $\pi : R \to R/\mathrm{Ker}(f)$ be the projection map. As $f_{|\mathrm{Ker}(f)} = 0$, we can write $f = \overline{f} \circ \pi$ with $\overline{f} : R/\mathrm{Ker}(f) \to k$, and it is very easy to check that $\overline{f}$ is a pseudo-character of degree d. Suppose that we have a k-algebra map $\overline{u} : R/\mathrm{Ker}(f) \to M_d(k)$ such that $\overline{f} = \mathrm{Tr} \circ \overline{u}$. Then $u = \overline{u} \circ \pi : R \to M_d(k)$ is a k-algebra map, and $f = \mathrm{Tr} \circ u$.

(2). Let $x \in \mathrm{rad}(R)$. As f is faithful, the Cayley-Hamilton theorem (question (6) of the previous part) gives $P_{x,f}(x) = 0$. The polynomial $P_{x,f}(t) \in k[t]$ is nonzero because its leading term is t^d, so we may write the equality $P_{x,f}(x) = 0$ as $x^r(c_0 + c_1x + \cdots + c_sx^s) = 0$, with $r \geq 0$ and $c_0, c_s \in k^\times$. As $x \in \mathrm{rad}(R)$, $c_0 + \cdots + c_sx^s$ is invertible (by proposition I.2.5 of chapter I), so we must have $r \geq 1$, and we get x^r, which show that x is nilpotent.

(3). First we note that, for every $x \in R$, $f(x^2) = 0$ implies that $f(x) = 0$. Indeed, we know that

$$0 = S_{d+1}(f)(x, \ldots, x) = \sum_{\sigma \in \mathfrak{S}_{d+1}} \mathrm{sgn}(\sigma) \prod_{i=1}^{l} f(x^{|c_i|}),$$

where $\sigma = c_1 \ldots c_l$ is the decomposition of σ as a product of cycles with disjoint supports. If $f(x^2) = 0$, then the only surviving term in the sum above is that for $\sigma = 1$, so we get $f(x)^{d+1} = 0$ in k. As k is a field, this gives $f(x) = 0$.

Now suppose that $x \in \mathrm{rad}(R)$. By question (2), x is nilpotent, so there exists $r \geq 0$ such that $x^{2^r} = 0$. In particular, we have $f(x^{2^r}) = 0$, and we have just seen that this implies that $f(x) = f(x^2) = \cdots = f(x^{2^{r-1}}) = f(x^{2^r}) = 0$.

(4). As $\mathrm{rad}(R)$ is an ideal of R (by corollary I.2.6 of chapter I), the result of question (3) implies that $\mathrm{rad}(R) \subset \mathrm{Ker}(f)$. But f is faithful, so $\mathrm{rad}(R) = 0$. As R is a finite-dimensional k-vector space, it's a left Artinian k-algebra, so, by theorem I.2.11 of chapter I, it is semisimple. As k is algebraically closed, we get by remark I.1.10.7 of chapter I an isomorphism $R \simeq M_{n_1}(k) \times \cdots \times M_{n_r}(k)$, with $n_1, \ldots, n_r \geq 1$.

For every $i \in \{1, \ldots, r\}$, we denote by e_i the unit element in $M_{n_i}(k)$ and by f_i the restriction of f to $M_{n_i}(k)$. Then f_i is a pseudo-character of degree $d_i \leq d$, so by question (5) of the first part, $f(e_i) = d_i$. By question (6) of the first part, $f_i = \frac{d_i}{n_i}\mathrm{Tr}$. Also, we have $1 = e_1 + \cdots + e_r$ in R, so $d = f(1) = d_1 + \cdots + d_r$ (where the first equality follows again from question (5) of the first part).

We want to show that all the quotient $\frac{d_i}{n_i}$ are nonnegative integers. Fix $i \in \{1, \ldots, r\}$, and

let e be the matrix $E_{11} \in M_{n_i}(k)$. We have $f(e) = \frac{d_i}{n_i}\mathrm{Tr}(e) = \frac{d_i}{n_i}$. Also, $e^2 = e$, so using the equality of question (3) of the first part and reasoning as in the proof if question (5) of that same part, we get that $f(e)(f(e)-1)\dots(f(e)-d) = S_{d+1}(f)(e\otimes\cdots\otimes e) = 0$. As k is a field, this shows that $\frac{d_i}{n_i} = f(e) \in \{0,\dots,d\}$.

Now let $u : R \to M_d(k)$ be the map sending $x = (x_1,\dots,x_r) \in M_{n_1}(k)\times\cdots\times M_{n_r}(k) = R$ to the $d\times d$ matrix with diagonal blocks $(x_1,\dots,x_1,\dots,x_r,\dots,x_r)$, where each x_i is repeated $\frac{d_i}{n_i}$ times. This u is obviously a morphism of k-alebras, and we have $f = \mathrm{Tr}\circ u$. Indeed, iff $x = (x_1,\dots,x_r)$ as before, then $x = x_1+\cdots+x_r$ in R, so

$$f(x) = \sum_{i=1}^{r} f(x_i) = \sum_{i=1}^{r} f_i(x_i) = \sum_{i=1}^{r} \tfrac{d_i}{n_i}\mathrm{Tr}(x_i) = \mathrm{Tr}(u(x)).$$

□

VII.7.3.5 Characters and representations

In this question, we suppose that k is an algebraically closed field and that $\dim_k(R) < +\infty$. (The most important example is when R is the group algebra of a finite group.) All the (left) R-modules are assumed to be finite-dimensional over k. If M is a R-module, its *character* $\chi_M : R \to k$ is by definition the composition of the structural map $R \to \mathrm{End}_k(M)$ and of the trace $\mathrm{End}_k(M) \to k$. By VII.7.3.2, this is a pseudo-character of degree $\leq \dim_k(M)$ if $M \neq 0$, and it is actually of degree $\dim_k(M)$ of $(\dim_k(M))!$ is invertible in k.

Let M be a R-module. Because we assumed that $\dim_k(M)$ is finite, M has a composition (= Jordan-Hölder) series $M = M_0 \supset M_1 \supset \cdots \supset M_r = 0$. (See sections I.1.5 of I.1.6 of chapter I.) The *semisimplification* of M is by definition the semisimple R-module

$$M^{ss} = \bigoplus_{i=1}^{r} M_{i-1}/M_i.$$

By the Jordan-Hölder theorem, it doesn't depend on the choice of the composition series.

(1). Let $V_1,\dots,V_r$ be the (isomorphism classes of) simple R-modules. Show that the functions $\chi_{V_1},\dots,\chi_{V_r}$ are linearly indepedent in the k-vector space of functions $R \to k$.

(2). Let M and M' be two R-modules such that $\dim_k(M) < char(k)$ and $\dim_k(M') < char(k)$. Show that $M^{ss} \simeq M'^{ss}$ if and only if $\chi_M = \chi_{M'}$.

Solution.

(1). As R is a finite-dimensional k-vector space, it is left Artinian, so $R/\operatorname{rad}(R)$ is semisimple by remark I.2.10 and theorem I.2.11 of chapter I. By the Artin-Wedderburn theorem (theorem I.1.10.5 of chapter I), $R/\operatorname{rad}(R) \simeq \prod_{i=1}^{r}\mathrm{End}_k(V_i)$. Note that all the

χ_{V_i} factor through $R/\operatorname{rad}(R)$, and that they correspond to the maps $\operatorname{Tr} \circ pr_i$, where $pr_i : \prod_{i=1}^r \operatorname{End}_k(V_i) \to \operatorname{End}_k(V_i)$ is the ith projection.

Suppose that we have $\sum_{i=1}^r a_i \chi_{V_i} = 0$, with $a_1, \ldots, a_r \in k$. Let $i \in \{1, \ldots, r\}$. We have a surjective map $R \to R/\operatorname{rad}(R) \simeq \prod_{j=1}^r \operatorname{End}_k(V_j) \simeq \prod_{j=1}^r M_{n_j}(k)$, where $n_j = \dim_k(V_j)$, and we choose an element $x_i \in R$ lifting $(0, \ldots, 0, e_i, 0, \ldots, 0)$, where e_i is an element of $M_{n_i}(k)$ of trace 1 (for example the elementary matrix E_{11}). Then $\chi_{V_i}(x_i) = 1$ and $\chi_{V_j}(x_i) = 0$ if $j \neq i$, so

$$0 = \left(\sum_{j=1}^r a_j \chi_{V_j} \right)(x_i) = a_i.$$

(2). If $M^{ss} \simeq M'^{ss}$, then clearly $\chi_M = \chi_{M'}$ (even without the condition on the dimensions of M and M'). Conversely, suppose that $\chi_M = \chi_{M'}$. Write $M^{ss} \simeq \bigoplus_{i=1}^r V_i^{\oplus n_i}$ and $M'^{ss} \simeq \bigoplus_{i=1}^r V_i^{\oplus m_i}$. Then we have $\sum_{i=1}^r n_i \chi_{V_i} = \sum_{i=1}^r m_i \chi_{V_i}$, so, by question (1), $n_i - m_i = 0$ in k for every $i \in \{1, \ldots, r\}$. As $\dim_k(M), \dim_k(M') < \operatorname{char}(k)$, we have $n_i, m_i < \operatorname{char}(k)$ for every i, and so the fact that $n_i = m_i$ in k implies that $n_i = m_i$ in $\mathbb{Z}$.

□

VII.7.3.6 Universal pseudo-character

In this question, we take G to be a group (not necessarily finite) and we fix a positive integer d.

(1). Show that there exists a unique pair (A^{univ}, f^{univ}), where A^{univ} is a commutative ring and $f^{univ} : A^{univ}[G] \to A^{univ}$ is a pseudo-character of degree d, satisfying the following condition : For every commutative ring A, for every pseudo-character $f : A[G] \to A$ of degree $\leq d$, there exists a unique morphism of rings $u : A^{univ} \to A$ such that $f_{|G} = u \circ f^{univ}_{|G}$.

Hint : The uniqueness should be easy. For the construction of A^{univ}, start with $\mathbb{Z}$, then for every list $(g_1, \ldots, g_r)$ of elements of G (r is variable) add an indeterminate (that is supposed to be $f^{univ}(g_1 \ldots g_r)$), then add some relations to make this work.

(2). (*) If G is a finitely generated group (i.e. generated by a finite subset), show that the ring $A^{univ}[1/d!]$ is a finitely generated $\mathbb{Z}[1/d!]$-algebra (i.e. a quotient of a polynomial algebra over $\mathbb{Z}[1/d!]$ with finitely many indeterminates).

(3). Suppose that G is a finite group. Show that, for every algebraically closed field k such that $d < \operatorname{char}(k)$, taking the character induces a bijection between the set of isomorphism classes of semisimple representations of G on k-vector spaces of dimension $\leq d$ and the set of ring morphisms $A^{univ} \to k$. [42] [43]

[42] This stays true for infinite groups.

[43] In the language of algebraic geometry, $\operatorname{Spec} A^{univ}$ is a $\mathbb{Z}$-scheme of finite type (if G is finitely generated) whose

Solution.

(1). Let's start with the uniqueness. Suppose that we have two pairs (A_1, f_1) and (A_2, f_2) satisfying the condition. Then the universal property of (A_2, f_2) (resp. (A_1, f_1)) gives a map $u : A_2 \to A_1$ (resp. $v : A_1 \to A_2$) such that $u \circ f_{2|G} = f_{1|G}$ (resp. $v \circ f_{1|G} = f_{2|G}$). Then we $(uv) \circ f_{1|G} = f_{1|G}$, so, by the uniqueness condition in the statement of the universal property for (A_1, f_1), $uv = \mathrm{id}_{A_1}$. Similary, $vu = \mathrm{id}_{A_2}$.

Let's show the existence. Consider the polynomial ring $B' = \mathbb{Z}[X_g, g \in G]$, and the ideal I of B' generated by the following elements :

- $X_{gh} - X_{hg}$, for every $g, h \in G$;
- for every $g_1, \ldots, g_{d+1} \in G$, the element $\sum_{\sigma \in \mathfrak{S}_{d+1}} \mathrm{sgn}(\sigma) \prod_{i=1}^{l} X_{g_{a_{i,1}} \cdots g_{a_{i,n_i}}}$, where $\sigma = c_1 \ldots c_m$ is the decomposition of σ into cycles with disjoint supports, and we have written $c_i = (a_{i,1} \ldots a_{i,n_i})$.

We take $B = B'/I$ and denote by $T : B[G] \to B$ the B-linear map sending every $g \in G$ to $X_g \mod I \in B$, and by $\pi : B' \to B$ the quotient map.

Let's show that the pair (B, T) has the universal property of the statement. Let A be a commutative ring and $f : A[G] \to A$ be a pseudo-character of degree $\leq d$. We have a map of rings $u' : B' \to A$ sending $X_g = T(g)$ to $f(g)$ for every $g \in G$, and $u'(I) = 0$ by the very definition of a pseudo-character of degree $\leq d$. So we get a map of rings $u : B \to A$ such that $u \circ T_{|G} = f_{|G}$. Suppose that $v : B \to A$ is another map of rings such that $v \circ T_{|G} = f_{|G}$, then $v \circ \pi = u'$ on every X_g, hence $v \circ \pi = u'$, hence $v = u$.

Finally, let's show that $T : B[G] \to B$ is a pseudo-character of degree d. The B-linear map $B[G] \otimes_B B[G] \to B$, $x \otimes y \longmapsto T(xy) - T(yx)$, is zero on all the elements of the form $g \otimes h$, with $g, h \in G$ (by definition of B), so it is zero everywhere because these elements generate $B[G] \otimes_B B[G]$. In other words, T is a central function. Similarly, the B-linear map $S_{d+1}T : B(G)^{\otimes d+1} \to B$ (the tensor product is again over B) is zero on all the elements of the form $g_1 \otimes \cdots \otimes g_{d+1}$ with $g_1, \ldots, g_{d+1} \in G$, by definition of B. As these elements generate $B[G]^{\otimes d+1}$, we see that $S_{d+1}T = 0$, and so T is a pseudo-character of degree $\leq d$. To see that the degree of T is exactly d, consider the $\mathbb{C}$-linear map $f : \mathbb{C}[G] \to \mathbb{C}$ sending every element of G to d. This is the character of the trivial representation of G on $\mathbb{C}^d$, and hence, by VII.7.3.2, it is a pseudo-character of degree d. So we get a map of rings $u : B \to \mathbb{C}$ such that $f_{|G} = u \circ T_{|G}$. As u is a map of rings, we have $S_d(f)(g_1 \otimes \cdots \otimes g_d) = u(S_d(T)(g_1 \otimes \cdots \otimes g_d))$ for all $g_1, \ldots, g_d \in G$. As the degree of f is d, $S_d(f)$ is not identically 0; as the elements $g_1 \otimes \cdots \otimes g_d$ generate $\mathbb{C}[G]^{\otimes d}$, $S_d(f)$ is nonzero on at least one of them, and then so is $S_d(T)$, which shows that $S_d(T)$ is not identically zero and hence that the degree of T is d.

Note that, by definition, B is generated as a $\mathbb{Z}$-algebra by the $T(g)$, $g \in G$.

---

k-points are naturally in bijection with isomorphism classes of dimension $\leq d$ semi-simple representations of G over k.

(2).

(3). Write $(A^{univ}, f^{univ}) = (A, T)$. Let $\mathrm{R}_{\leq d}$ be the set of isomorphism classes of semisimple representations of G on k-vector spaces of dimension $\leq d$, and $\mathrm{Hom}_{rings}(A, k)$ be te set of ring maps $A \to k$.

If $\rho : k[G] \to \mathrm{End}_k(V)$ is a representation of G on a k-vector space V such that $\dim_k V \leq d$, then $\mathrm{Tr} \circ \rho : k[G] \to k$ is a pseudo-character of degree $\dim_k V$ by VII.7.3.2, so by the universal property of (A, T) there exists a unique $u \in \mathrm{Hom}_{rings}(A, k)$ such that $u \circ T_{|G} = \mathrm{Tr} \circ \rho_{|G}$. As $\mathrm{Tr} \circ \rho$ only depends on the isomorphism class of the semisimplification of ρ, this gives a map $\varphi : \mathrm{R}_{\leq d} \to \mathrm{Hom}_{rings}(A, k)$.

Conversely, let $u \in \mathrm{Hom}_{rings}(A, k)$. Then $u \circ T_{|G} : G \to k$ extends to a k-linear map $f : k[G] \to k$, and we see as in the proof of (1) that this f is a pseudo-character of degree $d' \leq d$. By VII.7.3.4, there exists a representation $\rho : k[G] \to \mathrm{End}_k(V)$, with $\dim_k(V) = d'$, such that $f = \mathrm{Tr} \circ \rho$. Replacing ρ by its semisimplification (see VII.7.3.5), we may assume that ρ is semisimple. As $\dim_k(V) \leq d < \mathrm{char}(k)$, VII.7.3.5(2) says that ρ (with the property that $f = \mathrm{Tr} \circ \rho$) is unique up to semisimplification. This gives a map $\psi : \mathrm{Hom}_{rings}(A, k) \to \mathrm{R}_{\leq d}$.

The fact that $\psi \circ \varphi = \mathrm{id}_{\mathrm{R}_{\leq d}}$ follows from VII.7.3.5(2), and the fact that $\varphi \circ \psi = \mathrm{id}_{\mathrm{Hom}_{rings}(A,k)}$ follows from the remark, made in the proof of (1), that the elements $T(g)$, $g \in G$, generate A as a $\mathbb{Z}$-algebra.

□

VII.7.4 Schur-Weyl duality (chapters I, IV and VI)

In this problem, k is a field.

(1). If $A \subset A'$ are two k-algebras, the *centralizer of A in A'* is

$$Z_{A'}(A) = \{x \in A' | \forall y \in A, xy = yx\}.$$

Let V be a finite-dimensional k-vector space, let A be a subalgebra of $\mathrm{End}_k(V)$, and let $B = Z_{\mathrm{End}_k(V)}(A)$. Suppose that A is *semisimple*. Prove the following :

(a) B is semisimple.

(b) $A = Z_{\mathrm{End}_k(V)}(B)$.

(c) If k is algebraically closed, then, as a representation of $A \otimes_k B$, V is equal to $\bigoplus_{i \in I} V_i \otimes W_i$, where $(V_i)_{i \in I}$ (resp. $(W_i)_{i \in I}$)) is a complete set of representatives of isomorphism classes of irreducible representations of A (resp. B).

In particular, you get a bijection between the isomorphism classes of irreducible representations of A and B.

(2). We now assume that $\mathrm{char}(k)$ is prime to $n!$. We fix a nonzero finite-dimensional k-vector V and a positive integer n. We denote by T^nV and S^nV the nth tensor and symmetric powers of V (see problem VII.6.7). We make $\mathfrak{S}_n$ act on T^nV by permuting the factors (see problem VII.7.3.2(3)).

(a) Show that the quotient map $T^nV \to S^nV$ induces an isomorphism $(T^nV)^{\mathfrak{S}_n} \xrightarrow{\sim} S^nV$. We use this to identify S^nV to a subspace of T^nV in what follows.

(b) Show that S^nV is generated as a k-vector space by elements of the form v^n, $v \in V$. (Compare with problem VII.7.2.)

(c) Let A be a finite-dimensional associative k-algebra with unit. Then T^nA is also an associative k-algebra (we take $(a_1 \otimes \cdots \otimes a_n)(b_1 \otimes \cdots \otimes b_n) = (a_1b_1) \otimes \cdots \otimes (a_nb_n)$), and S^nV is a subalgebra (no need to prove this, it follows immediately from (a) anyway). For each $a \in A$, let

$$\Delta_n(a) = \frac{1}{n}(a \otimes 1 \otimes \cdots \otimes 1 + 1 \otimes a \otimes \cdots \otimes 1 + \cdots + 1 \otimes \cdots \otimes 1 \otimes a).$$

Show that, as a k-algebra, S^nA is generated by $\Delta_n(A)$.

(d) Now let $\mathfrak{g} = \mathfrak{gl}(V)$, and make this act on T^nV as in problem VII.6.7(5). Let $U\mathfrak{g}$ be the universal enveloping algebra of $\mathfrak{g}$. Show that the image of $U\mathfrak{g}$ in $\mathrm{End}_k(T^nV)$ is canonically identified with $S^n \mathrm{End}_k(V)$. (We see $S^n \mathrm{End}_k(V)$ as a subalgebra of $T^n \mathrm{End}_k(V)$ as in (c), and make $T^n \mathrm{End}_k(V)$ act on T^nV in the obvious way : $(\varphi_1 \otimes \cdots \otimes \varphi_n)(v_1 \otimes \cdots \otimes v_n) = \varphi_1(v_1) \otimes \cdots \otimes \varphi_n(v_n)$. This gives an injection $T^n \mathrm{End}_k(V) \hookrightarrow \mathrm{End}_k(T^nV)$, which is actually an isomorphism for dimension reasons.)

(e) Let A (resp. B) be the image of $k[\mathfrak{S}_n]$ (resp. $U\mathfrak{g}$) in $\mathrm{End}_k(T^nV)$. Show that A and B are semisimple, and that they are each other's centralizers in $\mathrm{End}_k(T^nV)$.

From now on, we take $k = \mathbb{C}$ and $V = \mathbb{C}^d$, so $\mathfrak{g} = \mathfrak{gl}_d(\mathbb{C})$.

(f) Let $Irr_{\mathfrak{g}}$ be the set of isomorphism classes of irreducible representations of $\mathfrak{g}$. We use the notation of chapter IV for partitions of n and irreducible representations of $\mathfrak{S}_n$. We also use the description of irreducible representations of $\mathfrak{g}$ in problem VII.6.19(4) (and the notation of this problem).

Show that there is a map

$$\begin{cases} \{\text{partitions of } n\} & \to & Irr_{\mathfrak{g}} \cup \{0\} \\ \lambda & \mapsto & W_\lambda \end{cases}$$

such that, as a representation of $\mathfrak{S}_n \times \mathfrak{g}$,

$$T^nV = \bigoplus_\lambda V_\lambda \otimes W_\lambda,$$

and that every W_λ comes by differentation (i.e. by the process of theorem VI.5.2 of chapter VI) from a continuous representation of $\mathrm{GL}_d(\mathbb{C})$, that we'll also denote by W_λ, and that is 0 or irreducible.

(g) Remember that $\dim V = d$. Let $\Delta \in k[x_1, \ldots, x_d]$ be the polynomial defined by

$$\Delta = \sum_{\sigma \in \mathfrak{S}_d} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} x_{\sigma(i)}^{d-i}.$$

Show that

$$\Delta = \prod_{1 \leq i < j \leq d} (x_i - x_j).$$

(h) Let $\lambda = (\lambda_1, \ldots, \lambda_r)$ be a partition of n with $r \leq d$. We define a polynomial $D_\lambda \in k[x_1, \ldots, x_d]$ by

$$D_\lambda = \sum_{\sigma \in \mathfrak{S}_d} \mathrm{sgn}(\sigma) \prod_{i=1}^{d} x_{\sigma(i)}^{\lambda_i + d - i},$$

where we take $\lambda_i = 0$ if $i > r$.

Let $S_\lambda = \frac{D_\lambda}{\Delta}$. Show that this is in $k[x_1, \ldots, x_d]$, i.e. a polynomial and not just a rational fraction. *(Hint : use the fact that D_λ is antisymmetric in $x_1, \ldots, x_d$.)*

Let $T_d \subset \mathrm{GL}_d(\mathbb{C})$ be the commutative subgroup of diagonal matrices. If $\rho : \mathrm{GL}_d(\mathbb{C}) \to \mathrm{GL}(W)$ is a continuous representation, remember that its character $\chi_W : \mathrm{GL}_d(\mathbb{C}) \to \mathbb{C}$ is defined by $\chi_W(g) = \mathrm{Tr}(\rho(g))$.

In the following questions, $g = \mathrm{diag}(a_1, \ldots, a_d)$ will be an element of T_d with diagonal entries $a_1, \ldots, a_d \in \mathbb{C}$ and σ will be an element of $\mathfrak{S}_n$. For every $l \in \mathbb{Z}_{\geq 1}$, we denote by i_l the number of cycles of length l in the decomposition of σ as a product of cycles with pairwise disjoint supports. Finally, we denote by Π the set of partitions $\lambda = (\lambda_1 \geq \cdots \geq \lambda_r)$ of n such that $r \leq d$.

Remember that we are using the notation of chapter IV.

(i) Show that

$$\mathrm{Tr}(g\sigma, T^n V) = \sum_{\lambda \in \mathscr{P}(n)} \chi_{V_\lambda}(\sigma) \chi_{W_\lambda}(g).$$

(j) Show that

$$\mathrm{Tr}(g\sigma, T^n V) = \prod_{l \geq 1} P_l(a_1, \ldots, a_d)^{i_l},$$

where, for every $l \geq 1$, $P_l(x_1, \ldots, x_d) = x_1^l + \cdots + x_d^l \in k[x_1, \ldots, x_d]$.

(k) Show that

$$\prod_{l\geq 1} P_l(x_1,\dots,x_d)^{i_l} = \sum_{\lambda\in\Pi} \chi_{V_\lambda}(\sigma) S_\lambda(x_1,\dots,x_d).$$

(Hint : Use the fact that the left-hand side is symmetric in $x_1,\dots,x_d$.)

(l) Show that $W_\lambda = 0$ if $\lambda \notin \Pi$, and that, for $\lambda \in \Pi$,

$$\chi_{W_\lambda}(g) = S_\lambda(a_1,\dots,a_d).$$

(m) Show that, in the notation of problem VII.6.19, if $\lambda \in \Pi$, then W_λ is the irreducible representation of $\mathfrak{g}$ corresponding to $(\lambda_1,\dots,\lambda_r,0,\dots,0) \in \Lambda^+_{\mathfrak{g}}$. (This justifies the notation W_λ a posteriori.)

Solution.

(1). Because A is semisimple, we can write $V = \bigoplus_{i\in I} V_i^{n_i}$ as a A-module, where $(V_i)_{i\in I}$ is the (finite) set of simple A-modules (up to isomorphism) and $n_i \geq 0$. Let $A = \prod_{i\in I} A_i$ be the corresponding decomposition of A into simple factors (see theorem I.1.10.5 of chapter I), that is, A acts on V_i through the projection $A \to A_i$. As the map $A \to \mathrm{End}_k(V)$ is injective, all the n_i are positive.

For every $i \in I$, let $\mathbb{D}_i = \mathrm{End}_A(V_i)$. This is a division algbera by Schur's lemma, and we have $B \simeq \prod_{i\in I} M_{n_i}(\mathbb{D}_i)$, so B is semisimple. Let $A' = Z_{\mathrm{End}_k(V)}(B)$; clearly, $A \subset A'$. By Schur's lemma again, $A' = \prod_{i\in I} E_i$, where E_i is the set of k-linear endomorphisms of V_i that commute with $\mathbb{D}_i = \mathrm{End}_A(V_i)$. By the double centralizer property (theorem I.1.9.2 of chapter I), the obvious map $A_i \to E_i$ is an isomorphism. So we get $A = A'$.

As B commutes with A, it preserves the isotypic components $V_i^{n_i}$ of V as a A-module. So to prove the rest of the last statement, we may assume that $A = A_i$ and $V = V_i^{n_i}$; then $B = M_{n_i}(\mathbb{D}_i)$. As B commutes with A, it acts on $\mathrm{Hom}_A(V_i, V)$ (through its action on V, so $(b\cdot f)(x) = bf(x)$ for all $b \in B$, $f \in \mathrm{Hom}_A(V_i,V)$ and $x \in V_i$), and, as a B-module, $\mathrm{Hom}_A(V_i,V)$ is isomorphic to $\mathbb{D}_i^{n_i}$, which is the unique simple B-module. So it suffices to show that the k-linear map $u : V_i \otimes_k \mathrm{Hom}_A(V_i,V) \to V$, $x \otimes f \longmapsto f(x)$, is an isomorphism of $A \otimes_k B$-module, where the $A \otimes_k B$ action on the left hand side is given by the A-action on the factor V_i and the B-action on the factor $\mathrm{Hom}_A(V_i,V)$. Let's check that u is $A \otimes_k B$-linear. Let $a \in A$, $b \in B$, $x \in V_i$ and $f \in \mathrm{Hom}_A(V_i,V)$. We have

$$u((a\otimes b)(x\otimes f)) = u((ax)\otimes(bf)) = (bf)(ax) = bf(ax) = baf(x) = abf(x),$$

where the last two inequalities come from the A-linearity of f and the fact that A and B commute. Also, u is surjective because $V \simeq V_i^{n_i}$ as a A-module. Finally, let's show that u is an isomorphism by computing dimensions. We use for the first time the hypothesis that k is algebraically closed. Because of this hypothesis, we have $A \simeq M_d(k)$, so $V_i \simeq V^d$, $\mathrm{End}_A(V_i) \simeq k$ and $\mathrm{Hom}_A(V_i,V) \simeq \mathrm{End}_A(V_i)^{n_i} \simeq k^{n_i}$. This gives

$$\dim_k V = n_i \dim_k V_i = (\dim_k \mathrm{Hom}_A(V_i,V))(\dim_k V_i),$$

hence the source and target of u have the same dimension.

(2). (a) Let $p : T^nV \to S^nV$ be the quotient map, and let $\pi : T^nV \to T^nV$ be the k-linear map defined by

$$v_1 \otimes \cdots \otimes v_n \longmapsto \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \sigma(v_1 \otimes \cdots \otimes v_n).$$

Then π is clearly a projection with image $(T^nV)^{\mathfrak{S}_n}$. Also, the map $V^n \to (T^n)^{\mathfrak{S}_n}V$, $(v_1, \ldots, v_n) \longmapsto \pi(v_1 \otimes \cdots \otimes v_n)$, is n-linear symmetric, so it induces a map $u : S^nV \to (T^nV)^{\mathfrak{S}_n}$, and it is very easy to see that u is an inverse of $p_{|(T^nV)^{\mathfrak{S}_n}}$.

(b) As in problem VII.7.2, we can do this two ways (at last if $k = \mathbb{C}$):

(i). The first solution of problem VII.7.2 gives that, for all $v_1, \ldots, v_n \in V$,

$$v_1 \ldots v_n = \frac{1}{n!} \sum_{\varnothing \neq S \subset \{1,\ldots,n\}} (-1)^{n-|S|} \left(\sum_{i \in S} v_i \right)^n$$

in S^nV, which immediately implies the result. This only uses the condition that $n! \in k^\times$.

(ii). If $k = \mathbb{C}$ (the case of interest later), then S^nV is an irreducible representation of $\mathfrak{sl}(V)$ by problem VII.6.4, hence of a group isomorphic to $\mathrm{SU}(\dim_{\mathbb{C}} V)$ by remark VI.8.3 of chapter VI, and the $\mathbb{C}$-subspace generated by the v^n, $v \in V$, is nonzero and stable by the action of this group, hence equal to S^nV.

(c) First, it is clear that $\Delta_n(A) \subset (T^nA)^{\mathfrak{S}_n}$, so the question makes sense. Let $a_1, \ldots, a_n \in A$. Then, in S^nV,

$$\Delta_n(a_1) \ldots \Delta_n(a_n) = a_1 \ldots a_n.$$

So the k-subalgebra of S^nA generated by $\Delta_n(A)$ contains all the elements of the form $a_1 \ldots a_n$, hence it is equal to S^nA.

(d) Let's make the injective map $S^n \operatorname{End}_k(V) \to \operatorname{End}_k(T^nV)$ explicit : If $u_1, \ldots, u_n \in \operatorname{End}_k(V)$ and $v_1, \ldots, v_n \in V$, then

$$(u_1 \ldots u_n)(v_1 \otimes \cdots \otimes v_n) = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} u_{\sigma(1)}(v_1) \otimes \cdots \otimes u_{\sigma(n)}(v_n).$$

Then the action of $\mathfrak{g}$ on T^nV is given by the composition of the map $\Delta_n : \mathfrak{g} \to S^n\mathfrak{g} = S^n \operatorname{End}_k(V)$ and of the map $S^n \operatorname{End}_k(V) \to \operatorname{End}_k(T^nV)$ that we just wrote. So the image of $U\mathfrak{g}$ is the k-subalgebra generated by the image of Δ_n, which by (c) is equal to $S^n \operatorname{End}_k(V)$.

(e) We already saw that $B = S^n \operatorname{End}_k(V)$. Hence B is a quotient of the semisimple k-algebra $T^n \operatorname{End}_k(V) \simeq \operatorname{End}_k(T^nV)$, and so it is semisimple.

On the other hand, $k[\mathfrak{S}_n]$ is semisimple because $n! = |\mathfrak{S}_n|$ is invertible in k (by Maschke's theorem, i.e. theorem I.3.2 of chapter I), so its quotient A is also semisimple. Also, it is clear on the formula for the action of $S^n \operatorname{End}_k(V)$ that this action commutes with the action of A, so $S^n \operatorname{End}_k(V) \subset Z_{k[\mathfrak{S}_n]}(\operatorname{End}_k(T^n V))$.

By (a), the centralizer of $k[\mathfrak{S}_n]$ in $\operatorname{End}_k(T^n V) \simeq T^n \operatorname{End}_k(V)$, i.e. $(T^n \operatorname{End}_k(V))^{\mathfrak{S}_n}$, is equal to $S^n \operatorname{End}_k(V) = B$. Finally, by question (1) (and the fact that A is semisimple), $A = Z_{\operatorname{End}_k(T^n V)}(B)$ (and we recover the fact that B is semisimple).

(f) We have seen (in theorem IV.3.3 of chapter IV) that we have a bijective map

$$\mathscr{P}(n) := \{\text{partitions of } n\} \to S_{\mathbb{C}}(\mathfrak{S}_n), \quad \lambda \longmapsto V_\lambda.$$

Also, by question (1), we have an isomorphism of $A \otimes_{\mathbb{C}} B$-modules

$$T^n V \simeq \bigoplus_{\lambda \in \mathscr{P}(n)} V_\lambda \otimes W_\lambda,$$

where, for every $\lambda \in \mathscr{P}(n)$, W_λ is either 0 (if the irreducible representation V_λ of $\mathfrak{S}_n$ doesn't appear in $T^n V$) or an irreducible representation of B, hence of $U\mathfrak{g}$, hence also of $\mathfrak{g}$. This gives the desired map. Note also that, if $\lambda, \mu \in \mathscr{P}(n)$ and such that $\lambda \neq \mu$ and $W_\lambda, W_\mu \neq 0$, then $W_\lambda \not\simeq W_\mu$ (again by (1)).

Finally, the representation of $\mathfrak{g}$ on $T^n V$ clearly lifts to a representation of $\mathrm{GL}_d(\mathbb{C})$ (given by the formula $g(v_1 \otimes \cdots \otimes v_n) = (gv_1) \otimes \cdots \otimes (gv_n)$, see remark VI.5.6 of chapter VI), and by remark VI.8.3 of chapter VI so do all its $\mathfrak{g}$-subrepresentation, hence so do the W_λ. By the same remark, each nonzero W_λ is irreductible as a representation of $\mathrm{GL}_d(\mathbb{C})$.

(g) It is clear on the definition of Δ that it is the determinant of the matrix $(x_i^{d-j})_{1 \leq i,j \leq d} \in M_d(k[x_1, \ldots, x_d])$. This is a Vandermonde matrix with the order of its columns reversed, so it determined is $(-1)^{d(d-1)/2} \prod_{1 \leq i < j \leq d}(x_j - x_i)$, which is exactly the desired formula for Δ.

(h) Make $\mathfrak{S}_d$ act on $k[x_1, \ldots, x_d]$ by $\sigma(f(x_1, \ldots, x_d)) = f(x_{\sigma(1)}, \ldots, x_{\sigma(d)})$, for every $\sigma \in \mathfrak{S}_d$ and $f \in k[x_1, \ldots, x_d]$. Then, for every $\sigma \in \mathfrak{S}_d$, $\sigma(D_\lambda) = \operatorname{sgn}(\sigma) D_\lambda$. (This is clear from the definition of D_λ.) In particular, if $i, j \in \{1, \ldots, j\}$ and $i \neq j$, then $\sigma(D_\lambda) = -D\lambda$ if $\sigma = (ij)$, hence $D_{\lambda | x_i = x_j} = 0$, so $x_i - x_j$ divides D_λ in $k[x_1, \ldots, x_d]$. As the ring $k[x_1, \ldots, x_d]$ is a unique factorization domain, we deduce that $\prod_{1 \leq i < j \leq d}(x_i - x_j)$ divides D_λ in $k[x_1, \ldots, x_d]$. But by (g), $\Delta = \prod_{1 \leq i < j \leq d}(x_i - x_j)$.

(i) This just follows from (f).

(j) The calculation exactly the same as in question (3) of VII.7.3.2 :

Let $(e_1, \ldots, e_d)$ be the canonical basis of $V = k^d$. Then a basis of $T^n V$ is given by $(e_{i_1} \otimes \cdots \otimes e_{i_n})_{1 \leq i_1, \ldots, i_n \leq d}$. Suppose that σ is a n-cycle. Then

$$(g\sigma)(e_{i_1} \otimes \cdots \otimes e_{i_n}) = (a_{\sigma^{-1}(i_1)} \ldots a_{\sigma^{-1}(i_n)})(e_{\sigma^{-1}(i_1)} \otimes \cdots \otimes e_{\sigma^{-1}(i_n)}),$$

and this is propotional to $e_{i_1} \otimes \cdots \otimes e_{i_n}$ if and only if $i_1 = \cdots = i_n$. So

$$\mathrm{Tr}(g\sigma, T^n V) = \sum_{i=1}^{d} a_i^n = P_n(a_1, \ldots, a_d).$$

Now if σ is any element of $\mathfrak{S}_n$, let $\sigma = c_1 \ldots c_\ell$ be its decomposition into cycles with disjoint supports $I_1, \ldots, I_\ell \subset \{1, \ldots, n\}$. Then we have (by the formula for the trace of a tensor product of maps, see the proof of proposition II.1.1.3 in chapter II)

$$\mathrm{Tr}(g\sigma, T^n V) = \prod_{i=1}^{\ell} \mathrm{Tr}(gc_i, V^{\otimes I_i}).$$

By the previous calculation, this is equal to $\prod_{i=1}^{\ell} P_{|I_i|}(a_1, \ldots, a_d)$, which gives the desired result.

(k) By definition of the S_λ, the formula we're trying to prove is equivalent to

$$\Delta \prod_{l \geq 1} P_l(x_1, \ldots, x_d)^{i_l} = \sum_{\lambda \in \Pi} \chi_{V_\lambda}(\sigma) D_\lambda(x_1, \ldots, x_d).$$

Note that, for every $tau \in \mathfrak{S}_d$, $\tau(D_\lambda) = \mathrm{sgn}(\tau) D_\lambda$, $\tau(\Delta) = \mathrm{sgn}(\tau)\Delta$ and $\tau(P_l) = P_l$. So both sides of the equality above are antisymmetric in the x_i, and so we only need to show that the coefficients of all the monomials of the form $x_1^{n_1} \ldots x_d^{n_d}$ with $n_1 > \cdots > n_d$ coincide. Note that $\Delta \prod_{l \geq 1} P_l^{i_l}$ is homogeneous of degree $d(d-1)/2 + \sum_{l \geq 1} l i_l = n + d(d-1)/2$, and that each D_λ is homogeneous of degree $\sum_{i=1}^{d} (d - i + \lambda_i) = n + d(d-1)/2$. So both sides are homogeneous of the same degree. Let $n_1 > \cdots > n_d \geq 0$ be integers such that $n_1 + \cdots + n_d = n + d(d-1)/2$, and write $n_i = d - i + \lambda_i$. Then we have $\lambda_1 \geq \cdots \leq \lambda_d = n_d \geq 0$, and $\lambda_1 + \cdots + \lambda_d = n$, so $\lambda := (\lambda_1, \ldots, \lambda_d) \in \Pi$. Also, by theorem IV.4.3 of chapter IV, the coefficient $\chi_{V_}lambda(\sigma)$ of $x_1^{n_1} \ldots x_d^{n_d}$ on the right hand side is equal to the coefficient of $x_1^{n_1} \ldots x_d^{n_d}$ in $\Delta \prod_{l \geq 1} P_l^{i_l}$, which is the left hand side. This proves the equality.

(l) This follows from (i), (j), (k) and the fact that the fonctions $\chi_{V_\lambda} : \mathfrak{S}_n \to \mathbb{C}$ form a linearly independent family (by corollary II.1.2.8 of chapter II).

(m) We have seen that W_λ is an irreducible representation of $\mathrm{GL}_d(\mathbb{C})$ and $\mathfrak{g}$, and that its character is given on $T_d \subset \mathrm{GL}_d(\mathbb{C})$ by the formula $(a_1, \ldots, a_d) \longmapsto S_\lambda(a_1, \ldots, a_d)$. By the Weyl character formula (theorem VI.13.2 of chapter VI), this is equal to the character of the representation W with highest weight λ of $\mathrm{GL}_d(\mathbb{C})$ on T_d. (See (6) of problem VII.6.19.) So $\chi_{W_\lambda} = \chi_W$ on T_d. As χ_W and χ_{W_λ} are both continuous and invariant by conjugation, and as the set of diagonalizable matrices in dense in $\mathrm{GL}_d(\mathbb{C})$, $\chi_W = \chi_{W_\lambda}$. Of course that should allow us to directly conclude that $W \simeq W_\lambda$ as representations of $\mathrm{GL}_d(\mathbb{C})$ (and hence of $\mathfrak{g}$), but technically we haven't proved this so we still need to do some work. What we can immediately conclude, however, is that :

- the characters of the representations of $\mathfrak{sl}_d(\mathbb{C})$ on W and W_λ are equal, so $W \simeq W_\lambda$ as representations of $\mathfrak{sl}_d(\mathbb{C})$, and in particular $\dim_{\mathbb{C}} W = \dim_{\mathbb{C}} W_\lambda$;
- the center $\mathbb{C}^\times$ of $\mathrm{GL}_d(\mathbb{C})$, which acts by homotheties, i.e. via continuous group morphisms $\psi_1, \psi_2 : \mathbb{C}^\times \to \mathbb{C}^\times$, on W and W_λ by Schur's lemma, acts via the same morphism on these two representations, i.e. $\psi_1 = \psi_2$ (indeed, for every $a \in \mathbb{C}^\times$, $\chi_W(a) = (\dim_{\mathbb{C}} W)\psi_1(a) = \chi_{W_\lambda}(a) = (\dim_{\mathbb{C}} W_\lambda)\psi_2(a)$);
- as a consequence of the second point, W and W_λ are isomorphic as representations of $\mathbb{C}I_d \subset \mathfrak{g}$.

Putting the first and third points together, we see that $W \simeq W_\lambda$ as representations of $\mathfrak{g}$ (and hence also as representations of $\mathrm{GL}_d(\mathbb{C})$).

□

Index

R-linear map, 12
R-module morphism, 12
p-elementary group, 69
p-regular, 72
p-singular element, 92
p-unipotent, 72
p-unramified ring, 241

action of a group on a module, 36
adjoint representation of a Lie algebra, 133
algebraic differential operator, 275
algebraic Peter-Weyl theorem, 285
annihilator of a R-module, 33
annihilator of an element, 33
Artin's theorem, 68
Artin-Wedderburn theorem, 31
Artinian R-module, 22
augmentation ideal, 36
augmentation map, 36

Baker-Campbell-Hausdorff formula, 148
Brauer group of a field, 190
Brauer's theorem, 69, 74, 89
Bruhat order, 160, 167

Casimir element, 165
Cauchy determinant, 105
Cauchy matrix, 105
central function, 55
central simple k-algebra, 188
centralizer, 73
character of a group (1-dimensional representation), 36
character of a Lie algebra representation, 168
character of a representation, 55, 115
character table of a group, 203
Clebsch-Gordon decomposition, 273
coinduction, 46
coinvariants, 47
column subgroup, 96
commutative Lie algebra, 132
commutator bracket, 132
compact operator, 127
complete discrete valuation ring, 237
completely reducible module, 15
completely reducible representation, 36
composition series, 20
continuous representation ring, 158
convolution product, 118
cuspidal representation of $\mathrm{GL}_2(\mathbb{F}_q)$, 222
cyclic module, 16

derivation, 185, 263
derivation of a Lie algebra, 133
differential polynomial ring, 185
direct sum of modules, 12
discrete valuation field, 236
discrete valuation ring, 235
division ring, 13
double centralizer property, 29
dual numbers, 276

eigenvalue, 128
equivalent Jordan-Hölder (or composition) series, 21
equivariant map, 36
exact sequence of modules, 12
exterior algebra of a module, 155

exterior powers of a module, 155

faithful Lie algebra representation, 132
faithful representation, 36
finite length R-module, 21
finitely generated module, 16
Fitting's lemma, 80
free algebra, 11
free module, 18
fundamental weight, 160

Gelfand-Graev representation of $\mathrm{GL}_2(\mathbb{F}_q)$, 218
group algebra, 11

Haar measure, 109
Hecke algebra, 213
Hensel's lemma, 238
highest weight, 160, 164
highest weight representation, 164
highest weight vector, 164
Hilbert space, 112

ideal, 13
ideal of a Lie algebra, 131
idempotent, 83
indecomposable module, 78
induction, 45, 197
injective module, 19
inner derivation, 185
invariant differential operators, 276
invariants, 48, 56
inverse, 13
invertible, 13
irreducible Lie algebra representation, 132
irreducible module, 15
irreducible representation, 36
isotypic components, 24
Iwasawa decomposition, 258

Jacobson radical, 33
Jacobson semisimple, 35
Jacobson-Morozov theorem, 273
Jordan-Hölder constituents, 22
Jordan-Hölder series, 20

Kostka numbers, 100
Krull-Schmidt-Remark theorem, 80

left Artinian ring, 22
left inverse, 13
left invertible, 13
left Noetherian ring, 22
left regular R-module, 12
left regular representation, 122
length of a R-module, 21
length of a Jordan-Hölder (or composition) series, 20
Lie algebra, 131
Lie algebra of a closed subgroup of $\mathrm{GL}_n(\mathbb{C})$, 140
Lie algebra of a linear algebraic group, 264
Lie bracket, 131
Lie subalgebra, 131
lifting of idempotents, 83
linear algberaic group, 263
local left Artinian ring, 78

Mackey's formula, 65
Mackey's irreducibility criterion, 67
Maschke's theorem, 36
matrix exponential, 137
matrix logarithm, 139
modular function, 255
module, 11
monomial representation, 74
morphism of algberaic groups, 282
morphism of Lie algebras, 131
multiplicity of a simple module, 24
multiplicity of a weight, 160, 162
multiplicity-free, 215

nil ideal, 83
nilpotent ideal, 35
Noetherian R-module, 22
noncommutative polynomial ring, 11

opposite ring, 29

order of a differential operator, 276
orthogonal group, 257
orthogonal idempotents, 19, 83
orthogonality of characters, 59

partition of an integer, 95
Peter-Weyl theorem, 124
Poincaré-Birkhoff-Witt, 275
Poincaré-Birkhoff-Witt theorem, 164
polar decomposition, 258
polarization, 287
polynomial function on a vector space, 269
positive roots, 163
principal series of $\mathrm{GL}_2(\mathbb{F}_q)$, 222
profinite group, 260
projection formula, 53
projective envelope (or cover), 83
projective module, 17
pseudo-character, 288

quaternions, 181
quotient module, 12

realizable over k (for a representation), 210
reduced trace, 190
regular functions on an algebraic group, 283
regular representation, 36
representation of a group on a module, 36
representation of a Lie algebra, 132
representation ring, 43
restriction, 45
right inverse, 13
right invertible, 13
right regular R-module, 12
right regular representation, 122
ring, 11
root system, 160
row subgroup, 96

Schur index, 211
Schur orthogonality, 59, 116
Schur's lemma, 20, 57
Schur-Weyl duality, 303
semisimple Lie algebra representation, 133
semisimple module, 15
semisimple representation, 36
semisimple ring, 17
semisimplification, 300
sign representation, 39
simple module, 15
simple ring, 26
simple roots, 163
Specht modules, 100
special orthogonal group, 257
special unitary group, 142, 257
spectrum of an operator, 128
standard representation, 157, 269
Steinberg representation of $\mathrm{GL}_2(\mathbb{F}_q)$, 222
submodule, 12
subrepresentation, 36
subrepresentation of a Lie algebra representation, 132
sum of modules, 12
symmetric algebra, 267
symmetric powers, 267
symplectic group, 257

tensor algebra of a module, 134
tensor powers of a module, 134
tensor product of two modules, 173
topological group, 109
trivial representation, 39
trivial representation of a Lie algebra, 133

unimodular group, 255
unitary group, 142, 257
unitary representation, 113
universal enveloping algebra, 135
Urysohn's lemma, 253

Verma module, 166

weight of a representation, 162
weight space, 162
weights of a representation, 160
Weil representation, 227
Weyl denominator, 163
Witt polynomials, 241

Witt vectors, 242

Young diagram attached to a partition, 96
Young projector, 97
Young tableau corresponding to a partition, 96

Zorn's lemma, 14

Bibliography

[1] I. D. Ado. The representation of Lie algebras by matrices. *Uspehi Matem. Nauk (N.S.)*, 2(6(22)):159–173, 1947.

[2] Joël Bellaïche. Ribet's lemma, generalizations, and pseudocharacters. `http://people.brandeis.edu/~jbellaic/RibetHawaii3.pdf`. Accessed: 2017-09-26.

[3] I. N. Bernˇ steĭn and A. V. Zelevinskiĭ. Representations of the group $GL(n, F)$, where F is a local non-Archimedean field. *Uspehi Mat. Nauk*, 31(3(189)):5–70, 1976.

[4] Armand Borel. *Linear algebraic groups*, volume 126 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[5] Kenneth S. Brown. *Cohomology of groups*, volume 87 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1994. Corrected reprint of the 1982 original.

[6] Daniel Bump. *Automorphic forms and representations*, volume 55 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1997.

[7] Gaëtan Chenevier. The p-adic analytic space of pseudocharacters of a profinite group and pseudorepresentations over arbitrary rings. In *Automorphic forms and Galois representations. Vol. 1*, volume 414 of *London Math. Soc. Lecture Note Ser.*, pages 221–285. Cambridge Univ. Press, Cambridge, 2014.

[8] Michel Demazure and Pierre Gabriel. *Groupes algébriques. Tome I: Géométrie algébrique, généralités, groupes commutatifs*. Masson & Cie, Éditeur, Paris; North-Holland Publishing Co., Amsterdam, 1970. Avec un appendice ıt Corps de classes local par Michiel Hazewinkel.

[9] Vladimir Dotsenko. Pseudo-characters. `http://www.maths.tcd.ie/~vdots/research/files/Pseudo-characters.pdf`. Accessed: 2017-10-03.

[10] M. Eichler. A new proof of the Baker-Campbell-Hausdorff formula. *J. Math. Soc. Japan*, 20:23–25, 1968.

[11] Per Enflo. A counterexample to the approximation problem in Banach spaces. *Acta Math.*, 130:309–317, 1973.

[12] Pavel Etingof, Oleg Golberg, Sebastian Hensel, Tiankai Liu, Alex Schwendner, Dmitry Vaintrob, and Elena Yudovina. Introduction to representation theory, 2009.

[13] William Fulton and Joe Harris. *Representation theory*, volume 129 of *Graduate Texts in*

Mathematics. Springer-Verlag, New York, 1991. A first course, Readings in Mathematics.

[14] James E. Humphreys. *Linear algebraic groups*. Springer-Verlag, New York-Heidelberg, 1975. Graduate Texts in Mathematics, No. 21.

[15] James E. Humphreys. *Introduction to Lie algebras and representation theory*, volume 9 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1978. Second printing, revised.

[16] Thomas Jech. *Set theory*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2003. The third millennium edition, revised and expanded.

[17] Irving Kaplansky. Projective modules. *Ann. of Math (2)*, 68:372–377, 1958.

[18] Anthony W. Knapp. *Lie groups beyond an introduction*, volume 140 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, second edition, 2002.

[19] Steven G. Krantz and Harold R. Parks. *A primer of real analytic functions*. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser Boston, Inc., Boston, MA, second edition, 2002.

[20] T. Y. Lam. *A first course in noncommutative rings*, volume 131 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 2001.

[21] Lynn H. Loomis. *An introduction to abstract harmonic analysis*. D. Van Nostrand Company, Inc., Toronto-New York-London, 1953.

[22] Louise Nyssen. Pseudo-représentations. *Math. Ann.*, 306(2):257–283, 1996.

[23] Daniel Perrin. *Cours d'algèbre*, volume 18 of *Collection de l'École Normale Supérieure de Jeunes Filles [Collection of the École Normale Supérieure de Jeunes Filles]*. École Normale Supérieure de Jeunes Filles, Paris, 1982. Edited with the collaboration of Marc Cabanes and Martine Duchene.

[24] Raphaël Rouquier. Caractérisation des caractères et pseudo-caractères. *J. Algebra*, 180(2):571–586, 1996.

[25] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.

[26] Walter Rudin. *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, second edition, 1991.

[27] Mark R. Sepanski. *Compact Lie groups*, volume 235 of *Graduate Texts in Mathematics*. Springer, New York, 2007.

[28] Jean-Pierre Serre. *Corps locaux*. Hermann, Paris, 1968. Deuxième édition, Publications de l'Université de Nancago, No. VIII.

[29] Jean-Pierre Serre. *Linear representations of finite groups*. Springer-Verlag, New York-Heidelberg, 1977. Translated from the second French edition by Leonard L. Scott, Graduate

Texts in Mathematics, Vol. 42.

[30] Jean-Pierre Serre. *Complex semisimple Lie algebras*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2001. Translated from the French by G. A. Jones, Reprint of the 1987 edition.

[31] Jean-Pierre Serre. *Lie algebras and Lie groups*, volume 1500 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2006. 1964 lectures given at Harvard University, Corrected fifth printing of the second (1992) edition.

[32] T. A. Springer. *Linear algebraic groups*. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, second edition, 2009.

[33] Terry Tao. Haar measure and the peter-weyl theorem. https://terrytao.wordpress.com/2011/09/27/254a-notes-3-haar-measure-and-the-peter-weyl-theorem/. Accessed: 2017-10-03.

[34] Terry Tao. The peter-weyl theorem, and non-abelian fourier analysis on compact groups. https://terrytao.wordpress.com/2011/01/23/the-peter-weyl-theorem-and-non-abelian-fourier-analysis-on-compact-gr Accessed: 2017-10-03.

[35] Richard Taylor. Galois representations associated to Siegel modular forms of low weight. *Duke Math. J.*, 63(2):281–332, 1991.

[36] A. Wiles. On ordinary λ-adic representations associated to modular forms. *Invent. Math.*, 94(3):529–573, 1988.