Representation Theory

v. 1 Representations of finite and compact groups.
Representations of simple Lie algebras

Dimitry Leites (Ed.)

Abdus Salam School
of Mathematical Sciences
Lahore, Pakistan

Summary. This volume consists of:

(1) lectures on representation theory of finite and compact groups for beginners (by A. Kirillov), and Lie theory on relation between Lie groups and Lie algebras (after É. Vinberg);

(2) introduction into representation theory of Lie algebras (by J. Bernstein); these lectures contain previously unpublished new proof of the Chevalley theorem and an approach to the classification of irreducible finite dimensional representations of simple finite dimensional Lie algebras by means of certain infinite dimensional modules.

(3) The lecture course is illustrated by a translation of a paper by V. Arnold in which the master shows how to apply the representation theory of one of the simplest of finite groups — the group of rotations of the cube — to some problems of real life. This part also contains open problems for a computer-aided study.

ISBN 978-969-9236-01-3

Contents

Editor's preface

What representation theory is. Representation theory of groups and Lie algebras and Lie rings is the branch of science that studies symmetries. Initially designed to study symmetries of (solutions of) algebraic equations it soon found its application in geology, namely, in crystallography (certain groups are now said to be *crystallographic*). Here finite groups were main characters. A bit later representation theory evolved to describe symmetries of (solutions of) differential equations wherefrom it came close to all branches of natural sciences and — via economics — to social sciences: Any equation describing reality (be it an elementary particle or stock market) must be invariant with respect to admissible (in the frameworks of the given model) changes of coordinates (and, perhaps, some other — hidden — parameters). The study of all such symmetries is the object of representation theory whose main characters are continuous (in particular, Lie [1)]) groups and their infinitesimal approximations — the Lie algebras.

The discovery of quarks (the most elementary of known particles — speaking about theoretical results) and construction of tomograph (what can be more down-to-earth and practical?!) are among well-known (and most spectacular) applications of representation theory. Rough classification of chemical elements can be described in terms of irreducible representations of one of the simplest finite groups (that of rotations of the cube); this is another of spectacular applications.

About this course. In 1997, at the Department of Mathematics of Stockholm University, I got a possibility to digress from the usual (somewhat boring) task of reading the Calculus course to students mainly not interested in math at all in order to give an introductory course on representation theory to interested listeners. I used the occasion to translate from the Russian the notes of A. A. Kirillov's brilliant lectures on the topic [Ki]. Kirillov [2)] lectured in Moscow to a very wide audience: "From first year students and occasional high school students from mathematical schools to gifted professors", in Kirillov's terms. The advantage of these lectures as compared with other courses is their simplicity and clarity. Besides, usual text books, except perhaps, Serre's one [Se] (see also [Si, V]), are far too thick for a half year or even a year course.

Serre and Manin taught us the best way to learn a new material: give a lecture course on the topic: "The first attempt might be not very successful, but after the second one the lecturer is usually able to learn the subject". Keeping this advice in mind, my *arrier pencée* was to try to learn in the process how to describe irreducible representations of simple finite supergroups, or, at least, get close to it. At least, figure out the definitions. Well, at least, in *some* cases. This was, and still is, an *Open Problem*. A more realistic goal (also not

---

[1] http://en.wikipedia.org/wiki/Sophus_Lie. *All unmarked footnotes in both volumes are due to me.*

[2] http://en.wikipedia.org/wiki/Alexandre_Kirillov

even approached) was to try to learn in the process about some interesting applications of finite groups to coding theory or cryptography. A question: *Is it possible to use simple finite supergroups in a similar way*?

Despite possible applications (say, to super Riemann surfaces, hence to high energy physics), nobody had described irreducible representations of simple finite supergroups; *consider at least one example*.

The transcripts of A.A. Kirillov's lectures I've translated did not cover the whole course: A part was not documented. So I inserted into the lectures some excerpts (mainly omitted here) from the monographs [OV], [V] and [St], [Go, G] to bridge the group theory with the Lie algebra theory.

I also could not resist the temptation to show the students how representations naturally arise in so-called "practical" questions, so I chose a short paper by V.I. Arnold [3)] (Crafoord prize [4)] winner, Wolf prize [5)] winner, Shaw prize [6)] winner, winner of numerous other prizes and member of most Academies) [Ar0], a paper *with several open problems*, a paper in which several disciplines embrace each other in what the French call "entrefécundation" (=mutual fertilization) to the admiration of the reader.

Kirillov concluded his course with applications. As such, he chose elements of Hamiltonian mechanics; for more detail, see [Ar]. Observe that it was Kirillov who proved that any Hamiltonian dynamical system can be realized on one of the orbits of the coadjoint representation of an appropriate Lie algebra, so any scholar studying dynamical systems has to learn at least some basics of representation theory. Let me point out here to Dynkin's short appendix in [D2] as probably one of the best summaries of Lie algebra theory, see also [D1]. Though written in slightly obsolete terms, this is still a yet unsurpassed masterpiece.

For the modern presentation of Lie algebra theory I've used (with minor editing) transcript of extremely transparent lectures by J. Bernstein [7)] in a summer school on representation theory: it squeezes all basics in just several lectures and contains a shortest known and most lucid proof of one important theorem. For further reading, I suggest a remarkable book by a remarkable person F. Adams [Ad], and [FH]; then [Go]. For references, see the universal at all times [CR], see also [C1, C2]. Notes of P. Etingof's lectures reflect interesting aspects of representation theory seldom mentioned in first courses, see [E].

The "super" version of these lectures forms Volume 2. For further reading on the theory of certain particular infinite dimensional Lie algebras, highly resembling simple finite dimensional Lie algebras (everything over $\mathbb{C}$), see Kac's book on Kac-Moody Lie algebras [K].

---

[3] http://en.wikipedia.org/wiki/Vladimir_Arnold
[4] http://en.wikipedia.org/wiki/Crafoord_Prize.
[5] http://en.wikipedia.org/wiki/Wolf_Prize.
[6] http://www.shawprize.org/en/index.html
[7] http://en.wikipedia.org/wiki/Joseph_Bernstein

Observe that the term "representation theory" is almost always applied in these lectures to **simple** (Lie) groups and Lie (super)algebras. More precisely, these simple objects constitute a natural core and all other objects, sometimes no less interesting in applications, are somehow derived from the simple ones. In these lectures I will mainly deal with the simple objects warning at the same time of the danger of overpricing them.

My task at ASSMS: Formulation. The noble initiative of the Pakistani government, which the Higher Education Commission tries to implement, is to educate national researchers to make them competitive at the international level. Given an incredibly low level of knowledge in mathematics the students used to get at National Universities so far, it is difficult to offer them reasonably interesting research problems, whereas life is too short to spend a good deal of it solving "purely educational" exercises somebody has solved long ago. Besides, even if the student is capable to memorize the lecture and get the highest score at the exam, (s)he is not necessarily capable of **doing research**, that is solve something that nobody yet did inventing the technique for the occasion. The only way to teach research I know of is to show examples performed by masters. This volume consists of such examples.

Some branches of mathematics (like graph theory) **seem** to be designed for the purpose: A bright student might find a long-standing problem (even with a name) that does not require anything but common sense and ability to solve "Olympiad-type" problems. But this road is usually a dead-end as far as research is concerned: The student learns almost nothing and, even a bright one, is seldom able to continue career of a research mathematician.

Most of the other branches of mathematics (like algebraic geometry), manifestly pregnant with important applications to real life, require a lot of time to master them sufficiently to come within reach of frontiers of research.

The representation theory is a rare field of science where there is room for everybody from bright first year students (who know almost nothing) to deep thinkers and broadly educated experts. The problems in representation theory are usually "with an open end", meaning that having solved one, the researcher can usually see how to extend it by generalizing in various ways ("superize", consider it over various fields, vary the dimension, vary the symmetry in question, and so on).

My task at ASSMS: An example of a solution. In 2004–2006, I read a course on representation theory of Lie algebras and Lie superalgebras at Max-Planck-Institute for Mathematics in the Sciences (MPIMiS, Leipzig), and in 2007, another one, at SMS, Lahore. The main purpose of these courses was to introduce the listeners to the main notions of the representation theory to enable them — as fast as possible (e.g., the time allocated for Ph.D. study at MPIMiS was 2 or 3 years from entrance to submission of the thesis) — to begin working on interesting open research problems.

I've decided to split the course into two volumes. Volume 1 contains the basics of the classical representations theory. (Nevertheless, it also contains several open problems.) Volume 2 contains basic information on the very frontiers of modern (2007) research needed to come to grasps with formulation of open problems.

The soundness of my approach to the choice of the topic and selection of problems is illustrated by results of my former Ph.D. students, my colleagues, and mine. About half of the volume contains totally new results obtained during the past two years. All the results are "with an open end" and in addition to scattered problems (easy to find from Index) a list of open problems is formulated explicitly at the end. The reader is encouraged to contact me (mleites@math.su.se) to avoid nuisance of queueing selecting problems and to inform if something is solved.

The chapters written by A. Lebedev and E. Poletaeva contain main parts of their respective Ph.D. theses. I hope these chapters will serve as models for the readers not yet having a Ph.D. diploma.

Computer-aided scientific research. Most of the open problems offered as possible topics for Ph.D. research in Volume 2 are easier to solve with the help of the `Mathematica`-based package `SuperLie` for scientific research designed by Pavel Grozman. Arnold's problems in this volume are also to be solved with computers' aid. This feature of these problems is another asset: It encourages to master certain basic skills useful in the modern society in general and for a university professor and researcher in particular.

Acknowledgements. I am thankful:

To my teachers who gave their permissions to use their results that constituted Volume 1, and, separately, to É. B. Vinberg and A. O. Onishchik.

I edited the transcripts of the first two Chapters very little trying to preserve the author's style, and, bar correcting typos, only translated Arnold's one. (All the mistakes and typos not edited or inserted are, of course, my responsibility. I will be thankful for all remarks sent to me (mleites@math.su.se) that will help to improve the text for the next printing.)

To my students who participated in Volume 2.

To the chairmen of the Department of Mathematics of Stockholm University (T. Tambour, C. Löfwall, and M. Passare) for the possibility to digress from the routine to do research.

To J. Jost, the Director of MPIMiS, for inviting me to MPIMiS, where I worked as Sophus Lie Professor, and to MPIMiS for excellent working conditions during 2004–6.

To TBSS [8)], and grants of the Higher Education Commission, Pakistan that supported parts of this project.

[8] Teknikbrostiftelse i Stockholm, the Swedish name of The Foundation Bridging Technology and Science in Stockholm.

Notation and several commentaries

$\mathbb{N}$, $\mathbb{Z}_+$, $\mathbb{Z}$ denote, respectively, the sets of positive, non-negative and all integers; $\mathbb{R}$ and $\mathbb{C}$, respectively, the sets of real and complex numbers;

I is the identity operator;

E or I is the unit matrix (E_n or I_n of size $n \times n$);

A_{ij} is the (i, j)th entry of the matrix A; and E_{ij} is the (i, j)th matrix unit;

irrep = irreducible representation (slang or blackboard abbreviation);

$\#(S)$= the cardinality of the set S;

$A := B$ a new notation for B.

Most basic preliminaries. Unless otherwise stated, the ground field k is $\mathbb{C}$ but further on we consider other fields.

V.I.Arnold teaches us that the notion of *group* should be given to students as it appears in actual problems, as groups of motions or symmetries of something. However, to study various properties of the group itself, one distinguishes its abstract features and thus comes to the following abstract definition. A *group* is a set G with a fixed element $e \in G$ called the *unit* (sometimes it is called *identity*, meaning the identity transformation) and operation (*multiplication*) $m \colon G \times G \longrightarrow G$ which to every ordered pair $f, g \in G$ assigns their product fg, and the *inversion map* $i \colon G \longrightarrow G$ which to every $g \in G$ assigns g^{-1} so that, for any $f, g, h \in G$, we have

$$f(gh) = (fg)h; \quad ge = eg = g; \quad gg^{-1} = g^{-1}g = e.$$

The group G is said to be *commutative* if the product satisfies $fg = gf$ for any $f, g \in G$; often the product of commutative groups is denoted by a + and the unit is denoted by 0 and called the *zero*.

A *ring* R is a commutative group with respect to + (called addition, and 0 is its *zero*) and with *another* operation (called multiplication) usually denoted, for brevity, by juxtaposition. The ring is said to be *commutative* if $ab = ba$ for any $a, b \in R$. If every nonzero element of R is invertible, R is called a *skew field*; the skew field with a commutative multiplication is called *field*. The identity with respect to the multiplication is called *unit* or *unity*. A ring A which is at the same time a module over a field and both structures of the module and the two ring multiplications are naturally compatible (distributive with respect to each other, and so on) is called an *algebra*.

A *representation* of the group G in the vector space V is a group homomorphism [9)] $T \colon G \longrightarrow GL_{\mathbb{K}}(V)$ of G into the group of invertible linear operators of a vector space V defined over a field $\mathbb{K}$. This vector space V is called the *space of representation* T or a *G-module*. Two representations $T_1 \colon G \longrightarrow GL(V_1)$ and $T_2 \colon G \longrightarrow GL(V_2)$ are are said to be *equivalent* and we write $T_1 \cong T_2$ if there exists an isomorphism $S \colon V_1 \longrightarrow V_2$ such that $T_2S = ST_1$. (To save space and effort of the typist, I avoid commuting diagrams whenever possible.)

---

[9] This means that $T(fg) = T(f)T(g)$ for any $f, g \in G$.

The following problems naturally arise:

(1) To distinguish and classify "simplest" representations (irreducible, indecomposable, "models", etc). To describe every given representation in terms of the simplest ones.

(2) To study and classify special functions (a term to be defined), operators and spaces that arise in the study of representations, e.g., matrix elements, spherical functions, intertwining operators, and so on.

(3) To study various structures on the space V of the representation r of G, e.g., topology, ring or algebra structures, unitary property, and so on.

Selected applications of representation theory (just to mention a few): A rough classification of chemical elements (an approximation to Mendeleev's table); a method for solution of various differential equations of mathematical physics (Maxwell, Einstein, Liouville, Korteveg–de Vries, Toda lattice, and many, many others). Here is another example, of a more down-to-earth level.

0.0.0.1. Remarks. (1) When a student, I was taught the following physical interpretation: "An irreducible representation" is approximately the same as "an elementary particle". Now I tend to think that this interpretation is passable for particles describable by means of gauge fields with compact gauge groups because their representations are completely reducible. Otherwise, we should consider particles of various levels of elementariness; indecomposable representations correspond to a rough sketch, irreducible ones to the "truly elementary" particles.

(2) All physical laws are invariant with respect to a certain group and are usually formulated in terms of invariant operators (intertwining for some representations).

(3) Unitary property is interpreted as positivity of energy in field theory.

0.0.0.2. A problem (after A.A.Kirillov). Before "perestroika" and collapse of the USSR, children were often sent to summer Pioneer (scout) camps. Life there was not easy, but cheap; a porridge of sorts was their staple food. The cook, however, was often drunk or just did not care, so the portions were highly uneven. Darwin's law "The fittest survives" ruled.

Under these conditions consider a round table with kids sitting at it, a bowl with porridge (some full to the brim, some almost or totally empty) in front of each kid. When the duty teacher (in charge of the discipline) digresses to take a sip of home-brewed vodka disguised as tea, all kids use both hands to apply Lenin's rule of redistribution (rob the robbers) to their nearest neighbors' bowls. Assuming that the on-duty teacher drinks at all times and never ceases, what is the limit distribution of porridge as time tends to infinity?

In strict mathematical terms, the problem reads like this: If initially the k-th kid had amount of porridge of x_k^0 gram, then after the $(i+1)$-st redistribution (s)he had $x_k^{i+1} = \frac{1}{2}(x_{k-1}^i + x_{k+1}^i)$, where $x_{N+1}^i = x_1^i$, the total amount of kids being N. Determine $x_k^\infty := \lim\limits_{i \longrightarrow \infty} x_k^i$.

If we replace the round table with a regular polygon or polyhedron, and the amount of porridge with, say, temperature, we obtain another, no less realistic, physical model.

Chapter 1

Lectures for beginners on representation theory: Finite and compact groups (A. A. Kirillov)

Unless otherwise stated, the ground field $\mathbb{K}$ in these lectures is the field $\mathbb{C}$ of complex numbers although in several Main Theorems only algebraic closedness of $\mathbb{K}$ is vital (sometimes with the same proof). Several crucial theorems are only true under certain restrictions on the characteristic of the ground field. (For example, Maschke's theorem.)

Lecture 1. Complex representations of Abelian groups

1.1. Exercises. Suppose G is a topological space. Let G be a group as well. We say that G is a *topological group* if the multiplication (product) and inversion maps

$$m\colon G \times G \longrightarrow G \quad (m(g,\, h) = gh) \quad \text{and} \quad i\colon G \longrightarrow G \quad (i(g) = g^{-1})$$

are *continuous* in the topology of G.

1.1.1. Prove that the set $\mathbb{R}$ with the usual topology is a topological group with respect to the $+$ operation. (Notation: $\mathbb{R}^{+}$; do not confuse with the set $\mathbb{R}^{\times} := \mathbb{R} \setminus \{0\}$ and the multiplication as the product operation.)

1.1.2. The representation T of the topological group G in a linear space V is *continuous* if the action map

$$a(g, v) = T(g)v \quad \text{for any } g \in G, v \in V$$

is continuous with respect to topologies in G and V.

Prove that every continuous 1-dimensional (over $\mathbb{C}$) representation (i.e., the one with 1-dimensional V) of $\mathbb{R}^{+}$ is of the form

$$T(x) = e^{ax} \quad \text{for a fixed } a \in \mathbb{C} \text{ and any } x \in \mathbb{R}^{+}.$$

Here we identify the operator in the 1-dimensional space over $\mathbb{C}$ with the factor by which the operator multiplies the basis vector of V; i.e., for every nonzero $v \in V$, we have

$$T(x)v = e^{ax} \cdot v.$$

1.1.3. The space V and the representation T in V are called *decomposable* if V can be represented as the direct sum of two subrepresentations, V_1 and V_2. This means that every nonzero $x \in V$ can be represented as $x = x_1 + x_2$, where $x_1 \in V$, $x_2 \in V_2$ and $T(g)x_i \in V_i$ for all $g \in G$.

If V is decomposable, we write $V = V_1 \oplus V_2$ and $T = T_1 \oplus T_2$, where $T_i = T|V_i$.

Prove that every representation is the direct sum of indecomposable representations.

1.1.4. The representation $T\colon G \longrightarrow GL(V)$ is *reducible* if V has a nontrivial ($\neq V, \{0\}$) invariant subspace.

For $G = \mathbb{R}^+$, define a representation T in $\mathbb{C}^2 = \mathrm{Span}(v_1, v_2)$ by setting

$$T(a)v_1 = v_1, \quad T(a)v_2 = v_2 + av_1.$$

Prove that T is reducible but indecomposable.

1.2. Hermitian or sesquilinear forms. The form $(\cdot,\cdot)\colon V \times V \longrightarrow \mathbb{C}$ is said to be *pseudo-hermitian* or *sesquilinear* if it is

1) linear in the first argument, i.e., $(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z)$ for any $\alpha, \beta \in \mathbb{C}$, $x, y, z \in V$;

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

If a pseudo-hermitian form is such that

$$(x, x) \geq 0 \ \text{ and } (x, x) = 0 \text{ if and only if } x = 0,$$

then the form is said to be *hermitian*.

1.2.1. Exercise. In every finite dimensional V with an hermitian form, there is an orthonormal basis, i.e., a set of vectors $e_1, \dots, e_n \in V$ which is a basis of V and such that $(e_i, e_j) = \delta_{ij}$.

Proof: An orthogonalization process. Start with any basis $\{e'_1, \dots, e'_n\}$. Set

$$\begin{aligned} e_1 &= a_{11}e'_1 \\ e_2 &= a_{21}e'_1 + a_{22}e'_2 \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ e_n &= a_{n1}e'_1 + \dots + a_{n1}e'_n. \end{aligned}$$

In a fixed basis, the passage "operator $\mapsto$ its matrix" will be denoted by $\circ\colon T(g) \mapsto T^\circ(g)$. $\square$

1.2.2. Exercise. $T_1 \cong T_2$ if, in appropriate bases of V_1 and V_2, we have $T_1^\circ(g) = T_2^\circ(g)$ for all $g \in G$.

1.2.3. Exercise. A given representation T of the group G is decomposable (resp. reducible) if, in an appropriate basis, $T^\circ(g)$ is of the form

$$\begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}, \quad \text{resp. } \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}, \quad \begin{matrix}\text{where matrices } A \text{ and } D \text{ are of the same size} \\ \text{for any } g \in G.\end{matrix}$$

Having selected a basis of V we identify $GL(V)$ with the group $GL(n)$ of invertible $n \times n$ matrices, where $n = \dim V$. If the field of definition, $\mathbb{K}$, has to be specified, we write $GL_{\mathbb{K}}(V)$ or $GL(n; \mathbb{K})$.

1.2.4. Unitarity. The element $A \in GL(n)$ is said to be *unitary* if $AA^* = E$, where $A^* = \overline{A}^t$. A representation $T\colon G \longrightarrow GL(V)$ in an hermitian space V is said to be *unitary*, if

$$(T(g)x, T(g)y) = (x, y).$$

Exercise. A representation $T\colon G \longrightarrow GL(V)$ is unitary if and only if in an orthonormal basis of V the matrices $T^\circ(g)$ are unitary for all $g \in G$.

1.3. The tensor product of representations. Let $V_1 = \operatorname{Span}(x_1, \dots, x_k)$ and $V_2 = \operatorname{Span}(y_1, \dots, y_l)$. Then the *tensor product* of V_1 and V_2 is defined by either of the definitions:

(a) $V_1 \otimes V_2 = \operatorname{Span}(x_i y_j \mid 1 \leq i \leq k, 1 \leq j \leq l)$
(b) $V_1 \otimes V_2 = V_1 \times V_2 / W$, where W is spanned by the pairs

$$\begin{aligned} &(x_1 + x_2, y) = (x_1, y) + (x_2, y), \\ &(x, y_1 + y_2) = (x, y_1) + (x, y_2), \\ &(cx, y) = (x, cy) = c(x, y) \quad \text{for any } c \in \mathbb{C}. \end{aligned}$$

The element $x_i y_j$ from definition (a) and the class of the element (x_i, y_j) (mod W) from definition (b) is denoted by $x_i \otimes y_j$.

Accordingly, the tensor product of representations T_1 and T_2 acting in V_1 and V_2, respectively, is defined by one of the following formulas:

(a) $((T_1 \otimes T_2)(g))\,(x_k \otimes y_l) = \sum\limits_{i,j} T_1^\circ(g)_{ki} T_2^\circ(g)_{lj} x_i \otimes y_j$.
(b) $((T_1 \otimes T_2)(g))\,(x \otimes y) = T_1(g)x \otimes T_2(g)y$.

Clearly, we have:

A) Let V_1 and V_2 be vector spaces of row-vectors of dimension k and l, respectively. There is an isomorphism $\varphi\colon V_1 \otimes V_2^* \simeq \operatorname{Mat}(k \times l)$, where $\operatorname{Mat}(k \times l)$ is the space of $k \times l$ matrices, given by

$$(v_1 \otimes v_2^*)(w) = \varphi(v_1 \otimes v_2^*) w_2^t \text{ for any } v_1 \in V_1,\ v_2, w \in V_2 \text{ and } z \in V_1 \otimes V_2^*.$$

B) If $T_1 \simeq T_1'$ and $T_2 \simeq T_2'$, then $T_1 \otimes T_2 \simeq T_1' \otimes T_2'$.

C) If at least one of the representations T_1 and T_2 is reducible (decomposable), then $T_1 \otimes T_2$ is reducible (decomposable).

D) If V_1 and V_2 are endowed with hermitian products H_1 and H_2, respectively, then

$$(x_1 \otimes y_2, x_1' \otimes y_2') := H_1(x_1, x_1')H_2(y_2, y_2')$$

is an hermitian product on $V_1 \otimes V_2$.

1.3.1. Two rings. A) On the set Reps (G) of equivalence classes of representations, we have two operations: The addition, +, which is $\oplus$, and the product, $\otimes$. These operators satisfy

$$T_1 \otimes (T_2 \oplus T_3) = (T_1 \otimes T_2) \oplus (T_1 \otimes T_3),$$

with $T_1 \oplus T_2$ acting in $V_1 \oplus V_2$ and $T_1 \otimes T_2$ acting in $V_1 \otimes V_2$. Clearly, the zero representation (for which $T(G) = 0$) acting in the 0-dimensional space $\mathbf{0} := \{0\}$ is the zero with respect to $\oplus$, but there is no opposite element. The trivial 1-dimensional module (for which $T(G) = 1$) is the unit with respect to tensoring, but there is no inverse element. So Reps (G) is not a ring, actually, but a semi-ring.

By formally defining $-T$ as an element such that $(-T) \oplus T = \mathbf{0}$ (as we do in order to define negative numbers and zero being given only positive ones) we can extend Reps (G) to a module $\mathrm{Reps}_{\mathbb{Z}}(G)$ over $\mathbb{Z}$ by setting:

$$(nT) := \begin{cases} \underbrace{T \oplus T \dots T \oplus T}_{n \text{ summands}} & \text{for any nonnegative } n \in \mathbb{Z}; \\ -(nT) & \text{for any nonpositive } n \in \mathbb{Z}. \end{cases}$$

Furthermore, we can make $\mathrm{Reps}_{\mathbb{Z}}(G)$ into a module $\mathrm{Reps}_{\mathbb{K}}(G)$ over the ground field $\mathbb{K}$ by setting

$$\mathbb{K} \otimes_{\mathbb{Z}} \mathrm{Reps}_{\mathbb{Z}}(G).$$

Thus, the expressions of the form $\sum c_i T_i$, where T_i are representations of G and the c_i belong to the ground field, constitute the *algebra of representations* of G.

Since, as we will see, all modules over finite (and compact) groups G are completely reducible, it suffices (at least, for such groups) consider only irreducible representations T_i. One of the first problems in the representation theory: Express the tensor product of two representations as the sum (perhaps, with multiplicities) of irreducible representations.

B) The ring $\mathbb{K}[G] = \Big\{ \sum\limits_{g_i \in G} c_i g_i \mid c_i \in \mathbb{K} \Big\}$ in which

$$\begin{aligned} \sum c_i g_i + \sum c_i' g_i &= \sum (c_i + c_i') g_i, \\ \sum c_i g_i \cdot \sum c_j' g_j &= \sum_{i,j} c_i c_j' g_i g_j \end{aligned}$$

is called the *group ring* of G.

1.3.1a. Example. Consider $\mathbb{Z}/2 = \{\bar{0}, \bar{1}\} \simeq \{\pm 1\}^\times$. For every representation $T\colon G \longrightarrow GL(V)$, we see that $T(e) = E \in GL(V)$.

So, to determine $T(\mathbb{Z}/2)$, it suffices to determine $T(-1)$. Observe that

$$T(-1)T(-1) = T(1) = E.$$

Exercise. If $A^2 = E$, then, in a certain basis, $A = \operatorname{diag}(\varepsilon_1, \dots, \varepsilon_n)$, where $\varepsilon_i = \pm 1$ for each i.

Hint. Observe that $V = V_+ \oplus V_-$, where $x_+ = \frac{1}{2}(x + Ax)$ and $x_- = \frac{1}{2}(x - Ax)$ are components of $x \in V$ belonging to $V_\pm$. Clearly, the $V_\pm$ are the eigenspaces of A, and A is the reflection with respect to V_+.

Corollary. *There are two distinct 1-dimensional representations of* $\mathbb{Z}/2$*; every representation is the direct sum of such representations. The presentation in the form of the direct sum is not unique.*

1.3.1b. Example. Consider

$$\mathbb{Z}/n = \{\bar{0}, \dots, \bar{n}\} \simeq \left\{1, \varepsilon = e^{\frac{2\pi i}{n}}, \dots, \varepsilon^{n-1} = e^{\frac{2(n-1)\pi i}{n}}\right\}^\times \simeq$$

the group of proper movements of a regular n-gon.

For any representation T of $\mathbb{Z}/n$, consider the following auxiliary operations:

$$P_0 = \frac{1}{n}[T(1) + T(\varepsilon) + \dots + T(\varepsilon^{n-1})],$$
$$P_1 = \frac{1}{n}[T(1) + \varepsilon T(\varepsilon) + \dots + \varepsilon^{n-1} T(\varepsilon^{n-1})],$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$P_{n-1} = \frac{1}{n}[T(1) + \varepsilon^{n-1} T(\varepsilon) + \dots + \varepsilon^{(n-1)^2} T(\varepsilon^{n-1})]$$

i.e.,

$$P_k = \frac{1}{n}\sum_{i=0}^{n-1} \varepsilon^{ik} T(\varepsilon^i).$$

Exercise. Prove that

(a) $P_k P_l = \begin{cases} P_k = P_l & \text{if } k \equiv l \pmod{n}, \\ 0 & \text{otherwise;} \end{cases}$

(b) $\sum\limits_{k=0}^{n-1} P_k = \mathrm{Id}$.

(c) Every operator P, such that $P^2 = P$, is the projection onto a subspace $V_+ \subset V$ parallel to a transversal subspace V_-, i.e., a subspace such that $V_+ \oplus V_- = V$.

Hint. Observe that $V_+ = \{x_1 \mid x_1 = Px\}$, and $V_- = \{x_2 \mid x_2 = x - Px\}$.

Exercise. Denote by V_k the image of P_k. Since $P_k P_l = 0$ for $k = l$, it follows that $V_k \cap V_l = 0$ for $k \neq l$; since $\sum P_k = \mathrm{Id}$, it follows that $V = \oplus V_k$.

Lecture 2. Schur's lemma, Burnside's theorem

2.1. Summary of Lecture 1. *Every $\mathbb{Z}/n$-module V is the direct sum of irreducible 1-dimensional representations.*

There are n distinct 1-dimensional spaces V_k in which representations of $\mathbb{Z}/n$, denote them T_k, where $k = 0, \ldots, n-1$, act so that $T_k(\varepsilon^l) = e^{il}$. The module V can be uniquely represented as the sum of isotypical components $\oplus\widetilde{V}_k$, whereas each $\widetilde{V}_k$ can be (non-uniquely) represented as the direct sum of several identical copies $\widetilde{V}_k = V_k \oplus \cdots \oplus V_k$.

Since the T_k are 1-dimensional, so is the tensor product of any two of them:

$$(T_k \otimes T_l)(\varepsilon^i) = T_k(\varepsilon^i) \otimes T_l(\varepsilon^i) = \varepsilon^{ik} \cdot \varepsilon^{il} = \varepsilon^{i(k+l)} = T_{k+l}(\varepsilon^i).$$

2.1.1. Corollary. *The ring of representations of $\mathbb{Z}/n$ is isomorphic to $\mathbb{C}[\mathbb{Z}/n]$.*

2.1.2. Exercise. $\mathbb{C}[\mathbb{Z}/n] \simeq \mathbb{C}^n$, as vector spaces.

Hint. For a basis, take $P_k = \dfrac{1}{n}\sum\limits_l \varepsilon^{lk}[\varepsilon^l]$, where $[\varepsilon^l] \in \mathbb{Z}/n$ is the element represented by $\exp\left(\dfrac{2\pi il}{n}\right)$.

2.1.3. Corollary (Main Corollary)**.** *The description of representations of any finite commutative group. (Formulate it on your own.)*

2.2. Non-commutative groups. The simplest example is S_3, the permutation group on 3 elements. Other realizations of S_3:

(1) all motions of the equilateral triangle,

(2) conformal group with $z \mapsto \dfrac{1}{z}$ and $z \mapsto 1 - z$ as generators,

(3) $SL(2; \mathbb{Z}/2)$,

(4) The group generated by a and b subject to the relations

$$a^3 = e, \quad b^2 = e, \quad bab = a^2.$$

2.2.1. Exercise. 1) Prove that definitions (1)–(4) of S_3 are equivalent, i.e., the groups obtained are isomorphic.

2) Consider the subgroup of S_3 generated by a; notation: $\langle a\rangle$. Prove that $\langle a\rangle \simeq \mathbb{Z}/3$. Describe this group in the other four incarnations given in subsect. 2.2.

Given a representation $T\colon S_3 \longrightarrow GL(V)$, we obtain a representation of $\mathbb{Z}/3$, as $T|\langle a\rangle$ determined by the operators E, $T(a)$, $T(a^2)$. Hence, $V = V_0 \oplus V_1 \oplus V_2$, where V_i is the eigenspace of $T(a)$ corresponding to the eigenvalue ε^i. Consider the action of $T(b)$ on each V_i: For any $x \in V_0$, set $T(b)x = y$. Then

$$T(a)y = T(a)T(b)x = T(ab)x = T(b^2ab)x = T(b)T(bab)x =$$
$$T(b)T(a^2)x = T(b)x = y.$$

Hence, $y \in V_0$. For any $x \in V_1$, set $y = T(b)x$. Then

$$T(a)y = \ldots = T(b)T(a^2)x = T(b)\varepsilon^2 x.$$

Hence, $T(b)(V_1) = V_2$.

We similarly prove that $T(b)(V_2) = V_1$.

If V is irreducible, then either (1) $V = V_0$ or (2) $V = V_1 \oplus V_2$.

In case (1), $T(a) = T(a^2) = E$ and we actually have the representation of $S_3/\langle a\rangle \simeq \mathbb{Z}/2$. There are two irreducible representations of $\mathbb{Z}/2$: With $T(b) = E$ and with $T(b) = -E$.

In case (2), take $x \in V_1$ and let $y = T(b)x \in V_2$. Then

$$T(b)y = T^2(b)x = x.$$

Hence, $\mathrm{Span}(x, y)$ is S_3-invariant; since it is an irreducible representation, we see that $V_1 \oplus V_2 = \mathrm{Span}(x, y)$ and the matrix realization is

$$T(a) = \begin{pmatrix} \varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix}, \quad T(b) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

2.2.2. Exercise*. *Every S_3-module is the direct sum of several copies of the above two distinct irreducible representations and the trivial one.*

One of the “most main” theorems in representation theory is the following one:

2.3. Theorem. *Each representation of the finite group G is the direct sum of its irreducible representations.*

Comment: This theorem implies that, for every subrepresentation there is a direct complement.

Proof of Theorem 2.3 follows from the next Lemmas 2.3.1–2.3.5.

2.3.1. Lemma. *Every representation $T\colon G \longrightarrow GL(V)$ of a given finite group G is equivalent to a unitary one, i.e., we may assume that V is endowed with an hermitian product $(\cdot,\cdot)$.*

Proof. By definition, $\widetilde{T} \simeq T \Longleftrightarrow \widetilde{T}(g) = Q^{-1}T(g)Q$ for an invertible Q. If $\widetilde{T}$ is unitary, then

$$\widetilde{T}(g)\widetilde{T}^* = E$$

or

$$Q^{-1}T(g)QQ^*T^*(g)(Q^*)^{-1} = E$$

or

$$T(g)QQ^*T^*(g) = QQ^*.$$

Set $A = QQ^*$; we have

$$TAT^* = A. \tag{1.1}$$

Set $S(g)\colon A \mapsto T(g)AT(g)$.

2.3.2. Exercise. Prove that $S(g)\colon G \longrightarrow GL(\mathrm{End}(V))$ is a representation. Eq. (1.1) means that A is S-invariant.

2.3.3. How to construct S-invariant matrices?

Exercise. Prove that $\mathring{A} := \sum\limits_{g\in G} S(g)A$ is S-invariant.

Remark. If A is S-invariant, then $\mathring{A} = |G| \cdot A$, and therefore considering $\mathring{A}$ for every $A \in \mathrm{End}(V)$, we obtain all S-invariant matrices.

So, to prove Lemma 2.3.1, we have to find, among S-invariant matrices A, a matrix of the form QQ^* for an invertible Q.

2.3.4. Exercise. We have $A = QQ^*$ for an invertible Q if and only if

1) $A^* = A$,

2) $(Ax, x) > 0$ for any nonzero $x \in V$. This property of the operator (matrix) A is denoted $A > 0$.

Hint. Define a new inner product in V by setting

$$\{x, y\} = (Ax, y).$$

Let $e_1, \dots, e_n$ be the initial orthonormal basis of V and $e'_1, \dots, e'_n$ an orthonormal basis for $\{\cdot, \cdot\}$. For Q, take the matrix of the passage $\{e_i\}_{i=1}^n \longrightarrow \{e'_i\}_{i=1}^n$. It remains to observe that if $A > 0$, then so are the matrices $S(g)A = T(g)AT^*(g)$ for all $g \in G$. Hence, so is $\mathring{A} = \sum\limits_{g\in G} S(g)A$. So if we take E for A, then E^0 is an invariant matrix of the form QQ^*. □

2.3.5. Lemma. *For unitary representations, indecomposability is equivalent to irreducibility.*

Proof. Irreducibility clearly implies indecomposability.

$\Longleftarrow$ Let T be indecomposable and unitary. Suppose it is not irreducible. Let $V_1 \subset V$ be an invariant subspace, V_2 its orthogonal complement, i.e.,

$$V_2 = V_1^\perp = \{x \in V \mid (x, y) = 0 \ \text{ for any } y \in V_1\}.$$

Then

$$(T(g)x,\ y) = (x,\ T^*(g)y) = (x,\ T(g^{-1})y) = 0$$

for all $x \in V_2$, $y \in V_1$. □

2.4. The regular representation. This is the name for the representation of the group G in the space of functions on G. Clearly, the regular representation is reducible. Remarkably, if G is finite, the regular representation contains each irreducible representation with multiplicity equal to its dimension.

Theorem (Burnside's theorem)**.** *Let G be a finite group, $|G| = N$; let $T_1, \dots, T_a$ be the full set of irreducible representations of G, let* $\dim V_i = n_i$*. Then*

$$N = \sum n_i^2.$$

Proof. Let $L(G) = G^{\mathbb{K}} = \mathrm{Maps}(G, \mathbb{K})$ the space of functions on G; let $\dim L(G) = N$. For the basis of $L(G)$ take δ*-functions*

$$\delta_g(h) = \begin{cases} 1 & \text{if } g = h, \\ 0 & \text{otherwise.} \end{cases}$$

In $L(G)$, introduce the hermitian product $(\cdot, \cdot)$ by setting

$$(f_1, f_2) = \sum_{g \in G} f_1(g)\overline{f_2(g)}.$$

2.4.1. Exercise. The functions δ_g form an orthonormal basis with respect to $(\cdot, \cdot)$.

2.4.2. Another basis. Let us construct another basis such that, to every n-dimensional irreducible representation, there correspond n^2 elements of this other basis.

Let T be an irreducible representation of dimension n. Then $T(g)$ can be considered as a matrix-valued function on G and each $T(g)_{ij}$ is a numerical function.

2.4.3. Exercise. Show that the $T_{ij}(g)$ are linearly independent functions.

Remark. For equivalent representations $\widetilde{T} \cong T$, where $\widetilde{T} = CTC^{-1}$, we obtain different functions $\widetilde{T}_{ij}$, but since

$$\begin{aligned} \widetilde{T}_{ij}(g) &= \sum_{k,l} C_{ik}(C^{-1})_{lj} T_{kl}(g), \\ T_{kl}(g) &= \sum_{i,j} (C^{-1})_{ki} C_{jl} \widetilde{T}_{ij}(g), \end{aligned}$$

we see that $\mathrm{Span}(T_{ij})_{i,j} \cong \mathrm{Span}(\widetilde{T}_{ij})_{i,j}$.

To complete the proof of Burnside's theorem, we have to verify that $L(G)$ is the direct sum of $\mathrm{Span}(T_{ij})$ for various T's, that $\mathrm{Span}(T_{ij})$ is orthogonal to $\mathrm{Span}(U)_{k,l}$ for $T \not\cong U$ and prove Exercise 2.4.3. $\square$

2.5. Lemma (I. Schur). *Let T_1 and T_2 be irreducible representations of G in V_1 and V_2, respectively, and let a linear map $C\colon V_1 \longrightarrow V_2$ be such that*

$$T_2 C = C T_1. \tag{1.2}$$

Then, over $\mathbb{C}$, either $C = 0$ or C is an isomorphism (hence, $T_1 \simeq T_2$). [1)]

The operators C satisfying condition (1.2) are called *intertwining operators* and

$$I(T_1, T_2) = \dim \text{ (the space of intertwining operators)}$$

the *intertwining number.*

2.5.1. Lemma. *Over $\mathbb{C}$, for irreducible representations T_1 and T_2, the expression* (1.2) *is equivalent to the following fact:*

$$I(T_1, T_2) = \begin{cases} 0 & \text{if } T_1 \not\simeq T_2, \\ 1 & \text{if } T_1 \simeq T_2. \end{cases}$$

Indeed, if $T_1 \simeq T_2$, then $T_1 = C T_2 C^{-1}$ i.e., eq. (1.2) holds. If C' is another intertwining operator, then $C' + \lambda C$ is also an intertwining operator. But $\det(C' + \lambda_0 C) = 0$ for some $\lambda_0 \in \mathbb{C}$. By Schur's lemma $C' + \lambda_0 C$ is not invertible, hence, $C' = -\lambda_0 C$. $\square$

2.5.2. Exercise. Consider the representation $T\colon \mathbb{Z}/4 \longrightarrow GL(2; \mathbb{R})$ given by

$$T(\overline{0}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad T(\overline{1}) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad T(\overline{2}) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad T(\overline{3}) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

(1) Prove that T is irreducible, but $I(T, T) = 2$.

(2) Is $T^{\mathbb{C}}$, given by the same formulas over $\mathbb{C}$, irreducible? If not, what are the irreducible components?

2.5.3. Exercise. Finish the proof of Burnside's theorem.

2.5.4. Exercise. Let G be a finite group, $T_1, \dots, T_a$ the full set of its irreducible representations, $V_i = \mathrm{Span}((T_i)_{kl} \mid \text{ all possible } k, l)$. Prove that V_i is invariant with respect to the left and right translations by G.

Lecture 3. The structure of the group ring

Recall that $\mathbb{K}[G] := \{\sum c_i g_i \mid c_i \in \mathbb{K}, g_i \in G\}$. In this section, $\mathbb{K} = \mathbb{C}$.

---

[1] Over algebraically non-closed fields, there are more possibilities, see, e.g., [Ad] for the case $\mathbb{K} = \mathbb{R}$.

3.1. Another interpretation of $\mathbb{C}[G]$. Let $\sum c_i g_i$ be considered as the function on G such that $\sum c_i g_i(g_j) = c_j$. How to describe the convolution, $*$, in $\mathbb{C}[G]$ in these terms? Let $f_1 = \sum\limits_{g\in G} f_1(g)g$, and $f_2 = \sum\limits_{h\in G} f_2(h)h$. Then

$$(f_1 * f_2) = \sum_{g,h} f_1(g)f_2(h)gh;$$

hence,

$$(f_1 * f_2)(t) = \sum_{gh=t} f_1(g)f_2(h).$$

Therefore,

$$(f_1 * f_2)(t) = \sum_g f_1(g)f_2(g^{-1}t) = \sum_g f_1(tg^{-1})f_2(g). \tag{1.3}$$

3.2. Theorem (Maschke's theorem)**.** *Let G be a finite group whose irreducible representations are of dimensions $n_1, \dots, n_k$. Then*

$$\mathbb{C}[G] \cong \mathrm{Mat}(n_1) \oplus \cdots \oplus \mathrm{Mat}(n_k).$$

In other words, $\mathbb{C}[G] \ni f \longleftrightarrow (C_1, \dots, C_k)$ such that

$$\begin{aligned} f + \widetilde{f} &\longleftrightarrow (C_1 + \widetilde{C}_1, \dots, C_k + \widetilde{C}_k)\\ f \cdot \widetilde{f} &\longleftrightarrow (C_1 \cdot \widetilde{C}_1, \dots, C_k \cdot \widetilde{C}_k)\end{aligned} \tag{1.4}$$

For example, if G is abelian, then $n_1 = \cdots = n_k = 1$ and $\mathbb{C}[G]$ is the "ring" $\mathbb{C}^k$ with component-wise multiplication.

Proof. Let $T_1, \dots, T_k$ be the full set of irreducible representations of G.

To every $f \in \mathbb{C}[G]$ we assign the collection of matrices:

$$f = \sum_g f(g)g \mapsto (T_1(f), \dots, T_k(f)), \text{ where } T_i(f) = \sum_{g\in G} f(g)T_i(g). \tag{1.5}$$

Clearly, the map (1.5) is linear, so it suffices to verify that $f \cdot g$ goes into the product of collections for delta-functions.

3.2.1. Exercise. Verify that the map (1.5) preserves multiplicativity.

Let us prove that the map (1.5) is onto. Indeed,

$$\dim \mathbb{C}[G] \overset{\text{Burnside's Th.}}{=} N = \sum n_i^2.$$

It remains to prove that the map (1.5) is *mono*.

3.2.2. Exercise. If $T_i(f) = 0$ for all irreducible representations T_i, then for any, not necessarily irreducible, representation T we have

$$T(f) = \sum f(g)T(\cdot) \equiv 0.$$

Hint. Represent T in the form of the direct sum of irreducible representations. Now, observe that if T is the regular representation, then $T(f) = 0 \iff f = 0$.

$\square$

3.3. Another proof of Theorem 3.2. Recall orthogonality relations of matrix elements (see [Ad]):

Let T_1 and T_2 be two irreducible representations of G. Replacing them, if needed, by equivalent ones, we may assume that T_1 and T_2 are unitary. If C is an intertwining operator, then

$$T_2^{-1}(g)CT_1(g) = C$$

(or, since T_1 and T_2 are unitary)

$$T_2^*(g)CT_1(g) = C.$$

Clearly, if $C\colon V_1 \longrightarrow V_2$ is an arbitrary operator, then

$$C^0 = \sum_{g\in G} T_2^*(g)CT_1(g)$$

is also an intertwining operator.

Now suppose $T_1 \not\simeq T_2$. Then $C^0 = 0$ for any C (by Schur's lemma).

$$C_{ij}^0 = \sum_{g\in G}\sum_{k,l}(T_2^*(g))_{ik}C_{kl}(T_1(g))_{lj} = \sum_{k,l} C_{kl}(\sum_{g\in G} T_{ki}^2(g)T_{lj}^1(g)). \tag{1.6}$$

Hence,

$$\sum_{g\in G} T_{ki}^2(g)T_{lj}^1(g) = 0 \text{ for any } i,j,k,l;$$

which means that $T_{ki}^2 \perp T_{lj}^1$.

Let now $T_1 = T_2 = T$. Clearly, E is an intertwining operator.

By Schur's lemma, $\dim I(T,T) = 1$; hence, any intertwining operator is of the form λE, where $\lambda \in \mathbb{C}$. Hence, for any C, we should have $C^0 = \lambda_0 E$. To determine λ_0, let us compute $\operatorname{tr} C^0$. We have:

$$\operatorname{tr} C^0 = \sum_{g\in G} \operatorname{tr} T^{-1}(g)CT(g) = N\cdot \operatorname{tr} C.$$

On the other hand,

$$\operatorname{tr} C^0 = \operatorname{tr} \lambda_0 E = \lambda_0 \cdot \dim T.$$

Hence,

$$\lambda_0 = \frac{N\cdot \operatorname{tr} C}{\dim T}.$$

Finally,

$$(C^0)_{ij} = \begin{cases} 0 & \text{if } i \neq j, \\ \dfrac{N\cdot \operatorname{tr} C}{\dim T} & \text{if } i = j. \end{cases} \tag{1.7}$$

Set $C := E_{kl}$. Having inserted E_{kl} into (1.6) we obtain

$$\sum_g \overline{T_{ki}(g)} T_{lj}(g) = \frac{N}{\dim T} \cdot \delta_{kl}\delta_{ij}. \tag{1.8}$$

Let us return to the second proof. Consider $f(g) = \overline{T^{(i)}_{kj}(g)}$. What collection of matrices corresponds to it? $C_l = 0$ for all $l \neq i$ and $C_i = \dfrac{N}{n_i} E_{kj}$, where $E_{kj} \in \mathrm{Mat}(n_i)$.

NB! The representations $\overline{T}$ and T are not necessarily equivalent: Consider $G = \mathbb{Z}/n$, where $n > 2$.

3.4. The center $Z(\mathbb{C}[G])$ of $\mathbb{C}[G]$.

3.4.1. Exercise. Prove that

$$Z(\mathbb{C}[G]) = \{(\lambda_1 E_{n_1}, \dots, \lambda_k E_{n_k}) \mid \lambda_i \in \mathbb{C}\}.$$

What functions correspond to the center?

Answer. As follows from the proof of Theorem 3.2, the functions that constitute $Z(\mathbb{C}[G])$ are exactly the functions

$$\chi_i(g) = \sum_{k=1}^{n_i} T^i_{kk}(g) = \operatorname{tr} T_i(g) \tag{1.9}$$

called the *characters of the representations* T_i.

Thus, the center of $\mathbb{C}[G]$ consists of the characters of G.

Corollary. *All the elements from the center of* $\mathbb{C}[G]$ *are constants on the conjugacy classes of* G, *i.e.,*

$$\chi(ghg^{-1}) = \chi(h) \text{ for any } \chi \in Z(\mathbb{C}[G]). \tag{1.10}$$

Indeed, $\operatorname{tr}\ ABA^{-1} = \operatorname{tr}\ A^{-1}AB = \operatorname{tr} B$.

3.4.2. Exercise. Prove the opposite statement:

$$\text{If } \chi(ghg^{-1}) = \chi(h), \text{ then } \chi \in \mathbb{Z}(\mathbb{C}[G]). \tag{1.11}$$

Hint. It suffices to check that $\chi\delta_g = \delta_g\chi$.

Thus,

$$\dim Z(\mathbb{C}[G] = \#(\text{characters of irreps}) \overset{\text{Maschke's Th.}}{=} \#(\text{conjugacy classes}).$$

3.4.3. Exercise. Prove that if $\dim T_i = 1$ for all irreducible representations of G, then G is abelian.

Hints. (1) Apply the isomorphism $\mathbb{C}[G] \cong \overset{k}{\underset{i=1}{\oplus}} \mathrm{Mat}(n_i)$.

(2) Prove (with the help of Burnside's and Maschke's theorems) that each conjugacy class has exactly one element. (Recall that Burnside's theorem is the statement "$N = \sum n_i^2$".)

3.5. Theorem. *N is divisible by n_i for each irreducible representation T_i.*

Proof. In $Z(\mathbb{C}[G])$, select two bases:

(1) The characters $\chi_1, \dots, \chi_k$ of irreducible representations,

(2) The characteristic functions $\varphi_1, \dots, \varphi_k$ of conjugacy classes.

Let $U = (U_{ij})$ be the transition matrix, i.e.,

$$\chi_i = \sum_j U_{ij} \varphi_j.$$

Let $V = U^{-1}$, then

$$\varphi_j = \sum_i V_{ji} \chi_i. \tag{1.12}$$

Since $\varphi_i * \varphi_j \in Z(\mathbb{C}[G])$, it follows that

$$\varphi_i * \varphi_j = \sum_l A^l_{ij} \varphi_l. \tag{1.13}$$

Observe that $A^l_{ij} \in \mathbb{Z}$. Indeed, if g belongs to the l-th conjugacy class, then the value of the right hand side at g is equal to A^l_{ij}; while, on the other hand, the value of the right hand side is equal to

$$(\varphi_i * \varphi_j)(g) = \sum_{hk=g} \varphi_i(h) \varphi_j(k) \in \mathbb{Z}.$$

3.5.1. Exercise. Describe the conjugacy classes and compute the coefficients A^l_{ij} for S_3 (in other words, compose the multiplication table for the functions φ_i).

On the other hand, inserting (1.12) into (1.13) we obtain

$$\left(\sum_s V_{is} \chi_s\right) * \left(\sum_p V_{jp} \chi_p\right) = \sum_{l,q} A^l_{ij} V_{lq} \chi_q.$$

But since $\chi_s * \chi_p = \delta_{ps} \cdot \dfrac{N}{n_p} \chi_p$, it follows that

$$\frac{N}{n_q} V_{iq} V_{jq} = \sum_l A^l_{ij} V_{lq}. \tag{1.14}$$

□

3.6. Let us show that $\dfrac{N}{n_q} V_{iq}$ is an integer algebraic number, i.e., a root of a polynomial $x^n + \ldots + a_1 x + a_0$ with $a_i \in \mathbb{Z}$. Indeed, set $B_{lj} = A^l_{ij}$, $x = (x_1, \ldots)$, where $x_i = V_{iq}$. Then (1.14) becomes

$$Bx = \lambda x.$$

Hence, λ is root of the characteristic polynomial of B, i.e., an integer algebraic number.

A reformulation: Set $D = \operatorname{diag}(d_1, \dots, d_k)$, where $d_i = \dfrac{N}{n_i}$. Then the elements of VD are integer algebraic numbers.

Observe that the elements of U are also integer algebraic numbers: Since $\chi_i = \sum U_{ik}\varphi_k$, it follows that $U_{ik} = \chi_i(g)$, where g belongs to the kth conjugacy class. But $\chi_i(g) = \operatorname{tr} T_i(g)$ and, since $(T_i(g))^R = E$ for some R (explain: Why?), the eigenvalues of $T_i(g)$ are integer algebraic numbers, moreover, roots of unity. Hence, $\chi_i(g)$ is equal to the sum of integer algebraic numbers; but the latter ones form a ring, so $\chi_i(g)$ is an integer algebraic number. But then the elements of $UVD = D$ are also integer algebraic numbers.

3.6.1. Exercise. Verify that a rational number q is an integer algebraic number if and only if $q \in \mathbb{Z}$. Thus, $\dfrac{N}{n_i} = d_i \in \mathbb{Z}$.

Lecture 4. Intertwining operators and intertwining numbers

4.1. For two representations $T_i\colon G \longrightarrow GL(V_i)$, where $i = 1, 2$, the operator $C\colon V_1 \longrightarrow V_2$ is said to be an *intertwining* one if the diagram

$$\begin{tikzcd} V_1 \arrow[r] \arrow[d, "T_1(g)"'] & V_2 \arrow[d, "T_2(g)"] \\ V_1 \arrow[r] & V_2 \end{tikzcd}$$

commutes for any $g \in G$. Clearly, the collection of intertwining operators can be naturally endowed with a linear space structure. This space is denoted by $\operatorname{Hom}_G(V_1, V_2)$ and let $I(T_1, T_2) := \dim \operatorname{Hom}_G(T_1, T_2)$. Then

$$\begin{aligned} I(T_1 \oplus T_2, T_3) &= I(T_1, T_3) + I(T_2, T_3) \\ I(T_4, T_5 \oplus T_6) &= I(T_4, T_5) + I(T_5, T_6), \end{aligned}$$

whereas the equality

$$I(T_1, T_2) = I(T_2, T_1) \tag{1.15}$$

is false, generally. Indeed:

4.1.1. Exercise. Let

$$\begin{aligned} &T_1\colon \mathbb{R}^+ \longrightarrow GL_{\mathbb{R}}(2), \quad && T_1(a) = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \\ &T_2\colon \mathbb{R}^+ \longrightarrow GL_{\mathbb{R}}(1), \quad && T_2(a) = 0. \end{aligned}$$

Then $I(T_1, T_2) = 2$, $I(T_2, T_1) = 1$.

For finite groups G, the identity (1.15) is, however, true. Indeed:

4.1.2. Exercise. Let $T_1, \dots, T_k$ be the full set of inequivalent irreducible representations of the finite group G. For any T, set $m_i(T) = I(T, T_i)$ Prove that

$$I(T, T') = \sum m_i(T) m_i(T').$$

4.1.3. Remarks. 1) Therefore, the bilinear form $I(\cdot, \cdot)$ determines something like an inner product on the space of representations with irreducible representations as an orthonormal basis.

2) For certain types of "geometric" representations, most often encountered in applications, it is possible to compute $I(T, T')$ without any knowledge of "coordinates" $m_i(T), m_i(T')$ and thus obtain some information concerning these coordinates.

4.2. The simplest type of representations. Let X be a finite set on which G acts, i.e., there is a map $a\colon G \times X \longrightarrow X$ such that:

$$a(g)\colon x \mapsto gx, \quad a(g_1 g_2)x = a(g_1)(a(g_2)x).$$

The action a gives rise to the representation a^* in the space $L(X)$ of functions on X:

$$(a^*(g)f)(x) = f(a(g^{-1})x).$$

Example (already considered) $X = G$ and $a(g) = R_g$ (or $a(g) = L_g$) the right (or left) translations by g.

Let a_1 and a_2 be actions of G on X and Y; let a_1^* and a_2^* be the induced representations in $L(X)$ and $L(Y)$, respectively. Let us try to compute $I(L(X), L(Y))$.

Clearly, the δ-functions δ_x and δ_y defined in $L(X)$ and $L(Y)$, respectively, form bases of these spaces. Then, clearly,

$$a_1^*(g)\delta_x = \delta_{gx}, \quad a_2^*(g)\delta_y = \delta_{gy}.$$

For an intertwining operator $C\colon L(X) \longrightarrow L(Y)$, denote by $c(x, y)$ its matrix coefficients relative to the bases δ_x, δ_y:

$$C\delta_x = \sum_{y \in Y} c(x, y)\delta_y \text{ for any } x \in X$$

The condition $Ca_1^* = a_2^* C$ takes the form:

$$Ca_1^*\delta_x = a_2^* C\delta_x$$

or (the tag displayed refers to the boxed formula)

$$\begin{array}{l} C\delta_{gx} = a_2^*(g) \sum c(x, y)\delta_y = \sum c(x, y)\delta_{gy} \\ \| \\ \sum c(gx, y)\delta y \Longleftrightarrow \boxed{c(gx, gy) = c(x, y)}. \end{array} \tag{1.16}$$

In other words, to find $I(a_1^*, a_2^*)$, it suffices to compute the dimension of the space Ω of functions $c(x,y)$ on $X \times Y$ satisfying (1.16). Since the G-action on $X \times Y$ splits into separate G-orbits, eq. (1.16) implies that the function $c(\cdot,\cdot)$ is constant on each orbit, and the characteristic functions $c_i(\cdot,\cdot)$ of orbits form a basis of Ω. Thus,

$$I(a_1^*, a_2^*) = \#(G\text{-orbits on } X \times Y).$$

4.3. Example. Let G be the group of cube's rotations (without inversions). Set

$$X_1 = \{\text{cube's vertices}\} \Longrightarrow \dim a_1^* = 8,$$
$$X_2 = \{\text{cube's faces}\} \Longrightarrow \dim a_2^* = 6.$$

According to the above, $\dim I(a_1^*, a_2^*) = \#(\text{ orbits in } X_1 \times X_2)$ and

$$(V_1, F_1) \sim (V_2, F_2) \text{ if a rotation sends } V_1 \text{ to } V_2 \text{ and } F_1 \text{ to } F_2.$$

Consider $\Omega_1 = \{(V,F) \mid V \in F\}$ and $\Omega_2 = \{(V,F) \mid V \notin F\}$.

4.3.1. Exercise. Prove that Ω_1 and Ω_2 are the only two orbits in $X_1 \times X_2$.

4.3.2. Corollary. $I(a_1^*, a_2^*) = 2$.

Moreover, in the space of intertwining operators, an explicit basis can be produced:

$$C_1 \colon f(V) \mapsto \varphi(F) = \sum_{V \in F} f(V),$$
$$C_2 \colon f(V) \mapsto \varphi(F) = \sum_{V \notin F} f(V).$$

The following basis is somewhat more convenient:

$$C_1' = C_1 + C_2 \colon f(V) \mapsto \varphi(F) = \sum f(V)$$
$$C_2' = C_1 - C_2 \colon f(V) \mapsto \varphi(F) = \sum_{V \in F} f(V) - \sum_{V \notin F} f(V).$$

The image of C_1' is the space of constants; the image of C_2' is the space of functions whose values on the opposite faces are opposite.

4.4. A method for description of irreducible representations of G. Find as many X_i on which G acts as possible and compute $I(a_i^*, a_j^*)$.

Example. G is the group of rotations of the cube. Consider:

$X_1 =$ the cube's vertices,

$X_2 =$ the cube's faces,

$X_3 =$ the big diagonals of the cube,

$X_4 =$ the regular tetrahedra inscribed in the cube,

$X_5 =$ the center of the cube.

So, $\dim X_i = 8, 6, 4, 2, 1$, respectively.

4.4.1. Exercise. Prove that $I(a_i^*, a_j^*)$ are of the form:

	a_1^*	a_2^*	a_3^*	a_4^*	a_5^*
a_1^*	4	2	2	2	1
a_2^*	2	3	1	1	1
a_3^*	2	1	2	1	1
a_4^*	2	1	1	2	1
a_5^*	1	1	1	1	1

Consider the main diagonal. Recall that

$$I(T,T) = \sum m_i(T)^2 \text{ for the set of all irreducible representations } T_i (= a_i^*)$$

Observe that T_5 is irreducible, T_3 and T_4 split into two inequivalent irreducible representations, T_2 is the direct sum of 3 inequivalent irreducible representations.

For T_1, we have an alternative: Either T_1 is the direct sum of four pair-wise inequivalent irreducible representations or the direct sum of two equivalent irreducible representations. In the second case, $I(T_1, T)$ is even for any T, while actually $I(T_1, T_5) = 1$; hence, T_1 is the direct sum of four inequivalent irreducible representations.

Further, T_5 enters each T_i with multiplicity 1. Let $T_i = T_i' \oplus T_5$. Then

	T_1'	T_2'	T_3'	T_4'
T_1'	3	1	1	1
T_2'	1	2	0	0
T_3'	1	0	1	0
T_4'	1	0	0	1

We see that T_1' consists of $T_3' \oplus T_4' \oplus T_2'$ and T_2'' is the direct sum of two irreducible representations distinct from T_3' and T_4'; call them T_2'' and T_2'''. The last two lines of the table imply that T_3' and T_4' are irreducible representations. Finally,

$$\begin{aligned}
&T_1 = A \oplus B \oplus C \oplus D,\\
&T_2 = A \oplus B \oplus E \text{ with } \dim E = 2, \dim B = 3,\\
&T_3 = A \oplus C \text{ with } \dim C = 3,\\
&T_4 = A \oplus D \text{ with } \dim D = 1,\\
&T_5 = A.
\end{aligned}$$

Since $1^2 + 1^2 + 2^2 + 3^2 + 3^2 = 24 = \#(G)$, we are done (by Burnside's theorem). (We could have obtained the same with the help of Maschke's theorem.)

4.4.2. Exercise. The group G has 5 conjugacy classes:

e;
the rotations about the vertex;
the rotations about the midpoint of the edge;
the two types of rotations about the center of the face: Through 90° and 180°.

4.4.3. It remains to describe the spaces in which irreducible representations act.

First, let us indicate them in $L(X_i)$.

Clearly, each $L(X_i)$ has an invariant irreducible representation of constants. In it, the representation A acts.

Further, among the intertwining operators $C\colon T_i \longrightarrow T_i$, there is the central symmetry operator, and hence, each $L(X_i)$ splits into the direct sum of invariant subspaces of symmetric and anti-symmetric functions. The former is the direct sum of the space of constants and its orthogonal complement, if any. For $i = 2, 3, 4$ (and 5) this suffices.

For $i = 1$, the space of anti-symmetric functions is reducible. It contains the 1-dimensional subspace generated by the function Sign that takes only values ± 1 and such that to neighboring vertices the opposite signs correspond. Thus,

$$\begin{aligned}
L(X_1) &= K \oplus L^\circ_S(X_1) \oplus L^\circ_a(X_1) \oplus L_{\text{Sign}},\\
L(X_2) &= K \oplus L^\circ_S(X_2) \oplus L_a(X_2),\\
L(X_3) &= K \oplus L^\circ_S(X_2),\\
L(X_4) &= K \oplus L_a(X_4),\\
L(X_5) &= K.
\end{aligned}$$

Moreover,

$$\begin{aligned}
&A \text{ acts in } K,\\
&B \text{ acts in } L_a(X_2) \quad \text{and } L^\circ_a(X_1),\\
&C \text{ acts in } L_S(X_3) \quad \text{and } L_S(X_1),\\
&D \text{ acts in } L_{\text{Sign}} \quad \text{and } L_a(X_4),\\
&E \text{ acts in } L_S(X_2).
\end{aligned}$$

Lecture 5. On representations of S_n

Let $X = \{1, \dots, n\}$. The following facts are assumed to be known:

1) The derived group of S_n is denoted $A_n = [S_n, S_n]$. This is the group generated by commutators $aba^{-1}b^{-1}$ for any $a, b \in S_n$. It is a normal subgroup of S_n. Its index is equal to 2.

2) Each permutation $g \in S_n$ can be represented as the product of cycles. Hence, X can be split into the disjoint union $X = X_1 \coprod X_2 \coprod \dots$ with g acting cyclicly on each X_i.

3) The elements $g, h \in S_n$ are conjugate if and only if, in their factorizations into products of cycles, these cycles are of equal lengths. Therefore, the number of classes of conjugate elements is equal to the number of representations of n as the sum of integers, e.g.,

$$3 = 1+1+1 = 2+1, \quad 4 = 1+1+1+1 = 3+1 = 2+2, \text{ and so on.}$$

We will classify irreducible representations of S_n by the same method that we used to describe irreducible representations of the group of rotation of the cube.

Let $\alpha = (n_1, \dots, n_k)$, where

$$n_1 \geq n_2 \geq \dots \geq n_k \quad \text{and} \quad n_1 + n_2 + \dots + n_k = n. \tag{1.17}$$

Clearly, the number of α's satisfying eq. (1.17) is equal to the number of conjugacy classes in S_n. For each α, consider the set X_α whose elements are partitions of the set $\{1, \dots, n\}$ into k groups of cardinality $n_1, \dots, n_k$.

Example: $n = 3$

$$X_{(3)} = (1,2,3), \quad X_{(2,1)} = \{(1,2)(3),\ (2,3)(1),\ (1,3)(2)\},$$
$$X_{(1,1,1)} = \{(1)(2)(3),\ (1)(3)(2),\ (2)(1)(3),\ (2)(3)(1),\ (3)(1)(2),\ (3)(2)(1)\}.$$

5.1. Exercise. Prove that $\#(X_\alpha)$, where $\alpha = (n_1, \dots, n_k)$, is equal to

$$\frac{n!}{n_1! \dots n_k!}.$$

Hint. Prove that X_α is a homogeneous space of S_n and find the order of the stationary subgroup of any point of X_α.

Let $L(X_\alpha)$ be the space of functions with the usual S_n-action

$$(T_\alpha(g)f)(x) = f(g^{-1}x).$$

Let now Sign be equal to 1-dimensional representation of S_n given by the formula

$$\operatorname{Sign}(g) = \begin{cases} 1 & \text{if } g \in A_n, \\ -1 & \text{if } g \notin A_n. \end{cases}$$

Set $t'_\alpha = \operatorname{Sign} \circ T_\alpha$. In other words,

$$(T'_\alpha(g)f)(x) = \operatorname{Sign}(g) f(g^{-1}x).$$

Our main goal: Find $I(T_\alpha, T'_\beta)$ for all α, β.

Let us order collections (1.17) lexicographically. Set

$$\alpha^* = (n_1^*, \dots, n_l^*), \quad \text{where } n_i^* = \#(n_j \mid n_j \geq i).$$

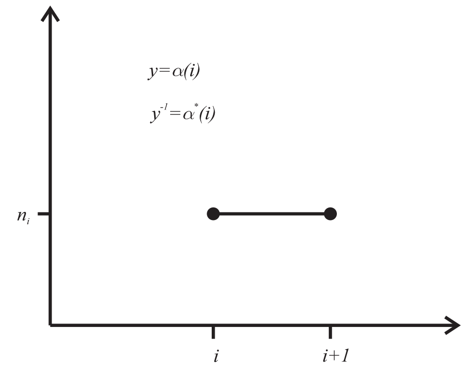

Fig. 1

5.2. Exercise. 1) α^* satisfies eq. (1.17).

2) $\alpha^{**} = \alpha$.

Hint: see Fig. 1

5.3. Lemma (J. von Neumann–H. Weyl)**.** 1) $I(T_\alpha, T'_\beta) \neq 0$ *only if* $\alpha \leq \beta^*$.

2) $I(T_{\beta^*}, T'_\beta) = 1$.

Proof. (1) Let $C\colon L(X_\alpha) \longrightarrow L(X_\beta)$ be an intertwining operator for T_α and T'_β. As in Lecture 4, define C by setting

$$(Cf)(y) = \sum_{x \in X_\alpha} c(x, y) f(x) \text{ for any } y \in X_\beta.$$

Since $CT_\alpha = T'\beta C$, we have $CT_\alpha f = T'_\beta Cf$ or

$$\sum_{x \in X_\alpha} c(x, y) f(g^{-1}x) = \operatorname{Sign}(g) \sum_{x \in X_\alpha} c(x, g^{-1}y) f(x) \text{ for any } y \in X_\beta. \quad (1.18)$$

Since eq. (1.18) holds for any f, it follows that

$$c(gx, gy) = \operatorname{Sign}(g) c(x, y) \text{ for any } x \in X_\alpha, y \in X_\beta, g \in S_n. \quad (1.19)$$

Now, recall that each $x \in X_\alpha$, where $\alpha = (n_1, \dots, n_k)$, is a partition of the set $\{1, \dots, n\}$ into k groups with n_1 numbers in the first group, and so on.

Similarly, for $y \in X_\beta$, where $\beta = (m_1, \dots, m_l)$.

Suppose there are two numbers, i and j, such that $1 \leq i, j \leq n$ that belong to one group in the partition x and to one group in the partition y.

5.3.1. Exercise. Does such a pair always exist?

Let $g \in S_n$ be the permutation of these two numbers and the identity on the other numbers. Then (the order of numbers inside the group is inessential)

$$gx = x \text{ and } gy = y.$$

5.3.2. Exercise. Moreover, $\mathrm{Sign}(g) = -1$.

Therefore, eq. (1.19) becomes

$$c(x, y) = -c(x, y) \Longleftrightarrow c(x, y) = 0. \tag{1.20}$$

Hence, the pairs that belong to one group in the partition x *must* belong to *different* groups in the partition y.

Let us show that this implies $\alpha \leq \beta^*$. Suppose $\alpha > \beta^*$. Then one of the following holds:

1) $n_1 > m_1^*$;
2) $n_1 = m_1^*, \quad n_2 > m_2^*$;
. . . .
i) $n_1 = m_1^*, \dots, n_{i-1} = m_{i-1}^*, n_i > m_i^*$.

Recall that m_1^* is the number of a group in the partition y. Hence, if $n_1 > m_1^*$, then the first group of the partition x has at least two elements that belong to one group of the partition y. Hence, case 1) can not take place.

Let $n_1 = m_1^*$. Then the number of elements in the first group of the partition $x = \#$ (groups of the partition y). Hence, each group of the partition y has exactly 1 element of the 1st group of the partition x.

Recall that m_2^* is the number of groups in the partition y with > 1 element. Since the elements of the 2nd group of the partition x should belong to different groups of the partition y, we see that $m_2^k \geq n_2$ contradicting the statement 2).

Similarly, each of the conditions i) leads to a contradiction. Heading 1) of Lemma is proved.

(2) Let now $\beta = \alpha^*$. Let us show that $I(T_\alpha, T'_{\alpha^*}) = 1$. Consider the Young tableau (YT)
Then

$$\alpha^* = (\text{the length of the 1st column}, \dots) = (n_1^*, \dots, n_l^*)$$

Placing the numbers $\{1, \dots n\}$ into cells, one number per cell we obtain a one-to-one correspondence:

$$YT \longleftrightarrow (x, y) \in X_\alpha \times X_{\alpha^*}.$$

Clearly, $g(YT) \longleftrightarrow (gx, gy)$.

Hence, all pairs (gx, gy) are distinct for distinct $g \in S_n$. Fix a Young tableau; let (x_0, y_0) be the corresponding pair. Define $c_0(x, y)$ as follows:

$$c_0(x, y) = \begin{cases} \mathrm{Sign}(g) & \text{if } x = gx_0, \ y = gy_0; \\ 0 & \text{otherwise.} \end{cases} \tag{1.21}$$

Clearly, the operator c given by (1.21) satisfies eq. (1.18), and therefore is an intertwining operator. It remains to prove that any other intertwining operator differs from c_0 by a scalar factor. Indeed,

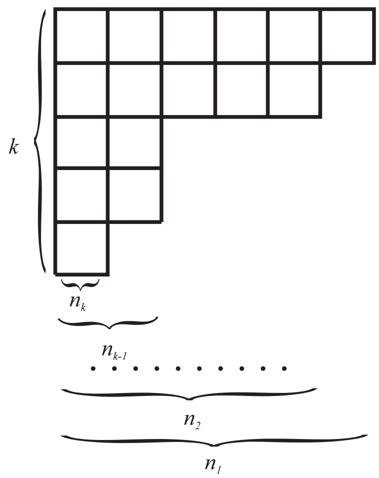

Fig. 2

> $c(x, y) \neq 0$ only if all numbers from one group in the partition x lie in different groups of the partition y.

Let us show that, under this condition, the pair (x, y) corresponds to the Young tableaux YT. To show this, place the numbers $1, \dots, n$ as follows: If the number belongs to the ith group of the partition x and the jth group of the partition y, let it occupy the (i, j)th slot of the Young tableaux YT.

5.3.3. Exercise. This way each slot gets exactly one number.

Finally: $c(x, y) \neq 0$ if and only if (x, y) is obtained from a Young tableaux YT, and hence is of the form (gx_0, gy_0). But then eq. (1.18) implies that

$$c(x, y) = c(x_0, y_0) \cdot c_0(x, y). \qquad \square$$

Heading 2) of Lemma 5.3 implies that T_α and T'_{α^*} have exactly one irreducible representation in common, call it U_α.

5.3.4. Theorem. *Irreducible representations U_α exhaust all distinct irreducible representations of S_n.*

Proof. First of all, we see that

$$\#\{U_\alpha \mid \text{ all } \alpha' s\} = \# \text{ (conjugacy classes of } S_n).$$

Hence, it suffices to prove that $U_\alpha \not\cong U_\beta$ for different α, β. For definiteness sake, let $\alpha > \beta$. Then

— on the one hand, $I(T_\alpha, T'_{\beta^*}) \geq 1$, since both representations have $U_\alpha \simeq U_\beta$ as a subrepresentation.

— on the other hand, heading (1) of Lemma 5.3 states that $I(T_\alpha, T'_\beta) = 0$ for $\alpha > \beta$. This is a contradiction. $\square$

Lecture 6. Invariant integration on compact groups

6.1. Recapitulation. For finite groups, we have established the following facts:

1) Every representation is equivalent to a unitary one.
2) Every representation is a direct sum of irreducible finite dimensional representations.
3) The matrix elements of irreducible representations form an orthogonal basis in the space of all functions on the group.

The proof of these facts was based on the existence of an invariant mean of every function. More exactly, to every function f on the finite group G, we can assign its mean

$$I(f) = \frac{1}{N}\sum_{g\in G} f(g), \text{ where } N = \#(G).$$

This mean is invariant with respect to the left and right translations by $h \in G$:

$$I(f_r) = I(f_l) = I(f), \text{ where } f_r(g) = f(gh), f_l(g) = f(hg).$$

It turns out that if G is compact, the invariant mean of each integrable function can be defined; hence, the above properties 1)–3) hold for compact groups as well.

First, consider an example, then a general theory.

6.2. Example. $G = S^1$, where multiplication is the addition of arguments of complex numbers $z = e^{i\varphi}$. We will try to "approximate" S^1 by a sequence of finite groups on which the invariant mean exists and hope that the limit exists. Let us see. Consider the groups

$$\mu_n = \{z \mid z^n = 1\} = \left\{e^{\frac{2\pi ik}{n}} \mid k = 0, 1, \dots, n-1\right\}.$$

Therefore, an "approximate" mean on S is the mean on μ_n:

$$I_n(f) = \frac{1}{n}\sum_{k=0}^{n-1} f(\varepsilon^k), \text{ where } \varepsilon = e^{\frac{2\pi i}{n}}.$$

Clearly, $I_n(f)$ is a Riemannian sum and for any integrable (e.g., continuous, or whatever) function f, there exist the limit

$$I(f) = \lim_{n\longrightarrow\infty} I_n(f) = \frac{1}{2\pi}\int_0^{2\pi} f(e^{i\varphi})d\varphi$$

which is the mean desired. Unfortunately, for some groups, there is no good approximation by finite groups (or, at least, no known such approximation). So we have to work out another method.

What are the properties of the invariant mean that we need?

a) $f \mapsto I(f)$ is a linear functional, i.e.,

$$I(af + bg) = aI(f) + bI(g) \text{ for any } a, b \in \mathbb{K}.$$

b) If $f \geq 0$, then $I(f) \geq 0$.

c) Set: $f_r(g) = f(gh)$, and $f_l(g) = f(hg)$. Then

$$I(f_l) = I(f_r) = I(f) \quad \text{for any } h \in G.$$

d) If $f \equiv 1$, then $I(f) \equiv 1$.

Besides, we would like to consider the mean on a sufficiently broad class of functions (e.g., on a class that contains all continuous functions).

Observe, that b) and a) imply that

$$\sup_{g \in G} f(g) \geq I(f) \geq \inf_{g \in G} f(g).$$

Let us show that properties a)–d) uniquely determine the value of $I(f)$, at least on continuous functions.

Let S be any finite subset of G. Define the left and right averaging operators by setting

$$(L_S f)(g) = \frac{1}{\#(S)} \sum_{s \in S} f(sg),$$

$$(R_S f)(g) = \frac{1}{\#(S)} \sum_{s \in S} f(gs).$$

Clearly, for any finite S, we have

$$I(L_S(f)) = I(R_S(f)) = I(f).$$

Suppose that for a given function f, we managed to find a set $S \subset G$ such that $L_S(f)$ is almost constant on G: For any given $\varepsilon > 0$, we have

$$|(L_S f)(g) - c| < \varepsilon \text{ for any } g \in G.$$

Since $I(f) = I(L_S(f))$, we can compute $I(f)$ to within ε making use of properties a)–d) only. Therefore, two means of one function f differ not more than by 2ε. If ε can be taken arbitrary, then the different means $I(f)$ and $I'(f)$ coincide proving the uniqueness of the mean.

Thus, for a given function f, we have to show that there exists a sequence of subsets $S_n \subset G$ such that the functions $f_n = L_{S_n}(f)$ uniformly converge to a constant. We split the proof into several steps.

6.3. Exercises

6.3.1. Exercise. Recall that the *oscillation* of a real-valued function f is

$$\omega(f) = \max_{g\in G} f(g) - \min_{g\in G} f(g).$$

Prove that if $\omega(f) > 0$ (i.e., f is not a constant), then there exists a finite set S such that $\omega(L_S(f)) < \omega(f)$.

Hint. Let $M = \max\limits_{g\in G} f(g)$ and $m = \min\limits_{g\in G} f(g)$. Consider the compact set

$$\mathfrak{F} = \left\{g \in G \mid f(g) \geq \frac{M+m}{2}\right\},$$

an open set

$$\mathcal{O} = \left\{g \in G \mid f(g) < \frac{M+m}{2}\right\}$$

and a finite covering of $\mathcal{F}$ by translations of $\mathcal{O}$.

Now, consider the collection $C(G)$ of all continuous functions on G. In $C(G)$, define the distance by setting

$$\rho(f_1, f_2) = \max_{g\in G} |f_1(g) - f_2(g)|$$

6.3.2. Exercise. Let f be a continuous function on a compact group G; denote:

$$A = \{f_h \mid f_h(g) = f(hg) \text{ for any} h \in G\}.$$

Prove that A is a compact subset in $C(G)$.

Hint. 1st method. Prove that f is uniformly continuous on G, i.e., for any $\varepsilon > 0$, there exists a neighborhood V of $e \in G$ such that if $g_1 g_2^{-1} \in V$, then $|f(g_1) - f(g_2)| < \varepsilon$. Now, construct a finite ε-net [2)] for A.

2nd method. Prove that the map $G \longrightarrow C(G)$ given by the formula $h \mapsto f_h$ is continuous for any continuous f. Therefore, A is the image of a compact under a continuous map, and hence A is a compact.

6.3.3. Exercise. Let f be a continuous function on G, and

$$B_f = \{g \in C(G) \mid g = L_S(f) \text{ for some } S\}.$$

Prove that B_f is a compact subset in $C(G)$.

[2] Let P be a probability distribution over some set A. An ε-*net* for a class $H \subset 2^A$ of subsets of A is any subset $S \subset A$ such that, for any $h \in H$, we have

$$p(h) \geq \varepsilon \Longrightarrow h \cap S \neq \emptyset.$$

(Intuitively S approximates a probability distribution.)

Hint. 1st method. Prove that all functions from B_f are uniformly continuous and uniformly bounded.

2nd method. Prove that if A is a compact set in a linear metric space, then its convex envelope is also a compact.

6.3.4. Exercise. Let f be a continuous function on a compact group G. Then

$$\inf_S \omega(L_S(f)) = 0.$$

Hint. Let $\inf\limits_S \omega(L_S(f)) = \omega_0 > 0$. Then there exists a sequence S_n of finite subsets of G such that $\lim\limits_{n\longrightarrow\infty} \omega(L_{S_n}(f)) = \omega_0$. Since the set B_f of all functions of the form $L_S(f)$ is compact (Exercise 6.3.4), there exists a subsequence of the sequence $L_{S_n}(f)$ that uniformly converges to a function $f_0 \in C(G)$.

Prove then that $\omega(f_0) = \omega_0$. Since $\omega_0 \neq 0$, it follows that f_0 is not a constant and there exists a finite set $T \subset G$ such that $\omega(L_T(f_0)) < \omega_0$. But then for sufficiently large n we have $\omega(L_T L_{S_n}(f)) < \omega_0$. This is a contradiction.

Exercises 6.3.1–6.3.4 easily imply the following statement (uniqueness of the invariant mean).

6.4. Lemma. *For any continuous function f on a given compact group G, there exists a sequence of finite subsets $S_n \subset G$ such that the functions $L_{S_n}(f)$ converge uniformly to a constant C.*

Clearly, if an invariant mean of a function f on G exists, it should be equal to $C = \lim\limits_{n\longrightarrow\infty} L_{S_n} f$. So the uniqueness is proved.

Observe that in our proof of uniqueness of the invariant mean we only used half of the property c): Since we only applied to f the left averaging operators, it suffices to assume the invariance of the mean with respect to left translations only.

6.5. Proof of existence of the invariant mean. One might think that for the mean $I(f)$ one can always take the constant C from Lemma 6.4. But suppose that for the same function f there exist two different sequences such that

$$\lim_{n\longrightarrow\infty} L_{S_n}(f) = c, \qquad \lim_{n\longrightarrow\infty} L_{S'_n}(f) = c'.$$

Clearly, in this case there is no invariant mean of f on G.

Let us show that in reality an invariant mean always exists. Here the right averaging operators become handy. Obviously, Lemma 6.4 holds for the right translations as well. Therefore, for any (continuous) function f, there exists a sequence of finite sets $T_n \subset G$ such that $\lim\limits_{n\longrightarrow\infty} R_{T_n}(f) = \widetilde{C}$.

We will not prove that all left means coincide. Instead, we will prove that any left mean is equal to any right mean.

6.5.1. Exercise. Prove that if $\lim\limits_{n\longrightarrow\infty} L_{S_n}(f) = c_l$ and $\lim\limits_{n\longrightarrow\infty} R_{T_n} = c_r$ uniformly, then $c_l = c_r$.

Hint. Show that L_S and R_T commute, i.e., $L_S R_T f = R_T L_S f$ and consider $\lim\limits_{\substack{n\longrightarrow\infty \\ m\longrightarrow\infty}} L_{S_m} R_{T_n}(f)$.

Summary. We have proved the existence of an invariant mean of the (continuous) function on the compact group and that the mean is invariant. Moreover, the "left mean", i.e., a quantity $I_l(f)$ such that

$$I_l(f_l) = I_l(f) \text{ for any } f_l(g) = f(hg) \text{ and any } h \in G$$

is automatically the "right mean", i.e., it satisfies

$$I_r(f_r) = I_r(f) \text{ for any } f_r(g) = f(gh) \text{ and any } h \in G.$$

6.5.2. Exercise. Let $f \in C(G)$ and $\widetilde{f}(g) = f(g^{-1})$. Prove that $I(\widehat{f}) = I(f)$.

Hint. $\widetilde{I}(f) = I(\widetilde{f})$ is an invariant mean of f, but we have uniqueness.

6.5.3. Exercise. Let G be the group of all isometric transformations of the circle. Each element of G is either a rotation r_φ through the angle φ, where $0 \leq \varphi \leq 2\pi$, or a reflection R_ψ in the diameter that forms an angle of $\dfrac{\psi}{2}$, where $0 \leq \psi \leq 2\pi$, with a fixed direction. Prove that the invariant mean on G can be given by the formula

$$I(f) = \frac{1}{4\pi}\left(\int_0^{2\pi} f(r_\varphi)d\varphi + \int_0^{2\pi} f(R_\psi)d\psi\right).$$

In what follows we will write $\int_G f(g)dg$ *instead of* $I(f)$.

6.6. Unitary representations of compact groups G. To fully use topology of G, e.g., the fact that G is compact, we assume that all G-modules (the spaces of representations) are endowed with a topology. We will demand that the assignment $g \mapsto T(g) \in GL(V)$ is continuous. This demand is nontrivial, since the space of linear operators on a given V can be rigged with *several* structures of a topological space and we have to select one of them.

We will assume that V is a normed linear space and the assignment $g \mapsto T(g)$ is continuous with respect to the *strong operator topology.*

Let us construct an invariant (with respect to G) inner product on V. To this end, suppose, first, that V^* contains a countable (infinite) set of linear functionals $\varphi_k \in V^*$, where $k \in \mathbb{N}$, sufficient to distinguish any two distinct elements from V. Clearly, without loss of generality, we may assume that $\| \varphi_k \| = 1$ for all $k \in \mathbb{N}$. Then

$$\{x, y\} = \sum \frac{1}{2^k}\varphi_k(x)\overline{\varphi_k(y)}$$

is an inner product in V (an Hermitian one). To obtain a G-invariant inner product, apply averaging:

$$(x, y) = \int_G (\{T(g)x, T(g)y\})dg,$$

Indeed, we only have to verify positivity of (x, x) for $x \neq 0$.

6.6.1. Exercise. Prove that if a continuous function f is ≥ 0 on G and $\neq 0$ at least at one point, then

$$\int_G f(g)dg > 0.$$

Now observe that we have realized our initial space V as a dense subset of a space $\widetilde{V}$ with an Hermitian inner product. Indeed, $\widetilde{V}$ is the completion of V with respect to the invariant inner product constructed. This construction is also applicable in the absence of the sequence of functionals that separate points of V. But then it leads to a degenerate inner product, i.e., $(x, x) = 0$ does not imply any more that $x = 0$. Hence, we thus embed into the Hilbert space not V, but a quotient space of V (modulo $x \neq 0$ such that $(x, x) = 0$).

6.7. The two problems. To what extent can the study of arbitrary representations be reduced to the study of irreducible representations? What does it mean that an infinite dimensional space is irreducible?

To the second question we have an answer:

1) Algebraic irreducibility: V has no non-trivial ($\neq \{0\}, V$) invariant subspaces.

2) Topological irreducibility: V has no nontrivial CLOSED invariant subspaces.

Clearly, for finite dimensional spaces, when all subspaces are closed, these notions coincide.

6.7.1. Theorem. *Every topologically irreducible representation T of the compact group G in a Hilbert space H is finite dimensional (hence, algebraically irreducible).*

6.7.2. Theorem. *Let T be a unitary representation of a given compact group G in a Hilbert space H. For T to be topologically irreducible, it is necessary and sufficient that any continuous operator $H \longrightarrow H$ commuting with T were a scalar one.*

Lecture 7. Algebraic groups and Lie groups (after É. Vinberg)

7.1. A subgroup of $GL(V)$ consisting of linear transformations is called a *linear*[3)] group. Having selected a basis in V we make $GL(V)$ into the matrix group $GL(n)$, where $n = \dim V$, and realize linear groups by matrices.

[3] The term indicates the realization currently used: "linear" groups may have non-linear representations (on coset spaces). "Linear" groups can be realized, for example, in terms of vector fields.

7.1.1. Example. The *classical groups* are the groups of the following examples 1)–4) and the real forms of the corresponding complex groups.

1) $GL(n;\mathbb{C})$, $GL(n;\mathbb{R})$, $GL(n;\mathbb{Z})$, $GL(n;\mathbb{K})$, where $\mathbb{K}=\mathbb{Z}_p$, $\mathbb{Z}/p$, $\mathbb{F}_q$, where $q=p^n$, ...

2) $SL(n;\mathbb{K})$, where $\mathbb{K}$ is as above;

The following examples are groups over $\mathbb{R}$:

3) $O_B(n)$ and $SO_B(n)=O_B(n)\cap SL(n;\mathbb{R})$, the *orthogonal groups* preserve the non-degenerate symmetric bilinear form B with matrix reduced to the canonical form (usually, 1_n or $\mathit{\Pi}_n=\operatorname{antidiag}_n(1,\dots,1)$, depending on the problem);

$O(p,q)$, the *pseudo-orthogonal group* preserves the bilinear form B with matrix $\operatorname{diag}(1_p,-1_q)$;

4) $Sp(2n)$, the *symplectic group* preserves the non-degenerate anti-symmetric bilinear form B with matrix J_{2n};

5) $U(p,q)$, the *pseudo-unitary group* preserves the sesquilinear form with matrix $\operatorname{diag}(1_p,-1_q)$. The the *unitary group*, $U(n)$, preserves the sesquilinear form with matrix 1_n.

The unitarity condition reads

$$\sum_k a_{ik}\overline{a_{jk}}=\delta_{ij}$$

Set $SU(n)=U(n)\cap SL(n;\mathbb{C})$.

6) The group $\left\{\begin{pmatrix}\lambda & 0\\ 0 & \lambda^a\end{pmatrix}\mid\lambda\in\mathbb{R}_+\right\}$, where $a\in\mathbb{R}$ is fixed.

A linear group is said to be *algebraic* if it is singled out by a system of *polynomial* equations. For example:

$$\begin{aligned}&GL(n;\mathbb{K})=\mathbb{K}[x_{ij},T]/(T\det(x_{ij})-1),\quad\text{where } 1\le i,j\le n;\\ &SL(n;\mathbb{K})=\mathbb{K}[x_{ij}]/(\det(x_{ij})-1);\\ &O(n;\mathbb{K})=\mathbb{K}[x_{ij}]/(\sum_k x_{ik}x_{jk}=\delta_{ij}\text{ for } 1\le i,j\le n).\end{aligned}$$

A linear group $G\subset GL(n;\mathbb{R})$ is called a d-dimensional *Lie* group if there is a neighborhood $U\subset\operatorname{Mat}(n;\mathbb{R})$ of the unit of G such that $G\cap U$ can be given parametrically as

$$\begin{aligned}&a_{ij}=a_{ij}(u_1,\dots,u_d),\quad\text{where } a_{ij}(u)\in C^\infty(V)\\ &\text{for }\ V\subset\mathbb{R}^d\text{ and }\operatorname{rk}\left(\frac{\partial a_{ij}}{\partial u_k}\right)=d\text{ on }V.\end{aligned}\tag{1.22}$$

7.1.2. Exercise. 1) Any finite group is algebraic.

2) The groups 1)–5) over $\mathbb{R},\mathbb{C}$ are algebraic except $GL(n;\mathbb{Z})$.

3) The group of example 6) is not algebraic for any $a\in\mathbb{R}$ (for different reasons for $a\in\mathbb{Q}$ and $a\neq\mathbb{Q}$).

4) Find $\dim G$ for G from examples 1)–6).

Let G be a linear group, $G \subset GL(n; \mathbb{R})$. The map

$$X \mapsto AX, \quad \text{where } X \in \operatorname{Mat}(n; \mathbb{R}) \text{ and } A \in G \text{ is fixed,}$$

sends G into itself, $I \mapsto A$. Therefore, as a set in $\operatorname{Mat}(n; \mathbb{R})$, the group G is "homogeneous": The m-neighborhoods of all points are all "alike". So, if G is a Lie group, then the eq. (1.22) holds not only for a neighborhood U of the unit $e \in G$ but for a neighborhood of any point $g \in G$.

7.1.3. Exercise. The groups 1)–5) are Lie groups. For which a is the group 6) a Lie group?

7.2. Theorem. *Any algebraic linear group is a Lie group.*

7.3. Topological properties of the groups. Since any linear group G is a subset in $\mathbb{R}^{n^2} = \operatorname{Mat}(n; \mathbb{R})$, it can be endowed with topological properties, such as compactness.

7.3.1. Example. The groups $O(n)$, $SO(n)$ are compact (try to prove it!); any finite group is, clearly, compact. The unitary group $U(n)$ is compact, the pseudo-unitary one, $U(p, q)$ with $pq \neq 0$, is not.

A linear group is said to be *connected* if any two of its points may be connected by a continuous curve completely lying in G. (For a connected group G, any two points can be connected by a *smooth* curve lying in G.) The set

$$\{G^A \mid g \in G \text{ that can be connected with } A \in G\}$$

is called the *connected component* of an element $A \in G$.

7.4. Theorem. *The set G^E for the unit $E \in G$ is a normal subgroup in the linear group G. Moreover, $G^A = A \cdot G^E$ for any $A \in G$.*

Proof. Let $A, B \in G^E$, let $A(t), B(t) \in G$ be continuous curves that connect E with A and B, respectively, as t varies from 0 to 1. The curve

$$C(t) = \begin{cases} A(t) & \text{if } 0 \leq t \leq 1, \\ AB(t-1) & \text{if } 1 \leq t \leq 2 \end{cases}$$

connects E with AB; the curve $A(t)^{-1}$ connects E with A^{-1}. If $P \in G$, then the curve $PA(t)P^{-1}$ connects E with PAP^{-1}. Hence, G^E is a normal subgroup.

The "moreover" part: Since the map $X \mapsto AX$ is one-to-one and continuous together with its inverse, it sends G^E to G^A. $\square$

7.4.1. Example. $O(n) \supset SO(n)$.

7.4.2. Lemma. *In an appropriate basis, any orthogonal transformation* $A \in O(n;\mathbb{R})$ *is of the form*

$$\operatorname{diag}(\pi(\varphi_1), \dots, \pi(\varphi_m), 1, -1), \tag{1.23}$$

where $\pi(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$ *and any of the last two elements (i.e.,* 1 *and* -1*) may be absent. In particular, if* $\det A = 1$*, then* -1 *is certainly absent.*

In other words, for any $A \in O(n)$, there exists $C \in O(n)$ such that CAC^{-1} is of the form (1.23).

Proof. Linear algebra, see, e.g., [P]. □

If the matrix (1.23) has no -1, consider the continuous curve connecting the matrix (1.23) with the unit matrix:

$$\operatorname{diag}(\pi(t\varphi_1), \dots, \pi(t\varphi_m), 1), \quad 0 \leq t \leq 1.$$

Hence, $SO(n)$ is a connected group.

7.5. If the connected component of $E \in G$ consists of E only, then G is called *totally disconnected*.

7.5.1. Examples. $GL(n;\mathbb{Q})$, all discrete groups (like $\mathbb{Z}$, $GL_n(\mathbb{Z})$), and all finite groups are totally disconnected.

7.5.2. Theorem. *Any totally disconnected normal subgroup* N *of a connected linear group* G *belongs to the center of* G.

Proof. Let $C \in N$, let $A(t)$ for $0 \leq t \leq 1$ be a continuous curve in G. Then

$$A(t)CA(t)^{-1} \in N^C = \{C\},$$

and hence $A(t)C = CA(t)$. Since G is connected, we see that C commutes with all elements of G. □

A connected group G is called *simply connected* if any closed curve (loop) can be continuously contracted into a point (in other words, as a topological space, G has no holes). We will show that $SO(n)$ is NOT simply connected for $n \geq 2$.

7.6. Unitary groups. Consider $\mathbb{C}^n$ as $\mathbb{R}^{2n}$: If the vectors $e_1, \dots, e_n$ span $\mathbb{C}^n$, then the vectors $e_1, \dots, e_n, ie_1, \dots, ie_n$ span $\mathbb{R}^{2n}$. If $A \in GL(n;\mathbb{C})$, then, in the above basis of $\mathbb{R}^{2n}$, the matrix of A considered as an element of $GL(2n;\mathbb{R})$ is of the form

$$\begin{pmatrix} \operatorname{Re} A & -\operatorname{Im} A \\ \operatorname{Im} A & \operatorname{Re} A \end{pmatrix}.$$

Is it clear that $U(n)$ and $SU(n)$ are compact? If not, prove that it is.

7.6.1. $U(n)$ and $SU(n)$ are connected.

Lemma. *Any unitary transformation of $\mathbb{C}^n$ has an orthonormal basis consisting of eigenvectors. The absolute value of any eigenvalue of any unitary transformation is equal to 1.*

Proof. (Linear algebra): For any $A = \operatorname{diag}(e^{i\varphi_1}, \dots, e^{i\varphi_n})$, consider the curve

$$\{A(t) = \operatorname{diag}(e^{i\varphi_i t}, \dots, e^{i\varphi_n t}) \mid 0 \leq t \leq 1\},$$

that connects the unit with A and lies inside $U(n)$. If we select $\varphi_1, \dots, \varphi_n$ so that $\varphi_1 + \dots + \varphi_n = 0$ (not just $2\pi k$), then $\det A(t) = 1$ for any t. $\square$

7.7. Invariants of compact linear groups. Let $S = \mathbb{R}[x]$, where $x \in \mathbb{R}^n$. Every invertible linear map $A\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ induces an automorphism of S:

$$(Af)(x) = f(A^{-1}x), \quad \text{for any } x \in \mathbb{R}^n.$$

7.7.1. Lemma (On an averaged polynomial)**.** *Let $G \subset GL(n, \mathbb{R})$ be a compact group. Any nonempty convex subset $S_+ \subset S$ contains a G-invariant polynomial if S_+ is G-invariant.*

Proof. Take any $f \in S_+$ and apply to it all transformations from G; let Gf be the set obtained. Its linear envelope S_l is of finite dimension because $\dim f_l \leq \dim f$ for any $f_l \subset S_l$. Since G is compact, Gf is bounded in S_l. Let $M \subset S_l$ be the convex envelope of Gf, let f_0 be the center of mass of M (we naturally assume that the mass of the body is proportional to its volume). Clearly, $f_0 \in M \subset S_+$. The polynomial f_0 is G-invariant as the center of mass of a G-invariant set M. $\square$

7.7.2. Example. Let

$$S_+ = \{\text{positive definite quadratic forms }\}.$$

For any compact linear group $\subset GL(n, \mathbb{R})$, there exists a positive definite quadratic form which is G-invariant.

In other words, for a suitable inner product preserved by $O(n)$, any compact linear group G may be considered as a subgroup of $O(n)$.

Over $\mathbb{C}$: Let

$$S_+ = \{\text{positive definite Hermitian forms }\}.$$

Then any compact linear group $\subset GL(n; \mathbb{C})$ is a subgroup of $U(n)$ (for a suitable Hermitian product).

7.7.3. Corollary. *For any G-invariant subspace $V \subset \mathbb{R}^n$ (or $\mathbb{C}^n$) of a compact linear group $G \subset GL(n, \mathbb{R})$ (or $GL(n, \mathbb{C})$) there exists an invariant complementary subspace W, i.e., $V \oplus W = \mathbb{R}^n$.*

Indeed, take $W = V^\perp$.

7.8. Theorem (Hilbert)**.** *The ring S^G is finitely generated.*

Lecture 8. Unitary representations of compact groups (Proof of Theorem 6.7.1)

The idea of the proof is very simple. Suppose we managed to prove the orthogonality relations obtained earlier for finite groups:

$$\int_G t_{ik}(g)\overline{t_{je}}(g)dg = \frac{1}{\dim H}\delta_{ij}\delta_{kl}. \tag{1.24}$$

Since, for $i = j$ and $k = l$, the left hand side is a positive number, not 0, it follows that $\dim H < \infty$.

Although the *idea* of the proof is easy, to *prove* eq. (1.24) is not easy, we need several auxiliary statements.

2) *Let $T\colon G \longrightarrow GL(H)$ be a unitary representation, H a Hilbert space. The representation T is topologically irreducible if and only if every continuous operator in H that commutes with all $T(g)$ is a scalar one.*

N e c e s s i t y. This is quite clear: Indeed, let the representation be reducible, i.e., let there exist a subspace $H_1 \subset H$ invariant with respect to $T(G)$. But then $H = H_1 \oplus H_1^\perp$. Hence, the projection onto H_1 parallel to $H_1^\perp$, clearly, commutes with every $T(g)$, where $g \in G$, and is not a scalar operator.

S u f f i c i e n c y[4)].

8.1. Exercise. Prove that if an operator A commutes with all $T(g)$, where $g \in G$, then $A = B + iC$, where $B = B^*$ and $C = C^*$ (self-adjoint operators) and both B and C commute with all $T(g)$, too.

Hint. Prove that A^* commutes with all $T(g)$, where $g \in G$. Therefore, in the proof of sufficiency, we can confine ourselves to self-adjoint operators.

For self-adjoint operators A and B, we will write $A \geq B$ if

$$(Ax, x) \geq (Bx, x) \quad \text{for any } x \in H.$$

8.2. Exercise. Let $aI \leq A \leq bI$ (here I is the identity operator) and let $P(x) \in \mathbb{R}[x]$ be a polynomial such that $P(x) \geq 0$ on $[a, b]$. Then $P(x) \geq 0$ on $\mathbb{R}$ as well.

Hint. Represent $P(x)$ in the form

$$P(x) = (x - a)Q(x) + (b - x)R(x) + S(x),$$

where Q, R and S are the sums of some polynomials squared.

8.3. Exercise. Let A be a self-adjoint operator and $0 \leq A \leq cI$. Prove that $\| A \| \leq c$.

---

[4] If one knows the spectral theory of operators acting in a Hilbert space, all is easy: Just apply the known theorems. Regrettably, I completely forgot these theorems. And the reader, probably, did not start to study them yet. So let us solve problems instead of taking a shortcut. Kirillov had split the task in digestible chunks.

Hint. Introduce a new inner product $[x, y] = (Ax, y)$ and apply the *Cauchy–Bunyakovsky–Schwarz*[5] *inequality*

$$[x, y]^2 \le [x, x]\ [y, y] \quad \text{for } y = Ax.$$

8.4. Exercise. Let $f(x)$ be any real function on $[a, b]$ and $P_n(x)$ any sequence of polynomials that converges to $f(x)$ uniformly on $[a, b]$. Prove that the sequence of operators $P_n(A)$ converges uniformly to an operator that does not depend on the choice of the sequence $P_n(x)$.

The operator determined by Exercise 8.4 will be denoted $f(A)$.

8.5. Exercise. Prove that a monotonic bounded sequence of operators in H converges in the strong operator topology. In other words, if

$$A_1 \le A_2 \le \dots \le A_n \le \dots \le B,$$

then there exists an operator A such that the sequence $A_n x$ converges to Ax for any $x \in H$.

Hint. The numerical sequence $(A_n x, x)$ is monotonic and bounded. Hence, it converges. Demonstrate that the sequence $A_n x$ is a fundamental one, i.e., for all $\varepsilon > 0$, there exists N such that $\| A_n x - A_m x \| < \varepsilon$ for any $n, m > N$. To this end, introduce the new inner product

$$[x, y] = ((A_n - A_m)x, y)$$

and apply the Cauchy–Bunyakovsky–Schwarz inequality.

8.6. Exercise (The spectral theorem)**.** Prove that, for any self-adjoint operator A, there exists a family of subspaces $H_c \subset H$ labeled by real numbers c that belong to a subset $R \subset \mathbb{R}$ and with the properties

1) $H_c \subset H_d$ if $c < d$.

2) $\bigcup\limits_{c \in R} H_c = H$, and $\bigcap\limits_{c \in R} H_c = \{0\}$.

3) *For $c < d$, denote the orthogonal complement to H_c in H_d by $H_{c,d}$. Then*

$$c(x, x) \le (Ax, x) \le d(x, x) \text{ for any } x \in H_{c,d}.$$

4) *The spaces H_c are invariant with respect to all operators that commute with A.*

Hint. Let χ be the characteristic function of the set $\{x \in \mathbb{R} \mid x < c\}$; let $f_n(x)$ be an monotonically growing sequence of continuous nonnegative functions converging to χ_c. Then $P_n = \lim\limits_{n \longrightarrow \infty} f_n(A)$ is the projection operator onto H_c.

---

[5] Named, in respective countries, after Augustin Louis Cauchy, Viktor Yakovlevich Bunyakovsky and Hermann Amandus Schwarz.

8.7. Exercise. If, for a self-adjoint operator A, all the spaces H_c coincide with either H or $\{0\}$, then A is a scalar operator.

Clearly, the results of Exercises 8.6 and 8.7 prove the sufficiency condition in Theorem 6.7.1. $\square$

8.8. Now, let A be an arbitrary bounded operator in a Hilbert space H. Let also H be the space of a unitary representation T of a compact group G. Set

$$\mathring{A} = \int_G T(g)AT(g^{-1})dg.$$

8.8.1. Exercise. Prove that $\mathring{A}$ commutes with all $T(g)$, $g \in G$.

Hint. Consider the Banach space V consisting of all bounded operators in H and the operator $S\colon A \mapsto T(g)AT(g^{-1})$ in V.

To solve this exercise (and not only for this), we have to integrate over the group not only the usual numerical functions, but also vector-valued or operator-valued functions. We will reduce such problems to the usual integral of the usual functions.

For any function f continuous on a compact group G and with values in a Banach space V and for any $\chi \in V^*$, construct the numerical function $g \mapsto \langle f(g), \chi\rangle$, where $\langle \cdot, \cdot \rangle$ is the pairing of V and V^*. Define $\int_G f(g)dg$ as the vector $v \in V$ such that

$$\langle v, \chi\rangle = \int_G \langle f(g), \chi\rangle dg \text{ for any } \chi \in V^*. \tag{1.25}$$

The condition (1.25) uniquely determines v, but it is unclear why such a v exists. To prove the existence, consider the collection

$$K = \{v \in V \mid v = \sum_{k=1}^{n} C_k f(g_k), \text{ where } C_k \geq 0, \sum_{k=1}^{n} C_k = 1 \text{ for some } n \in \mathbb{N}\},$$

and let $\overline{K}$ be the closure of K.

8.8.2. Exercise. Prove that $\overline{K}$ is a compact.

Hint. Proceed as we did to solve Exercise 8.8.1.

8.8.3. For any $\chi \in V^*$ and any $\varepsilon > 0$, let $\overline{K_{\chi,\varepsilon}}$ be the part of $\overline{K}$ consisting of the vectors v that satisfy

$$\left| \langle v, \chi\rangle - \int_G \langle f(g), \chi(g)\rangle dg \right| \leq \varepsilon.$$

8.8.4. Exercise. Prove that the intersection of finitely many sets of the form $\overline{K_{\chi,\varepsilon}}$ is nonempty.

Since the sets $\overline{K_{\chi,\varepsilon}}$ are compact (as closed subsets of $\overline{K}$), it follows that their intersection is also nonempty. This intersection, clearly, consists of one point, the point $v \in V$ desired, which satisfies (1.25).

Remark. Observe the following property of the integral constructed that likens it with the Riemann integral:

The vector $v = \displaystyle\int_G f(g)dg$ is the limit of linear combinations of the form $\displaystyle\sum_{k=1}^{n} C_k f(g_k)$,

where $C_i \geq 0$, and $\sum\limits_{i=1}^{n} C_i = 1$.

8.8.5. Exercise. Prove that, for any continuous operator $S\colon V \longrightarrow W$, where W is a normed space, we have

$$\int_G Sf(g)dg = S\int_G f(g)dg.$$

Hint. Prove that, for any $\chi \in W^*$, we have

$$\left\langle \int_G Sf(g)dg, \chi \right\rangle = \left\langle S\int_G f(g)dg, \chi \right\rangle.$$

Now, apply the result of Exercise 8.8.3.

8.9. Orthogonality relations. To deduce these relations, we have to introduce in an infinite dimensional Hilbert space H a particular class of operators, *trace-class operators*, for which trace is defined.

First, consider the *finite dimensional operators*, i.e., operators A with $\dim A(H) < \infty$.

8.9.1. Exercise. Prove that every finite dimensional operator A in H is of the form

$$Av = \sum_{k=1}^{n} (v, x_k)y_k \text{ for a finite subset } (x, y) = x_1, \dots, x_n, y_1, \dots, y_n \in H. \tag{1.26}$$

Hint. For $y_1, \dots, y_n$ take a basis of H.

8.9.2. Exercise. Prove that for any representation of A in the form (8.9.1), the value $\sum\limits_{k=1}^{n} (y_k, x_k)$ does not vary, i.e., it does not depend on the choice of the vectors x, y.

The invariant quantity determined in Exercise 8.4.2 (in non-invariant terms) will be called the *trace* of A and denoted by $\operatorname{tr} A$.

8.9.3. For any finite dimensional operator A, set

$$|||A||| = \inf \sum_{i=1}^{n} \| x_i \| \cdot \| y_i \|,$$

where inf is taken over all possible representations (8.9.1).

Exercise. Prove that $A \mapsto |||A|||$ is a norm.

8.9.4. Exercise. Prove that $|||A||| > \| A \|$.

Hint. Represent $\| A \|$ in the form

$$\| A \| = \sup_{\|\xi\|=1, \|\eta\|=1} \operatorname{Re}(A\xi, \eta).$$

8.9.5. Properties of the norm $|||\cdot|||$. Consider the collection $\mathcal{O}_f$ of all finite dimensional operators with the metric

$$\rho(A, B) = |||A - B|||.$$

Clearly, $\mathcal{O}_f$ is a linear space. It is not manifest but still true that $\mathcal{O}_f$ is not complete with respect to the metric $\rho(\cdot, \cdot)$.

Let $\{A_n\}_{n=1}^{\infty}$ be a fundamental sequence in $\mathcal{O}$. This sequence is, clearly, fundamental with respect to the usual metric

$$\rho_n(A, B) = \| A - B \| .$$

8.10. Statement (See any textbook in Functional Analysis)**.** *The space of all bounded operators is complete with respect to the metric* $\rho_n(\cdot, \cdot)$.

Therefore, the completion of the space $\mathcal{O}_f$ of all finite dimensional operators with respect to the metric ρ can be embedded into the space of all bounded operators. Denote this completion by $\mathcal{S}$.

8.10.1. Exercise. Prove that every operator A from $\mathcal{S}$ can be defined by the formula

$$Av = \sum_{k=1}^{\infty} (v, x_k) y_k, \text{ where } \sum_{k=1}^{\infty} \| x_k \| \cdot \| y_k \| < \infty.$$

8.10.2. Exercise. Prove that any $A \in \mathcal{S}$ is of trace class. Namely, for any orthonormal basis $\{v_i\}_{i=1}^{\infty}$ of H the series $\sum\limits_{i=1}^{\infty} (Av_i, v_i)$ converges and the sum does not depend on the choice of the basis.

Hint. Prove the statement of Exercise 8.10.2 for finite dimensional operators and apply the inequality $|\operatorname{tr} A| < |||A|||$.

8.10.3. Exercise. For $A \in \mathcal{S}$, set

$$\mathring{A} = \int_g T(g)A(T(g))^* dg.$$

Prove that $\mathring{A} \in \mathcal{S}$ and $\operatorname{tr} \mathring{A} = \operatorname{tr} A$.

Hint. Apply the statement of Exercise 8.10.2 to the normed space $\mathcal{S}$ and the map $A \mapsto \operatorname{tr} A$.

8.11. Theorem (The orthogonality relations)**.** *Let $T\colon G \longrightarrow GL(H)$ be an irreducible representation of a compact group G in a Hilbert space H. Let $\{v_i\}_{i=1}^{\infty}$ be an orthonormal basis of H. Set*

$$t_{ij}(g) = (T(g)v_i, v_j).$$

Then the functions $t_{ij}(g)$ are pairwise orthogonal and their norms in $L_2(G)$ [6] *coincide.*

Proof. Set

$$E_{ij}v_k = \delta_{kj}v_i,\ E_{ij}^0 = \int_G T(g)E_{ij}T(g^{-1})dg.$$

Clearly, $E_{ij} \in \mathcal{S}$ and $E_{ij} = \delta_{ij}$.

But E_{ij}^0 commutes with all operators $T(g)$, and so $E_{ij}^0 = \lambda I$ by Theorem 6.7.1. But $E_{ij}^0 \in \mathcal{S}$, and hence $\operatorname{tr} E_{ij}^0 = \operatorname{tr} E_{ij} < \infty$. But this is only possible if $\dim H < \infty$.

Let us compute the (k,l)-th matrix element of E_{ij}^0. Since

$$E_{ij}^0 = \lambda \dim H = \operatorname{tr} E_{ij} = \delta_{ij},$$

it follows that

$$(E_{ij}^0)_{kl} = \frac{1}{\dim H}\delta_{kl}\delta_{ij}.$$

On the other hand,

$$(E_{ij}^0)_{kl} = (E_{ij}^0 v_k, v_l) = \int_G (T(g)E_{ij}^0 T(g)^* v_k, v_l)dg = \int_G t_{li}(g)\overline{t_{kj}(g)}dg =$$
$$(t_{li}, t_{kj})_{L_2(G)}.$$

Thus,

$$(t_{li}, t_{kj})_{L_2(G)} = \frac{1}{\dim H}\delta_{kl}\delta_{ij}. \qquad \square$$

Observe that we not only proved the orthogonality relations but proved Theorem 6.7.1 on the finite dimension of irreducible representations of compact groups.

---

[6] Recall that $L_2(X)$ is the space of all square-integrable functions on a given space X.

8.12. A refinement of Theorem 6.7.1. *Let T be a unitary representation of the compact Lie group G in a Hilbert space H. Then there exists a nonzero finite dimensional G-invariant subspace $H_0 \subset H$.*

In the proof of this refinement, the following fact is crucial.

The spaces $H_{c,d}$ are finite dimensional for all $c, d > 0$ if $A \in \mathcal{S}$ is self-adjoint.

We will split the proof of this fact into several problems.

8.12.1. Exercise. Prove that the operators of class $\mathcal{S}$ are completely continuous, i.e., they send any bounded set into a compact one.

Hint. Prove that all finite dimensional operators are completely continuous and the passage to the limit with respect to the norm $\| \cdot \|$, and, even more so, with respect to the norm $||| \cdot |||$ preserves this property.

8.12.2. Exercise. Prove that no completely continuous operator in the infinite dimensional space can have a bounded inverse.

Hint. Prove that completely continuous operators form an ideal in the ring of all bounded operators. Prove that the identity operator is not completely continuous.

8.12.3. Exercise. Prove that A has a bounded inverse on the spaces $H_{c,d}$ for all $c, d > 0$.

Hint. Prove that the collection of vectors $\{Ax \mid x \in H_{c,d}\}$ is dense in $H_{c,d}$ and on this collection the following inequality holds:

$$\| A^{-1} \| \leq \max(|c|^{-1}, |d|^{-1}).$$

The refinement 8.12 of Theorem 6.7.1 may now be proved as follows. Let A be a positive self-adjoint operator of $\mathcal{S}$ acting on H. Then $\mathring{A}$ is a positive self-adjoint operator of $\mathcal{S}$ commuting with all the $T(g)$, where $g \in G$. Among the spaces $H_{c,d}$ for any $c, d > 0$, at least one space is nonzero (since otherwise the operator A would not have been positive). By Exercises 8.12.2 and 8.12.3 this nonzero space is of finite dimension, as was required. □

Lecture 9. The regular representations of the compact groups

Let X be any set with a G-action for a compact group G. In other words, there is given a map (called *action*)

$$a \colon G \times X \longrightarrow X, \quad a(g, x) = gx$$

continuous with respect to the topologies in G and X and such that

$$(gh)x = g(hx) \text{ for any } g, h \in G \text{ and } x \in X.$$

This gives rise to a representation a^* of G in the space of functions on X:

$$(a^*(g)(f)(x) = f(a(g^{-1})x).$$

Usually, for infinite sets X, we confine ourselves to a subclass of all functions (to continuous, smooth, integrable, measurable, and so on, functions).

Among the sets with G-action, there is one, a “most natural” one, namely, G itself. On it, there are three “most natural” G-actions (right, left and adjoint):

$$R_g\colon h \mapsto hg^{-1};\quad L_g\colon h \mapsto gh;\quad \mathrm{Ad}_g\colon h \mapsto ghg^{-1}.$$

Let us investigate L_g^*. Consider the space $C(G)$ of continuous functions on G with the inner product given by the formula

$$(f, g) = \int_G f(g)\overline{h(g)}dg.$$

The completion of $C(G)$ with respect to this inner product (or the metric associated with it) will be denoted by $L_2(G)$ (the *space of square integrable functions*).

We will briefly denote by T the left regular representation L_g^* of G in $L_2(G)$.

9.1. Theorem. *Let T_α, where $\alpha \in A$, for some set A, be all irreducible representations of G (up to equivalence); let n_α be the dimension of T_α. Then $L_2(G) = \bigoplus\limits_{\alpha\in A} n_\alpha T_\alpha$, and each irreducible representation T_α enters $L_2(G)$ with multiplicity equal to* $\dim T_\alpha$.

Proof. Let $v_1, \dots, v_{n_\alpha}$ be an orthonormal basis in H_α, the space of representation T_α. Then the matrix elements $t_{ij}^{(\alpha)}$ of the operator $T_\alpha(g)$ with respect to this basis satisfy the orthogonality relations

$$\int_G t_{ij}^{(\alpha)}\overline{t_{kl}^{(\alpha)}}dg = \frac{1}{n_\alpha}\delta_{ik}\delta_{jl}.$$

Let us show that $H_\alpha = \mathrm{Span}\left(\overline{t_{ij}^{(\alpha)}}\ \text{ for all }\ i, j\right)$ is T_α-invariant. Indeed,

$$\begin{aligned}
&T_\alpha(g^{-1}h) = T_\alpha(g^{-1})T_\alpha(h) \Longleftrightarrow\\
&t_{ij}^{(\alpha)}(g^{-1}h) = \sum_k t_{ik}^{(\alpha)}(g^{-1})t_{kl}^{(\alpha)}(h) \overset{\langle Xv,w\rangle=\langle v,(X^t)^{-1}w\rangle}{\Longleftrightarrow}\\
&T_\alpha(g)\overline{t_{ij}^{(\alpha)}(h)} = \overline{t_{ij}^{(\alpha)}(g^{-1}h)} =\\
&\sum_k t_{ki}^{(\alpha)}(g)\overline{t_{kj}^{(\alpha)}(h)}.
\end{aligned}$$

Hence, H_α is T_α-invariant. Moreover, denote by $c_{ij}^{(\alpha)}$ the coordinates in H_α with respect to the basis $t_{ij}^{(\alpha)}$. Then

$$T_\alpha(g)(c_{ij}^{(\alpha)}) = \sum_{i=1}^{n_\alpha} t_{ki}^{(\alpha)}(g)c_{ij}^{(\alpha)} = \widetilde{c}_{ij}^{(\alpha)}.$$

Therefore, $H_\alpha = \underbrace{V_\alpha \oplus \cdots \oplus V_\alpha}_{n_\alpha-\text{ many summands}}$. As for finite groups, we prove that

$$\int_G t_{ij}^{(\alpha)}(g) = \sum_k t_{kl}^{(\beta)}(g)dg = 0 \quad \text{for } \alpha \neq \beta.$$

It remains to prove that the spaces H_α exhaust $L_2(G)$. Let H' be the orthogonal complement to $\bigoplus\limits_\alpha H_\alpha$ in $L_2(G)$.

9.2. Exercise. Show that H' has an invariant finite dimensional subspace V.

Let $\widetilde{V}^*$ be an irreducible component of H', let T_α the representation in $\widetilde{V}$ for some α (since T_α exhaust all irreducible representation of G, it follows that $\widetilde{V}$ must be among one of them). Let us show that $\widetilde{V} \subset H_\alpha$. This implies a contradiction since $\widetilde{V} \perp (\bigoplus\limits_\alpha H_\alpha)$ and the inner product is non-degenerate.

9.2.1. Remark. For *anti*-symmetric products, even non-degenerate ones, $V \perp V$ does not imply $V = 0$.

Let $w_1, \dots, w_{n_\alpha}$ be a basis of $\widetilde{V}$. Then

$$w_i(g^{-1}h)) = (T_\alpha(g)w_i)(h) = \sum_k \overline{t_{ij}^{(\alpha)}(g^{-1})}w_(h).$$

Setting $h = e$ and replacing g by g^{-1}, we obtain

$$w_i(g) = \sum_k \overline{w_k(e)}t_{ij}^{(\alpha)}(g).$$

But, by assumption, $\widetilde{V} \perp (\bigoplus\limits_\alpha H_\alpha)$. Hence, $V = 0$. □

Lecture 10. Lie algebras (after É. Vinberg)

10.1. Linear groups. Their Lie algebras. A *linear Lie group* G is, by definition, a submanifold in the manifold $\mathrm{Mat}(n;\mathbb{R})$ of $n \times n$-matrices. Denote the tangent space to G at point g by T_gG. This space is spanned by all vectors tangent to the curves in G passing through $g \in G$.

Since G is a Lie group, there exists a parametric representation of G in a neighborhood of g such that $g = g(0, \dots, 0)$. The matrices $\frac{\partial g(u)}{\partial u_i}|_{u=0}$, where $i = 1, \dots, d$, form a basis of T_gG. Clearly, $\dim T_gG = \dim G$ for any g.

The tangent space T_eG at the unit is called the *Lie algebra* of G and is denoted (after Bourbaki) by small Gothic letters corresponding to the capital Latin characters that denote the Lie group, e.g., $\mathfrak{g} = Lie(G)$. Obviously, the G-action sends one tangent space into another:

$$g \cdot \mathfrak{g} = T_g G.$$

For any fixed $g \in G$, define the *adjoint* automorphism Ad_g of G by setting

$$\mathrm{Ad}_g(h) = ghg^{-1}.$$

These adjoint automorphisms of G are called the *inner* ones. They generate the group $\mathrm{Ad}\, G$ of inner automorphisms of G.

Since Ad_g preserves the unit e of G, it follows that Ad_g induces a transformation of the tangent space at e; we will denote the induced transformation also by Ad_g. Clearly,

$$\mathrm{Ad}_g(X) = gXg^{-1} \quad \text{for any } X \in \mathfrak{g}.$$

Further, let $g(t)$ be any curve in G such that

$$\left.\frac{dg(t)}{dt}\right|_{t=0} = Y \in \mathfrak{g} \quad \text{and } g(0) = I.$$

We have

$$\frac{d}{dt}(g(t)Xg(t)^{-1})\,|_{t=0} = \frac{d}{dt}\left[(I + tY + o(t))X(I - tY + o(t))\right] = [Y, X] \in \mathfrak{g}.$$

That is, for any $X, Y \in \mathfrak{g}$ the matrix $[Y, X]$ is also in $\mathfrak{g}$. This means that the commutator of matrices from $\mathfrak{g}$ makes $\mathfrak{g}$ into an algebra. This algebra is called the *Lie algebra* and the product (multiplication) in it is called the *bracket* or *commutator* for obvious reasons.

More generally, any subspace $\mathfrak{g} \subset \mathrm{Mat}(n; \mathbb{R})$ closed with respect to the bracket is called a *linear* Lie algebra. The point is there are also *abstract* Lie algebras, or, at least, Lie algebras realized not via matrices but somehow else. To define them, observe that the commutator satisfies the following identities:

$$\begin{aligned}
&[X, Y] = -[Y, X] \text{ (anti-commutativity)},\\
&[X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]] \text{ (Jacobi identity)}
\end{aligned}$$

The algebra $\mathfrak{g}$ whose product satisfies the two above identities is called a *Lie algebra*.

10.2. Examples. 1) $\mathrm{Lie}(GL(n; \mathbb{R})) \cong \mathfrak{gl}(n; \mathbb{R})$, i.e., as a space, the Lie algebra of $GL(n; \mathbb{R})$ is isomorphic to the space $\mathrm{Mat}(n; \mathbb{R})$ of all $n \times n$ matrices. But in $\mathrm{Mat}(n; \mathbb{R})$ two operations are natural:

(a) just the product (juxtaposition) of matrices with respect to which $\mathrm{Mat}(n; \mathbb{R})$ is an associative algebra;

(b) the bracket which makes an associative algebra $\mathrm{Mat}(n; \mathbb{R})$ into a nonassociative (see the Jacobi identity), namely, L i e algebra. To distinguish these algebras we denote the latter by $\mathfrak{gl}(n; \mathbb{R})$.

2) For any associative algebra A, let A_L be the Lie algebra whose space is a copy of A and the bracket is the commutator in A.

3) $\mathrm{Lie}(SL(n;\mathbb{R})) = \mathfrak{sl}(n;\mathbb{R}) = \{X \in \mathfrak{gl}(n;\mathbb{R} \mid \operatorname{tr} X = 0\}$. Indeed, if X belongs to a small neighborhood of I_n, i.e., if X is of the form

$$X = I_n + tY + o(t),$$

then

$$\det X = 1 + t \operatorname{tr} Y + o(t)$$

implying

$$\det X = 1 \Longleftrightarrow \operatorname{tr} Y = 0.$$

Therefore $\mathrm{Lie}(SL(n;\mathbb{R}) \subset \mathfrak{sl}(n;\mathbb{R})$. Then dimension consideration

$$\dim G = \dim Lie(G)$$

completes the proof.

Observe that we could have gotten the same result from a useful formula known (at least, it *should* be known) from Algebra courses, namely:

$$\det X = \exp \operatorname{tr} \log X \text{ for } X \text{ from a small neighborhood of } I_n \in \mathrm{Mat}(n;\mathbb{R}).$$

Hence,

$$\det X = 1 \Longleftrightarrow \operatorname{tr} \log X = 0.$$

4) The Lie algebras of the Lie groups $O(n)$ and $SO(n) = O(n) \cap SL(n;\mathbb{R})$ are identical. Indeed, let

$$A(t) \in O(n) \text{ for any } |t| < \varepsilon; \quad A(0) = I_n, \quad A'(0) = X.$$

By differentiating the identity

$$A(t)A^t(t) = I_n \tag{1.27}$$

at $t = 0$ we obtain

$$X + X^t = 0.$$

The group $O(n)$ is determined by $\frac{1}{2}n(n+1)$ equations

$$\sum_k A_{ik}A_{jk} = \delta_{ij} \text{ for any } i \leq j. \tag{1.28}$$

Moreover, the rank of the Jacobi matrix at I_n for each of these equations is also equal to $\frac{1}{2}n(n+1)$ (Exercise: Why?), i.e., the condition for existence of the inverse function is fulfilled. Therefore,

$$\dim O(n) = n^2 - \frac{1}{2}n(n+1) = \frac{1}{2}n(n-1) = \dim \mathfrak{o}(n),$$

where

$$\mathfrak{o}(n) = \{X \in \mathfrak{gl}(n) \mid X + X^t = 0\}.$$

10.2.1. Exercise. $O(n) = SO(n) \coprod A \cdot SO(n)$, where A is any matrix from $O(n)$ with $\det A = -1$.

10.3. A main method. One of the principal methods in the study of representations of Lie groups is reduction to the same question in the context of Lie algebras.

Let us demonstrate, first of all, that each connected linear Lie group is uniquely determined by its Lie algebra.

Let $A(t)$ be a smooth curve in a linear Lie group G. By definition of the tangent space

$$A'(t) \in T_{A(t)}G = A(t)\mathfrak{g} \text{ for any } t.$$

Hence,

$$A'(t) = A(t) \cdot X(t) \text{ where } X(t) \in \mathfrak{g}. \tag{1.29}$$

We can consider the matrix equation (1.29) for any continuous curve $X(t) \in \mathrm{Mat}(n; \mathbb{R})$; clearly, this is a linear system of ordinary differential equations. Therefore, eq. (1.29) has a solution on the whole domain of definition of $X(t)$, and if $A_0(t)$ is the solution for the initial value $A(0) = I_n$, then the solution corresponding to the initial value $A(0) = C$ is $CA_0(t)$.

10.3.1. Lemma. *Let $X(t) \in \mathfrak{g}$ be an arbitrary continuous curve, $A_0(t)$ the solution of eq.* (1.29) *corresponding to the initial value $A(0) = I_n$. Then $A(t) = G$.*

Proof. Let $A(u)$, where $u = (u_1, \dots, u_d)$, be a parametric representation of a neighborhood of the unit of G; let $A(0) = I_n$. Since $A(u)X(t) \in T_{A(u)}G$, it follows that

$$A(u)X(t) = \sum \varphi_i(u; t)\frac{\partial A(u)}{\partial u_i}.$$

(Indeed, $\dfrac{\partial A(u)}{\partial u_i}$ constitute a basis of $T_{A(u)}G$.) Consider the system of differential equations

$$u_i' = \varphi_i(u; t) \text{ for } i = 1, \dots, d.$$

If $u = u(t)$ is a solution of this system for initial values $u_i = 0$, then $\widetilde{A}(t) = A(u(t))$ is a solution of eq. (1.29) for the initial value $A(0) = I_n$. Hence, $\widetilde{A}(t) = A(u(t))$ and $A(t) \in G$ for sufficiently small values of t.

Let $t_1 = \sup\{t \mid A(t) \in G\}$. Set $X_1(t) = X(t_1 + t)$ and let $A_1(t)$ be the solution of the initial value problem

$$A_1'(t) = A_1(t)X_1(t), \quad A_1(0) = I_n.$$

By the above, $A_1(t) \in G$ for sufficiently small t. But, on the other hand,

$$A(t) = A(t_1)A_1(t - t_1) \text{ for any } t \text{ close to } t_1.$$

Take $t_2 = t_1 - \varepsilon$ for a sufficiently small $\varepsilon > 0$. We have

$$A(t_1) = A(t_2)(A_1(-\varepsilon))^{-1} \in G.$$

Therefore, $A(t) \in G$ for any t sufficiently close to t_1, even for $t > t_1$. This contradicts the choice of t_1 and proves Lemma. □

10.3.2. Theorem. (1) *Let G, H be linear Lie groups, $\mathfrak{g} = \mathrm{Lie}(G)$ and $\mathfrak{h} = \mathrm{Lie}(H)$. If G is connected and $\mathfrak{g} \subset \mathfrak{h}$, then $G \subset H$.*

(2) *If G and H are connected and $\mathfrak{g} \cong \mathfrak{h}$, then $G \cong H$.*

Proof. If G is connected, then there exists a smooth curve $A(t)$ that connects $A \in G$ with $e \in G$ and we may assume that $A(0) = I$. The matrix-valued function $A(t)$ satisfies eq. (1.29). If $\mathfrak{g} \subset \mathfrak{h}$, then $X(t) \in \mathfrak{h}$ and, by Lemma 10.3.1, $A(t) \in H$. This proves the first statement. The second statement of Theorem obviously follows from the first one. □

10.3.3. Theorem. *Let G be a connected linear Lie group realized in $\mathbb{R}^n$. A subspace $V \subset \mathbb{R}^n$ is G-invariant if and only if V is invariant under $\mathfrak{g} = Lie(G)$.*

Proof. Let $A(t)$ be a smooth curve in G such that

$$A(0) = I, \quad A'(0) = X \in \mathfrak{g}. \tag{1.30}$$

For any $a \in V$, we have

$$Xa = \frac{d}{dt}(A(t)a)\,|_{t=0}\ . \tag{1.31}$$

Therefore, if V is G-invariant, then it is $\mathfrak{g}$-invariant.

Now, suppose V is $\mathfrak{g}$-invariant. Let $v_1, \dots, v_n$ be a basis of V. Take a smooth curve in G such that $A(0) = I$. Let

$$X(t)v_i = \sum_j \varphi_{ij}(t)v_j.$$

Then

$$\frac{d(A(t)v_i)}{dt} = \sum_j \varphi_{ij}(t)A(t)v_j. \tag{1.32}$$

Let $\lambda \in (\mathbb{R}^n)^*$ be a linear form that vanishes on V. Set

$$\lambda_i(t) = \lambda(A(t)v_i).$$

Clearly, $\lambda_j(0) = 0$. Eq. (1.32) implies that

$$\lambda_i'(t) = \sum_j \varphi_{ij}(t)\lambda_j(t).$$

Hence, $\lambda_i(t) = 0$ for *any* t, not only at 0. This means that $A(t)v_j \in V$. □

10.3.4. Theorem. *Let G be a connected linear Lie group. A vector a from the representation space of G is G-invariant if and only if*

$$Xa = 0 \ \ \textit{for all } X \in \mathfrak{g}.$$

Proof. Let $A(t)$ be a smooth curve in G such that $A(0) = I$, $A'(0) = X \in \mathfrak{g}$. If a is G-invariant, then eq. (1.31) implies that $Xa = 0$.

Now suppose $Xa = 0$ for all $X \in \mathfrak{g}$ and let $A(t)$ be a smooth curve in G satisfying eq. (1.29) with $X(t) \in \mathfrak{g}$. Then

$$\frac{d(A(t)a)}{dt} = A(t)X(t)a = 0.$$

If $A(0) = I$, then $A(t)a = a$ for all t. $\square$

10.4. The adjoint group. To every element g of $G \subset GL(\mathbb{R}^n)$ assign the linear transformation Ad of $\mathrm{Mat}(n;\mathbb{R})$:

$$\mathrm{Ad}_g\colon x \mapsto gxg^{-1}$$

Clearly, $\mathrm{Ad} : G \longrightarrow GL(n^2;\mathbb{R})$ is a linear homomorphism. As follows from sects. 10.3.1–10.3.3, $\mathfrak{g} \subset \mathrm{Mat}(n;\mathbb{R})$ is G-invariant.

To find the differential of Ad, take a curve $A(t) \in G$ with $A(0) = I$ and $A'(0) = X \in \mathfrak{g}$; and differentiate the equation

$$\mathrm{Ad}_{A(t)}(C) = A(t)(C)A(t)^{-1}.$$

We obtain

$$\mathrm{ad}_X(C) = [X, C], \text{ where } \mathrm{ad} = d(\mathrm{Ad}).$$

By Theorem 10.3.3 the subspace $\mathfrak{g} \subset \mathrm{Mat}(n;\mathbb{R})$ is $\mathrm{ad}\,\mathfrak{g}$-invariant. As is easy to verify directly,

$$\mathrm{ad}_{[X,Y]} = [\mathrm{ad}_X, \mathrm{ad}_Y],$$

i.e., ad is a Lie algebra homomorphism. The group $\mathrm{Ad}\,G$ is called the *adjoint group* of G; this is a group of linear transformations of G.

Let G be a connected linear Lie group. A subspace $\mathfrak{h} \subset \mathfrak{g}$ is $\mathrm{Ad}\,G$-invariant if and only if it is $\mathrm{ad}\,\mathfrak{g}$-invariant, i.e., if

$$[X, Y] \in \mathfrak{h} \quad \text{for any } X \in \mathfrak{g}, Y \in \mathfrak{h},$$

in other words, if $\mathfrak{h}$ is an ideal of $\mathfrak{g}$.

10.4.1. Theorem. *Let G be a connected linear Lie group, H its Lie subgroup. If H is normal in G, then $\mathfrak{h}$ is an ideal in $\mathfrak{g}$. Conversely, if $\mathfrak{h}$ is an ideal in $\mathfrak{g}$ and H is connected, then H is normal in G.*

Proof. The group H is normal in G if and only if

$$\mathrm{Ad}_A(H) = AHA^{-1} = H \text{ for any } A \in G.$$

The Lie algebra of $\mathrm{Ad}_A(H)$ is $(\mathrm{Ad}_A)(\mathfrak{h})$. Hence, if H is normal, then $\mathfrak{h}$ is $\mathrm{Ad}\,G$-invariant.

Conversely, if $\mathfrak{h}$ is $\mathrm{Ad}\,G$-invariant and H is connected, then for any $A \in G$ the connected groups AHA^{-1} and H have the same Lie algebra $\mathrm{Ad}_A(\mathfrak{h}) = \mathfrak{h}$. Hence, H is a normal subgroup. We observed above that $\mathfrak{h}$ is $\mathrm{Ad}\,G$-invariant if and only if $\mathfrak{h}$ is an ideal in $\mathfrak{g}$. This completes the proof. $\square$

10.4.2. Theorem. *Let $G \subset GL(\mathbb{R}^n)$. A matrix $C \in \mathrm{Mat}(n;\mathbb{R})$ commutes with all matrices from G if and only if C commutes with all matrices from $\mathfrak{g}$.*

Proof. The commutativity condition can be expressed as

$$\mathrm{Ad}_A(C) = C \text{ for all } A \in G.$$

As follows from Theorem 10.3.4, this is equivalent to the condition

$$\mathrm{ad}_X(C) = 0 \text{ for all } X \in \mathfrak{g}.$$

□

Corollary. *A connected Lie group G is commutative if and only if its Lie algebra $\mathfrak{g}$ is commutative, i.e., if $[X,Y]=0$ for any $X,Y \in \mathfrak{g}$.*

10.5. The exponential map. Any solution $A(t)$ of the initial value differential equation

$$A'(t) = A(t)X, \quad A(0) = I, \tag{1.33}$$

where $X \in \mathrm{Mat}(n;\mathbb{R})$, is a curve lying in $GL(n;\mathbb{R})$. Set

$$\exp(tX) = A(t).$$

Eq. (1.33) is a particular case of equation (1.29). Therefore, if G is a linear Lie group and $X \in \mathfrak{g}$, then $\exp(X) \in G$. The map exp is called the *exponential map*. If G is a linear Lie group, then this map is the usual exponent of matrices.

Let us prove that

$$\exp(s+t)X = \exp sX \cdot \exp tX \text{ for any } X \in \mathrm{Mat}(n;\mathbb{R}).$$

To this end, fix s and consider two curves in $GL(n;\mathbb{R})$:

$$B(t) = \exp(s+t)X \quad\text{and}\quad C(t) = \exp sX \cdot \exp tX.$$

We have

$$B'(t) = B(t)X \quad\text{and}\quad C'(t) = C(t)X.$$

Moreover,

$$B(0) = C(0) = \exp sX.$$

Hence, $B(t) = C(t)$ for all t.

Property (1.33) indicates that $\{\exp(tX) \mid t \in \mathbb{R}\}$ is a group. This group is called a *one-parameter group generated* by X.

10.6. Lie group homomorphisms. Let G, H be linear Lie group and $A(u)$, where $u \in \mathbb{R}^d$, a parametrization of a neighborhood U of the unit of G. The map $\varphi\colon U \longrightarrow H$ is called a *local homomorphism* of Lie groups if

1) $\varphi(AB) = \varphi(A)\varphi(B)$ whenever $AB \in U$;
2) $\varphi(A(u))$ is a continuously differentiable function of u.

The differential $d\varphi$ of the local homomorphism $\varphi\colon G \longrightarrow H$ at the unit $I \in G$ is a linear map $d\varphi\colon \mathfrak{g} \longrightarrow \mathfrak{h}$. Let us show that this linear map is a Lie algebra homomorphism, i.e., satisfies

$$d\varphi([X, Y]) = [d\varphi(X), d\varphi(Y)].$$

Indeed, let $C(t)$ be a smooth curve in G such that $C(0) = I$ and $C'(0) = X \in \mathfrak{g}$. Then

$$d\varphi(X) = \frac{d}{dt}\varphi(C(t))\,|_{t=0}\ .$$

Let us differentiate the identity

$$\varphi(AC(t)A^{-1}) = \varphi(A)\varphi(C(t))\varphi(A^{-1}) = \varphi(A)\varphi(C(t))(\varphi(A))^{-1}.$$

We obtain

$$d\varphi(\mathrm{Ad}_{A(X)}) = \mathrm{Ad}_{\varphi(A)}(d\varphi(X)). \tag{1.34}$$

Now, let $A(t)$ be a smooth curve in G such that $A(0) = I$ and $A'(0) = Y \in \mathfrak{g}$. Substituting $A(t)$ for A in (1.34) and differentiating with respect to t we obtain the desired.

10.6.1. Theorem. *Let $\varphi, \psi\colon G \longrightarrow H$ be local homomorphisms of Lie groups. If $d\varphi = d\psi$, then $\varphi = \psi$ in a neighborhood of the unit of G. If, moreover, G is connected and φ, ψ are global homomorphisms, then $\varphi = \psi$ on the whole of G.*

Proof. Let $A(t) \in G$ be a smooth curve given by equation (1.29) with the initial condition $A(0) = I$, the unit of G.

Let us prove that, for the differential $d\varphi|_A$ of $d\varphi$ at $A \in G$, we have

$$d\varphi|_A(AX) = \varphi(A)d\varphi(X) \quad \text{for any } X \in \mathfrak{g}. \tag{1.35}$$

Indeed, if $C(t) \in G$ is a curve such that $C(0) = I$, $C'(0) = X$, then AX is tangent to the curve $AC(t)$ at $t = 0$. To find $d\varphi|_A(AX)$, take the vector tangent at $t = 0$ to the curve

$$\varphi(AC(t)) = \varphi(A)\varphi(C(t)).$$

But this vector is exactly $\varphi(A)d\varphi(X)$.

By eq. (1.35)

$$\frac{d\varphi(A(t))}{dt} = \varphi(A(t))d\varphi(X(t)). \tag{1.36}$$

Since $d\varphi(X(t)) = d\psi(X(t))$, it follows that $\psi(A(t))$ also satisfies eq. (1.36). Hence,

$$\varphi(A(t)) = \psi(A(t)) \text{ on the common domain of definition.}$$

This completes the proof. □

10.6.2. Theorem. *Let G, H be linear Lie groups, $\Phi\colon \mathfrak{g} \longrightarrow \mathfrak{h}$ a Lie algebra homomorphism. Then there exists a local homomorphism $\varphi\colon G \longrightarrow H$ such that $d\varphi = \Phi$. If G is connected and simply connected, then φ can be determined in large, i.e., on the whole G.*

Proof. For a smooth curve $A(t) \in G$ satisfying eq. (1.29) with initial value $A(0) = I$, define its image $B(t) = \varphi(A(t)) \in H$ as the solution of the equation

$$B'(t) = B(t)\Phi(X(t)); \quad B(0) = I.$$

To prove that this $B(t)$, hence, φ, is well-defined in any connected simply connected domain $U \subset G$, we need the following

Lemma (On the map of a rectangle into a Lie group)**.** *Let G be a Lie group, $X(s,t)$, $Y(s,t) \in \mathfrak{g}$ be continuously differentiable functions on the rectangle*

$$R = \{(s,t) \mid 0 \leq s \leq s_0, 0 \leq t \leq t_0\}$$

with values in $\mathfrak{g}$. The G-valued function $A(s,t)$ defined on R and satisfying the initial value problem

$$\begin{cases} \dfrac{\partial A}{\partial s} = A(s,t)X(s,t), \\ \dfrac{\partial A}{\partial t} = A(s,t)Y(s,t), \end{cases} \qquad A(0,0) = I \tag{1.37}$$

exists if and only if the identity

$$\frac{\partial X}{\partial t} - \frac{\partial Y}{\partial s} = [X(s,t), Y(s,t)] \ \text{ for all } s, t \in \mathbb{R} \tag{1.38}$$

holds.

Observe that eq. (1.38) is the Frobenius integrability condition for the system (1.37), so Lemma is a tautology.

Let $A_1(t), A_2(t)$ be two smooth curves in $U \subset G$ satisfying the boundary value problem

$$\begin{aligned} &A_k'(t) = A_k(t)X_k(t), \text{ where } X_k(t) \in \mathfrak{g} \text{ for } k = 1, 2; \\ &A_1(0) = A_2(0) = I, \quad A_1(t_0) = A_2(t_0) = A. \end{aligned} \tag{1.39}$$

Let us construct curves $B_1(t)$ and $B_2(t)$ — solutions of the equations

$$\begin{aligned} &B_k'(t) = AB_k(t)X_k(t), \text{ where } X_k(t) \in \mathfrak{g} \text{ for } k = 1, 2; \\ &B_1(0) = B_2(0) = I \end{aligned}$$

and prove that $B_1(t_0) = B_2(t_0)$.

There exists a smooth homotopy of the curve A_1 into A_2 inside U, i.e., there exists a smooth function

$$A\colon R \longrightarrow U \text{ such that } A(0,t) = A_1(t);\quad A(1,t) = A_2(t);$$
$$A(s,0) = I;\quad A(s,1) = A.$$

The function $A(s,t)$ satisfies eq. (1.37), where $X(s,0) = X(s,t_0) = 0$. Therefore, eq. (1.38) hold.

The integrability condition for the system

$$\begin{cases} \dfrac{\partial B(s,t)}{\partial s} = B(s,t)\Phi(X(s,t)), \\ \dfrac{\partial B(s,t)}{\partial t} = B(s,t)\Phi(Y(s,t)), \end{cases} \qquad A(0,0) = I \tag{1.40}$$

is of the form

$$\frac{\partial \Phi(X)}{\partial t} - \frac{\partial \Phi(Y)}{\partial s} = [\Phi(X(s,t)), \Phi(Y(s,t))] \text{ for all } s,t \in \mathbb{R} \tag{1.41}$$

and follows from (1.38) since Φ is a Lie algebra homomorphism (i.e., $\Phi([X,Y] = [\Phi(X),\Phi(Y)]$.) Hence, there exists a solution of (1.40) satisfying the initial condition $B(0,0) = I$. Clearly, $B(0,t) = B_1(t)$, $B(1,t) = B_2(t)$. Since $X(s,t_0) = 0$, it follows that $B(s,t_0) = \text{const}$, in particular, $B_1(t_0) = B_2(t_0)$. Therefore, φ is well-defined.

Let us show that φ is a local homomorphism. Let $A_1, A_2, A_1A_2 \in G$. Take smooth curves $A_1(t), A_2(t) \subset U$ that connect I to A_1 and A_2, respectively; let

$$A_k'(t) = A_k(t)X_k(t) \text{ for } k = 1,2$$

Then the curve

$$A(t) = \begin{cases} A_1(t) & \text{for } 0 \le t \le 1, \\ A_1A_2(t-1) & \text{for } 1 \le t \le 2 \end{cases}$$

connects I to A_1A_2 and satisfies eq. (1.29) for

$$X(t) = \begin{cases} X_1(t) & \text{for } 0 \le t \le 1, \\ X_2(t-1) & \text{for } 1 \le t \le 2. \end{cases}$$

The image $B(t) = \varphi(A(t))$ of this curve satisfies, thanks to the definition of map φ, equation (1.29), and therefore

$$B(t) = \begin{cases} B_1(t) & \text{for } 0 \le t \le 1, \\ (B_1B_2)(t-1) & \text{for } 1 \le t \le 2, \end{cases}$$

where $B_k(t) = \varphi(A_k(t))$ and $B_k(1) = B_k$ for $k = 1,2$. In particular,

$$\varphi(A_1A_2) = \varphi(A(2)) = B(2) = B_1B_2 = \varphi(A_1)\varphi(A_2).$$

The fact that φ satisfies condition 2) of the definition of the local Lie group homomorphism follows from the theorem on smooth dependence of the solution of the system of differential equations on initial/boundary values.

To complete the proof of Theorem, it only remains to verify that $d\varphi = \Phi$. Let $A(t) \in G$ be a smooth curve, $A(0) = I$, $A'(0) = X \in \mathfrak{g}$. Let $B(t) = \varphi(A(t))$. By the definition of φ we have

$$B'(0) = B(0)\Phi(X) = \Phi(X).$$

But $B'(0)$ is exactly $d\varphi(X)$. $\square$

10.7. Two Lie groups G and H are said to be *locally isomorphic* if there is a one-to-one map φ of a neighborhood of the unit of G on a neighborhood of the unit of H such that φ and φ^{-1} are local homomorphisms.

Theorem. *Lie groups G and H are locally isomorphic if and only if their Lie algebras are isomorphic. The connected simply connected Lie groups are locally isomorphic if and only if they are isomorphic.*

Proof. Let φ be a local isomorphism. Set $\psi = \varphi^{-1}$. The Lie algebra homomorphisms $d\varphi\colon \mathfrak{g} \longrightarrow \mathfrak{h}$ and $d\psi\colon \mathfrak{h} \longrightarrow \mathfrak{g}$ are inverse to each other since $d\varphi d\psi = d(\varphi\psi) = \mathrm{Id}$ and $d\psi d\varphi = d(\psi\varphi) = \mathrm{Id}$. Hence, $\mathfrak{g} \cong \mathfrak{h}$.

Conversely, let $\mathfrak{g} \cong \mathfrak{h}$ and let $\Phi\colon \mathfrak{g} \longrightarrow \mathfrak{h}$ and $\Psi\colon \mathfrak{h} \longrightarrow \mathfrak{g}$ be mutually inverse Lie algebra homomorphisms. By Theorem 10.6.2 there exist local Lie group homomorphisms $\varphi\colon G \longrightarrow H$ and $\psi\colon H \longrightarrow G$ such that $d\Psi = \psi$ and $d\Phi = \varphi$. Moreover, by Theorem 10.6.2, if G and H are connected and simply connected, φ and ψ can be determined globally. $\square$

The composition $\varphi\psi\colon H \longrightarrow H$ is a local homomorphism whose differential at the unit is equal to $\Psi\Phi$. Theorem 10.6.1 implies that $\varphi\psi = \mathrm{id}$ and $\psi\varphi = \mathrm{id}$. Hence, φ is a local isomorphism; moreover, if G and H are connected and simply connected, then φ is an isomorphism in the large.

10.8. According to Theorem 10.7 in the class of locally isomorphic groups there is at most one, up to isomorphism, connected simply connected group. Let $\widetilde{G}$ be a connected simply connected group locally isomorphic to G. It is unclear yet if such a $\widetilde{G}$ exists! By Theorem 10.6.2 an isomorphism Φ is the differential of a homomorphism φ defined globally. In a sufficiently small neighborhood of the unit of $\widetilde{G}$ the homomorphism φ is a local isomorphism; hence, its kernel N is discrete.

Theorem. *A totally disconnected normal subgroup N of a connected Lie group G belongs to the center of G.*

Proof. If $C \in N$ and $A(t)$ is a continuous curve in G, then $A(t)CA(t)^{-1}$ belongs to N^C, the connected component of N connecting C, which is equal to $\{C\}$. Hence, $A(t)C = CA(t)$. Since G is connected, C commutes with all elements of G. $\square$

By this theorem N is contained in the center of $\widetilde{G}$. The subgroup $\varphi(\widetilde{G})$ of G contains a neighborhood of the unit of G.

Lemma. *Let G be a connected Lie group. If a subgroup H contains a neighborhood of the unit of G, then $H = G$.*

Proof. Take a smooth curve $A(t) \in G$ such that $A(0) = I$. Clearly, $A(t) \in H$ for sufficiently small t. Let $u = \sup\{t \mid A(t) \in H\}$. Take $v = u - \varepsilon$; for a sufficiently small ε we have

$$A(v)^{-1}A(u) \in G \text{ and } A(u) = A(v)\left(A(v)^{-1}A(u)\right) \in H.$$

Hence, $A(t) \in H$ for all t sufficiently close to u, but this contradicts the choice of u. □

The comparison of the above arguments and this Lemma implies

10.9. Theorem. *Let G be a connected Lie group, $\widetilde{G}$ the locally isomorphic to G connected simply connected Lie group. There exists a homomorphism $\widetilde{G} \longrightarrow G$ whose kernel is a discrete normal subgroup of $\widetilde{G}$ lying in the center of G.*

Observe that the existence of $\widetilde{G}$ is still to be proved.

10.10. Let $\varphi\colon G \longrightarrow H$ be a homomorphism of linear Lie groups. Let us show that

$$\varphi(\exp(X)) = \exp d\varphi(X) \text{ for any } X \in \mathfrak{g}.$$

Let $\varphi(\exp(tX)) = B(t) \in H$. We have

$$B'(t) = d\varphi\,|_{\exp(tX)}(\exp(tX)(X)) = B(t)d\varphi(X).$$

Since $B(0) = I$, it follows that $B(t) = \exp(td\varphi(X))$.

10.10.1. Lemma. *A given matrix $C \in \mathrm{Mat}(n;\mathbb{R})$ commutes with X if and only if C commutes with $\exp(tX)$ for all t.*

Proof. The identity

$$C \cdot \exp(tX) = \exp(tX) \cdot C \tag{1.42}$$

implies that $CX = XC$: differentiate eq. (1.42) with respect to t at $t = 0$.

Conversely, if $CX = XC$, set $A(t) = \exp(tX) \cdot C$. Clearly,

$$A'(t) = A(t)X, \quad A(0) = C.$$

Hence, $A(t) = C\exp(tX)$, as was required. □

10.10.2. Lemma. *If $[X, Y] = 0$, then*

$$\exp(X + Y) = \exp(X) \cdot \exp(Y).$$

Proof. Set $A(t) = \exp(tX) \cdot \exp(tY)$. By Lemma 10.10.1, it follows that the initial value problem

$$A'(t) = A(t)(X + Y), \quad A(0) = I$$

has $A(t) = \exp t(X + Y)$ as a solution. $\square$

10.10.3. Lemma. *For any* $X \in \mathrm{Mat}(n; \mathbb{R})$*, we have* $\exp(X) = I + X + o(X)$*.*

Proof. By the theorem on smooth dependence of the solution of the system of differential equation on the parameters, the map $\exp\colon \mathrm{Mat}(n; \mathbb{R}) \longrightarrow GL(n; \mathbb{R})$ is continuously differentiable. Since $\exp(0) = I$, then by Taylor's formula

$$\exp(X) = I + L(X) + o(X),$$

where $L\colon \mathrm{Mat}(n; \mathbb{R}) \longrightarrow GL(n; \mathbb{R})$ is the differential of exp at 0. For the fixed X, the definition of exp implies that

$$\exp(tX) = I + tX + o(t),$$

hence, $L(X) = X$. $\square$

10.11. Theorem. *Let G be a connected commutative Lie group. The map* $\exp\colon \mathfrak{g} \longrightarrow G$ *is a homomorphism of the commutative group* $\mathfrak{g}$ *(with respect to addition) onto G. The kernel of this homomorphism is discrete.*

Proof. Lemma 10.10.2 shows that exp is a homomorphism, Lemma 10.10.3 that the Jacobian of exp does not vanish at 0, and hence exp is a one-to-one map of a neighborhood of the unit (which is the origin 0 for $\mathfrak{g}$) of $\mathfrak{g}$ on a neighborhood of $I \in G$. This implies that

(1) the kernel of exp is discrete;

(2) $\exp G$ contains a neighborhood of $I \in G$.

Since $\exp G$ is a subgroup of G, then Lemma 10.10.3 implies that $\exp G = G$. $\square$

Recall that any discrete (enumerable but not finite) subgroup of $\mathbb{R}^d$ is a *lattice*, i.e., the $\mathbb{Z}$-span of a collection of vectors $X_1, \dots, X_r$, where $r \leq d$.

10.12. Exercise. If $A \in SL(2; \mathbb{R})$ with $\operatorname{tr} A < -2$, then A cannot be represented as $\exp(X)$ for any $X \in \mathrm{Mat}(2; \mathbb{R})$.

Lecture 11. Selected applications. Hamiltonian mechanics

The basic notion of the classical mechanics is that of the *phase space*. The points of the phase space correspond to the various possible states of the dynamical system, both positions and momenta, the functions on the phase space describe various physical characteristics of the system considered.

A mathematical model of the phase space is given by a *symplectic* manifold (M, ω), i.e., a manifold M with a non-degenerate closed differential 2-form ω.

11.1. Theorem (Darboux). [7] *Any non-degenerate closed differential* 2*-form can be locally reduced to the form*

$$\omega = \sum_{i=1}^{n} dp_i \wedge dq_i.$$

The canonical coordinates that locally exist, thanks to Darboux's theorem, on each symplectic manifold describe *positions*, q, and *momenta*, p, respectively.

11.2. Examples of symplectic manifolds. 1) If $M = T^*N$, then N is called the *configuration* space.

2) An example of a symplectic manifold which is not of the form T^*N for any N is S^2 on which ω is the area element. (Clearly, no compact manifold can be of the form 1).

The dynamics on the symplectic manifold (M, ω) is given with the help of a *Hamiltonian function* H via the formula, where we tacitly assume that all functions considered depend on t,

$$\dot{f} = \{f, H\}_{P.B.} \text{ for any } f \in \mathcal{F}(M), \tag{1.43}$$

where $\{\cdot, \cdot\}_{P.B.}$ is the *Poisson bracket* to be defined shortly, and the collection of functions $\mathcal{F}(M) := C^\infty(M)$, or of some other class (sometimes rather strange), depending on the problem; the dynamical parameter t is interpreted as Time which means that we tacitly replace the phase space M by $M \times \mathbb{R}$ (at least, locally) and we set

$$\dot{f} \overset{\text{def}}{=} \frac{d}{dt} f.$$

For reasons unknown, in practice, only functions quadratic in p usually appear as Hamiltonians.

The local expression of the *Poisson bracket* is given by the formula

$$\{f, g\}_{P.B.} = \sum \left(\frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} \right). \tag{1.44}$$

For the symplectic manifolds of the form different from T^*N, there might appear another expression for the Poisson bracket that depends on the embedding of the symplectic manifold M into an ambient space. For example, on $M = S^2$ singled out by the formula

$$x^2 + y^2 + z^2 = R^2 \text{ in } \mathbb{R}^3,$$

[7] For a conceptual (algebraic) proof of this seemingly analytic statement, see [LPS, GL] and volume 2. The obstructions to representability of a given non-degenerate differential 2-form ω in a canonical form not just at a given point, but in an infinitesimal neighborhood of this point, is $d\omega$. This fact is easy to see using methods of volume 2 and is absolutely mysterious otherwise. Having understood this, we make the Darboux theorem a tautology.

the Poisson bracket can be expressed globally as

$$\{f,g\}_{P.B.} = \frac{z}{R}\left(\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}\right) + \frac{y}{R}\left(\frac{\partial f}{\partial z}\frac{\partial g}{\partial x} - \frac{\partial f}{\partial x}\frac{\partial g}{\partial z}\right) + \frac{x}{R}\left(\frac{\partial f}{\partial y}\frac{\partial g}{\partial z} - \frac{\partial f}{\partial z}\frac{\partial g}{\partial y}\right). \tag{1.45}$$

The variation of the state of the dynamical system in question is given by substituting p_i and q_i for f into (1.43). Accordingly, inserting the coordinate functions considered to be depending on t into (1.43) with the bracket given by formula (1.45) instead of (1.44), where H is the *Hamiltonian*, or *energy* of the system, we obtain a description of movements of a solid rod of length R with a fixed endpoint.

11.3. Classification of elementary mechanical systems. A mechanical system (M, ω, H) with a symmetry group G is said to be *elementary* if G acts *transitively* on M, i.e., if, for any two points $a, b \in M$, there exists $g \in G$ such that $a = gb$.

The following[8)] construction classifies the elementary mechanical systems for the case when G is any Lie group.

Let $\mathfrak{g} = \mathrm{Lie}(G)$. Then G naturally acts on $\mathfrak{g}^*$ via Ad^*. From the Lie group theory ([OV]) it is known that each orbit Ω of the G-action on $\mathfrak{g}^*$ is a smooth manifold. Let $\Omega = G\varphi$ for some $\varphi \in \mathfrak{g}^*$ and let G_φ be the *stationary subgroup* of G, i.e., a subgroup that preserves φ. Set

$$\mathfrak{g}_\varphi = \mathrm{Lie}(G_\varphi).$$

Clearly, $\mathfrak{g}/\mathfrak{g}_\varphi \simeq T_\varphi\Omega$.

Recall that two dynamical systems (M, ω, H) and (M', ω', H') are considered to be *equivalent* if there exists a *symplectomorphism* $F\colon M \longrightarrow M'$, i.e., a diffeomorphism such that

$$\begin{aligned}&F(\omega) = \omega' \text{ and } F(H) = H'; \text{ moreover,}\\ &F \text{ is invertible, } F^{-1}(\omega') = \omega \text{ and } F^{-1}(H') = H.\end{aligned}$$

11.3.1. Statement ([Ki1])**.** *The kernel of the bilinear form*

$$\omega\colon X, Y \mapsto \varphi([X,Y]) \text{ for any } X, Y \in \mathfrak{g} \text{ and a fixed } \varphi \in \mathfrak{g}^*$$

is equal to $\mathfrak{g}_\varphi$ and ω is non-degenerate on $\mathfrak{g}/\mathfrak{g}_\varphi$. The collection of forms ω for φ running over Ω determines a symplectic structure on Ω. Two orbits Ω and Ω' determine equivalent dynamical systems (M, ω, H) and (M', ω', H') for a G-invariant $\varphi \in \mathfrak{g}^*$.

Let ω_Ω be the symplectic form thus defined on the orbit Ω.

[8] Due to Kirillov.

11.3.2. Generalizations. The above construction of elementary systems can be generalized:

1) Let G_1 be a central extension of G with the help of the Lie group Z; in other words, Z belongs to the center of G_1 and $G_1/Z = G$. If Ω is a G_1-orbit in $\mathfrak{g}_1^*$, then Z trivially acts on Ω, and therefore the actual symmetry group of Ω is $G_1/Z = G$.

2) If M is a homogeneous space with the G-action (i.e., $M = G/H$ for a Lie subgroup $H \subset G$) endowed with a symplectic structure, and Γ is a discrete subgroup of *canonical transformations* of M (i.e., Γ preserves the symplectic structure) and the G-action on M commutes with the Γ-action, then $M' = M/\Gamma$ is also a homogeneous symplectic manifold with symmetry group G.

3) If M is a non-simply connected symplectic manifold homogeneous with respect to the Lie group G, then the simply connected cover $\widetilde{M}$ of M is a symplectic manifold homogeneous with respect to $\widetilde{G}$, the simply connected cover of G.

11.3.3. Exercise. The dynamical systems obtained from each other by the mutually inverse procedures 2) and 3) are equivalent.

11.4. Digression: Lie derivative. Recall that the *Lie derivative* along the vector field D on $\Omega^{\boldsymbol{\cdot}}(M)$, the exterior algebra of differential forms, is defined by the formulas

$$L_D(f) = D(f); \quad L_D(df) = d(D(f)) \text{ for any } f \in \mathcal{F}(M)$$

and the Leibniz rule:

$$L_D(\omega_1 \wedge \omega_2) = L_D(\omega_1) \wedge \omega_2 + \omega_1 \wedge L_D(\omega_2).$$

The vector fields $D \in \mathfrak{vect}(M)$ satisfying (1.46), i.e., such that $L_D(\omega) = 0$, are called *Hamiltonian* vector fields.

11.4.1. Exercise. Let ω be the symplectic form on the symplectic manifold M. Verify that Hamiltonian vector fields on M constitute a Lie algebra $\mathfrak{h}(M)$, i.e., if $L_{D_1}(\omega) = 0$ and $L_{D_2}(\omega) = 0$, then $L_{[D_1,D_2]}(\omega) = 0$.

To describe the Hamiltonian fields, consider them as elements of $\Omega^1(M)$. For the definition of the inner derivative, see any textbook on Differential Geometry.

11.4.2. Digression. Running somewhat ahead, we give the following transparent definition of the inner derivative in terms of supermanifolds; for details, see [L] and volume 2 of these lectures. In coordinates $x = (q, p)$ and $\hat{x} := dx$ on the supermanifold $(M, \Omega(M))$, the inner derivative along the vector field $D \in \mathfrak{vect}(M)$ is of the form

$$\iota_D := \sum f_i(x)\partial_{\hat{x}_i} \text{ for any } D = \sum f_i(x)\partial_{x_i}.$$

Since ω is non-degenerate, the following equation for α_D

$$\alpha_D(X) = \omega(D, X) \text{ for any } X \in \mathfrak{vect}(M)$$

has a unique solution $\alpha_D \in \Omega^1(M)$. Explicitly:

$$\alpha_D = \iota_D(\omega), \text{ where } \iota_D \text{ is the inner derivative by } D.$$

Since d is $\mathfrak{vect}(M)$-invariant, i.e., $L_D(d\omega) = dL_D(\omega)$, it follows that D is a Hamiltonian vector field if and only if ω is closed, i.e., $d\omega = 0$. On every symplectic manifold, this is true by definition. In particular, to each function f there corresponds a Hamiltonian vector field D_f such that

$$\alpha_{D_f} = df.$$

Such a field D_f is called (strictly) *Hamiltonian* one and f is its *generating function.* Locally, in Darboux canonical coordinates, (i.e., if $\omega = \sum dp_i \wedge dq_i$), we have

$$D_f = \sum \left(\frac{\partial f}{\partial p_i} \frac{\partial}{\partial q_i} - \frac{\partial f}{\partial q_i} \frac{\partial}{\partial p_i} \right).$$

Since $[D_f, D_g]$ is again a Hamiltonian field, so it has a generating function that we denote by $\{f, g\}_{P.B.}$.

11.4.3. Exercise. The explicit formula for the Poisson bracket is given by formula (1.44) above.

11.4.4. Remark. Generally, the form α_{D_f} is of the form *df* only locally; the quotient space of all Hamiltonian functions modulo strictly Hamiltonian ones is isomorphic to $H^1(M; \mathbb{R})$.

11.5. Theorem. *Examples of elementary dynamical systems with the symmetry group G are exhausted, up to equivalence, by the orbits introduced in Lemma.*

Proof (A sketch). Let (M, ω) be a symplectic manifold with a transitive action of a Lie group G. To every $X \in \mathfrak{g}$ there corresponds a Hamiltonian vector field H_X, i.e., a vector field on M such that

$$L_{H_X}(\omega) = 0, \tag{1.46}$$

where L_D is the Lie derivative along the vector field D. □

11.6. The moment map. The map $m\colon M \longrightarrow \mathfrak{g}^*$ given by the formula

$$m(x)(X) = f_{H_X}(x), \text{ where } H_X = D_{f_{H_X}} \text{ for any } X \in \mathfrak{g},$$

is called the *moment map.* Here f_{H_X} is the generating function for the Hamiltonian field H_X determined by $\mathfrak{g}$.

Lemma. *Locally, the moment map is one-to-one; it commutes with the G-action, sends M into an orbit $\mathcal{O}$ of G in $\mathfrak{g}^*$ and the form ω on M into the form $\omega_{\mathcal{O}}$ constructed in Statement 11.3.1. More exactly, $\omega = m^*(\omega_{\mathcal{O}})$.*

Proof. Since M is a homogeneous G-manifold, it follows that, for every $x \in M$, there exists a collection of vector fields H_{X_i}, that form a basis of T_xM. (This means that we can shift the point x in any direction.) Then the set of differentials $d(f_{X_i})$ constitutes a basis of T^*_xM. By the inverse function theorem, the collection of functions f_{X_i} can be taken for the local coordinates on M in a neighborhood of x. Hence, the map m is, locally, one-to-one. □

Suppose the generating functions satisfy

$$f_{aX+bY} = af_X + bf_Y, \quad f_{[X,Y]} = \{f_X, f_Y\}_{P.B.} \text{ for any } X, Y \in \mathfrak{g}. \tag{1.47}$$

The comparison of (1.47) with the coadjoint action Ad^* implies that the moment map commutes with the G-action.

Since M is a homogeneous manifold and m commutes with the G-action, we deduce that $m(M)$ is a G-orbit in $\mathfrak{g}^*$.

The assertion about the image of $\omega_{\mathcal{O}}$ under m^* is subject to a direct verification. (The reader may take it as an Exercise.)

Thus, Theorem 11.4 is proved provided (1.47) holds. Generally, only a weaker condition takes place:

$$H_{aX+bY} = af_{H_X} + bf_{H_Y}, \quad H_{[X,Y]} = [H_X, H_Y] \text{ for any } X, Y \in \mathfrak{g}. \tag{1.48}$$

The weaker conditions (1.48) imply that the stronger conditions (1.47) are only satisfied modulo center of the Poisson algebra. Therefore, the totality of all generating functions for all fields H_X, where $X \in \mathfrak{g}$, constitutes a Lie algebra $\widehat{\mathfrak{g}}$ whose quotient modulo the center spanned by the constants is isomorphic to $\mathfrak{g}$, i.e., the following sequence is exact:

$$0 \longrightarrow \mathfrak{z} \longrightarrow \widehat{\mathfrak{g}} \longrightarrow \mathfrak{g} \longrightarrow 0.$$

So, in the general case, replace $\mathfrak{g}$ with $\widehat{\mathfrak{g}}$ and G with the group $\widehat{G}$ corresponding to $\widehat{\mathfrak{g}}$ and apply Statement 11.3.1. Theorem 11.5 is proved in full generality. □

Example. Let $G = \mathbb{R}^2$, for $\widehat{G}$ take the group of matrices of the form

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} = (x, y, z)$$

with multiplication

$$(x, y, z)(a, b, c) = (x + a, y + b, z + c + xb)$$

Clearly, $Z = \left\{ \begin{pmatrix} 1 & 0 & z \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}$ and $\widehat{G}/Z \simeq G$. In appropriate coordinates p, q, r the $\widehat{G}$-action on $\widehat{\mathfrak{g}}^*$ is given by the formula

$$\mathrm{Ad}^*_{(x,y,z)}(p,q,r) = (p + rx, q + ry, r).$$

The respective vector fields are

$$X = r\frac{\partial}{\partial p}, \quad Y = r\frac{\partial}{\partial q}, \quad Z = 0.$$

The orbits are the planes $r = const \neq 0$ and the points $(p, q, 0)$. On the planes $r = const \neq 0$ the symplectic form is as follows:

$$\omega = \frac{1}{r} dp \wedge dq.$$

What are the discrete transformation groups of $\mathbb{R}^2$ commuting with all the translation? There are several types of such groups:

(a) the trivial one (the identity transformation);

(b) the group generated by translations by a fixed vector $v \in \mathbb{R}^2$;

(c) the group generated by translations in two linearly independent vectors $v_1, v_2 \in \mathbb{R}^2$;

(d) What is the group of the fourth type?

The corresponding elementary systems are:

(a) $\mathbb{R}^2$,
(b) $\mathbb{R} \times S^1$ (cylinder),
(c) $T^2 = S^1 \times S^1$,
(d) $*$ (one point).

The systems (a), (b), (d) are of the form T^*M, where the configuration space M is a line, a circle and a point, respectively.

2) $G = SO(3)$. It is known (see, e.g., Kostrikin's books [Ko]) that every central extension of $SO(3)$ is the direct product of either $SO(3)$ or its two-sheeted covering $SU(2)$ by the one-dimensional commutative Lie group Z. Therefore, the passage to $\widehat{G}$ does not lead to new examples of the dynamical systems and the G-action in $\mathfrak{g}^*$ coincides with the identity G-action id in the 3-dimensional space. (Indeed, the G-actions in $\mathfrak{g}^*$ and $\mathfrak{g}$ coincide thanks to the Killing form, the G-module $\mathfrak{g}$ is irreducible and G has only one irreducible module in every odd dimension, so $\mathfrak{g} \simeq \mathfrak{g}^* \simeq \mathrm{id}$.) The orbits $\mathcal{O}$ are the origin, and the spheres with the center in the origin with ω proportional to the element of area.

There is only one nontrivial transformation of the sphere that commutes with rotations, namely, the central symmetry. Thus, the elementary systems for $G = SO(3)$ are

a) S^2,
b) $\mathbb{R}P^2$ (the real projective space),
c) $*$ (one point).

References

[Ad] Adams J. Frank, *Lectures on Lie groups.* W. A. Benjamin, Inc., New York-Amsterdam 1969 xii+182 pp.

[Ar] Arnold V. *Mathematical methods of classical mechanics.* Translated from the 1974 Russian original by K. Vogtmann and A. Weinstein. Corrected reprint of the second (1989) edition. Graduate Texts in Mathematics, 60. Springer, New York, 199?. xvi+516 pp.

[Ar0] Arnold V. On evolution of magnetic field under the action of transition and diffusion. In: V. Tikhomirov (ed.) *Selected problems of modern calculus*, Moscow State Univ. Press, Moscow, 1984, 8– 21 (In Russian)

[B] Berezin F. *Introduction to algebra and analysis with anticommuting variables.* Moscow University, Moscow, 1983 (enlarged and revised English translation: *Introduction to superanalysis.* Edited and with a foreword by A. A. Kirillov. With an appendix by V. I. Ogievetsky. Translated from the Russian by J. Niederle and R. Kotecký. Translation edited by D. Leites. Mathematical Physics and Applied Mathematics, 9. D. Reidel Publishing Co., Dordrecht, 1987. xii+424 pp.)

[CR] Curtis C., Reiner I., *Methods of representation theory with applications to finite groups and orders. Vol. I.* Reprint of the 1981 original. Wiley Classics Library. A Wiley-Interscience Publication. John Wiley and Sons, Inc., New York, 1990. xxiv+819 pp.
id, ibid, *Vol. II.* Pure and Applied Mathematics. A Wiley-Interscience Publication. John Wiley and Sons, Inc., New York, 1987. xviii+951 pp; id., *Representation theory of finite groups and associative algebras.* Reprint of the 1962 original. Wiley Classics Library. A Wiley-Interscience Publication. John Wiley and Sons, Inc., New York, 1988. xiv+689 pp.

[C1] Curtis, Ch. *Representation theory of Hecke algebras.* Notes taken by Peter Fleischmann. With an introduction by Fleischmann. Vorlesungen aus dem Fachbereich Mathematik der Universität Essen [Lecture Notes in Mathematics at the University of Essen], 15. Universität Essen, Fachbereich Mathematik, Essen, 1987. iv+86 pp.

[C2] Conway, J. H.; Curtis, R. T.; Norton, S. P.; Parker, R. A.; Wilson, R. A. *Atlas of finite groups. Maximal subgroups and ordinary characters for simple groups.* With computational assistance from J. G. Thackray. Oxford University Press, Oxford, 1985. xxxiv+252 pp.

[D1] Dynkin E.B. Semisimple subalgebras of semisimple Lie algebras. Matem. Sbornik, 1952, v.30 (72), 349–462 (Russian) = English transl. in Moscow Math. Soc. Translations Ser. 2, v.6, 111–244

[D2] Dynkin E.B., The review of principal concepts and results of the theory of linear representations of semi-simple Lie algebras. Complement to the paper: Dynkin E.B. Maximal subgroups of classical groups. Trudy Moskov. Mat. Obšč. 1, (1952). 39–166 (Russian = English transl. in Moscow Math. Soc. Translations Ser. 2, v.6, 245–378)

[E] Etingof, P. (notes taken by Oleg Golberg, Sebastian Hensel, Tiankai Liu, Alex Schwendner, Elena Udovina, Dmitry Vaintrob) *Introduction to representation theory.* `arXiv:0901.0827`

[FH] Fulton, W., Harris, J., *Representation theory. A first course.* Graduate Texts in Mathematics, 129. Readings in Mathematics. Springer, New York, 1991. xvi+551 pp

[Go] Gorenstein D., *Finite simple groups. An introduction to their classification.* University Series in Mathematics. Plenum Publishing Corp., New York, 1982. x+333 pp.

[G] Gorenstein D., *The classification of finite simple groups. Vol. 1. Groups of noncharacteristic* 2 *type.* University Series in Mathematics. Plenum Publishing Corp., New York, 1983. x+487 pp.; Gorenstein, Daniel; Lyons, Richard; Solomon, Ronald *The classification of the finite simple groups.* Mathematical Surveys and Monographs, 40.1. American Mathematical Society, Providence, RI, 1994. xiv+165 pp.;
id. *The classification of the finite simple groups. Number 2. Part I. Chapter G. General group theory.* Mathematical Surveys and Monographs, 40.2. American Mathematical Society, Providence, RI, 1996. xii+218 pp.
id., *The classification of the finite simple groups. Number 3. Part I. Chapter A. Almost simple K-groups.* Mathematical Surveys and Monographs, 40.3. American Mathematical Society, Providence, RI, 1998. xvi+419 pp.

[GL] Grozman P., Leites D., The nonholonomic Riemann and Weyl tensors for flag manifolds, Theor. and Mathem. Phys. v. 152(2), 2007, 1511–1538 `arXiv:math.DG/0509399`

[K] Kac V. *Infinite-dimensional Lie algebras.* Third edition. Cambridge University Press, Cambridge, 1990. xxii+400 pp.

[Ki] Kirillov A. A., *Lectures on representation theory of finite and compact groups.* I–IV, Moscow Univ. Press, Moscow, 1970 (Preprints in Russian)

[Ki1] Kirillov, A. A. *Elements of the theory of representations.* Translated from the Russian by Edwin Hewitt. Grundlehren der Mathematischen Wissenschaften, Band 220. Springer, Berlin-New York, 1976. xi+315 pp

[Ko] Kostrikin, A. I. *Introduction to algebra.* Translated from the Russian by Neal Koblitz. Universitext. Springer, New York-Berlin, 1982. xiii+575 pp.; Kostrikin, A. I. (ed.) *Exercises in algebra.* A collection of exercises in algebra, linear algebra and geometry. Expanded English edition. Translated from the Russian by V. A. Artamonov. Edited by Alexei I. Kostrikin. Algebra, Logic and Applications, 6. Gordon and Breach Science Publishers, New York, 1996. xii+464 pp.

[L] Leites D., Introduction to the supermanifold theory. Russian Math. Surveys, v. 35, n.1, 1980, 3–53; *Supermanifold theory*, Karelia Branch of the USSR Acad. of Sci., Petrozavodsk, 1983, 200 pp. (in Russian)

[LPS] Leites D., Poletaeva, E., Serganova V., On Einstein equations on manifolds and supermanifolds, J. Nonlinear Math. Physics, v. 9, 2002, no. 4, 394–425; `arXiv:math.DG/0306209`

[OV] Onishchik A. L., Vinberg È. B., *Seminar on algebraic groups and Lie groups*, Springer, 1990

[P] Prasolov V., *Problems and theorems in linear algebra*, AMS, 1991, 226 pp.

[Se] Serre J.-P., *Linear representations of finite groups*, Springer, 1993

[Si] Simon B., *Representation of finite and compact groups*, American Mathematical Society, 1996

[St] Steinberg, R., *Lectures on Chevalley groups.* Notes prepared by John Faulkner and Robert Wilson. Yale University, New Haven, Conn., 1968. iii+277 pp.

[V] Vinberg È. B., *Representations of compact groups*, Springer, 1990;
id., *Linear representations of groups.* Translated from the Russian by A. Iacob. Basler Lehrbücher [Basel Textbooks], 2. Birkhäuser, Basel-Boston, MA, 1989. iv+146 pp.

Chapter 2

Lectures on Lie algebras (J. Bernstein)

Summary This is a lecture course for beginners on representation theory of simple finite dimensional complex Lie algebras. It is shown how to use infinite dimensional representations (Verma modules) to derive the Weyl character formula. The Harish–Chandra and Kostant's theorems on the center of the enveloping algebra are given with lucid (and shortest known) proof. This proof of the Harish–Chandra theorem was instrumental in A. Sergeev's version of the Harish–Chandra theorem for Lie superalgebras.

Introduction

We [1)] will consider finite dimensional representations of semi-simple finite dimensional complex Lie algebras. The facts presented here are well-known (see [Bu], [Di], [S]) and in a more rigorous setting. But our presentation of

---

[1] These notes, first preprinted in Proceedings of my *Seminar on supersymmetries* at the Department of Mathematics, Stockholm University, originate from a draft of the transcript of J. Bernstein's lectures in the Summer school in Budapest in 1971. For the lectures addressed to the advanced part of the audience, see Gelfand I. (ed.), *Representations of Lie groups and Lie algebras*, Acad. Kiado, Budapest, 1975. The beginners' part was released a bit later (Kirillov A. (ed.), *Representations of Lie groups and Lie algebras*, Acad. Kiado, Budapest, 1985). It contains a review by Feigin and Zelevinsky which *expands and expounds* Bernstein's lectures but with (almost totally) different emphasis and has practically no intersection with Bernstein's lectures.

Meanwhile there appeared several text books and reference books on Lie algebra representations, some of them absolutely marvelous each in its way ([OV], [FH]). However, even the best ([FH], [OV]) do not approach finite dimensional representations by means of infinite dimensional ones (but of simpler structure) — one of the main "punch lines" of **these** lectures.

Besides, these excellent text books are **thick books**. Therefore, the demand in a short and informal guide for the beginners still remains, we were repeatedly told. So here it is.

these facts is comparatively new (at least, it was so in 1971) and is based on the systematic usage of the Verma modules M^χ. In this it cardinally differs from any other of the modern text book devoted to the study of finite dimensional representations of Lie algebras, where infinite dimensional representations are being avoided.

The reader will see how manifest the Weyl character formula is in terms of (infinite dimensional) Verma modules.

The reader is supposed to be acquainted with the main notions of Linear Algebra ([W] or a more recent [P] will be just fine). The knowledge of the first facts and notions of Lie algebra theory will not hurt but is not required.

The presentation is arranged as follows:

§ 2.0 provides without proof with results on the structure of the semi-simple complex Lie algebras and their universal enveloping algebras.

§ 2.1 introduces the category $\mathcal{O}$ and the Verma modules M^χ; several of their properties are listed.

In § 2.2, the highest (lowest) weight representations of the simplest simple Lie algebra $\mathfrak{sl}(2)$ are described.

In § 2.3, for every semi-simple Lie algebra $\mathfrak{g}$, a supply of irreducible finite dimensional representations ($\mathfrak{g}$-modules) is given.

In § 2.4, the Harish–Chandra theorem on the center $Z(U(\mathfrak{g}))$ of the enveloping algebra of $\mathfrak{g}$ is formulated. For the proof, see § 2.8.

In § 2.5, certain properties of the category $\mathcal{O}$ that follow from the Harish–Chandra theorem are derived.

In § 2.6, we prove that every finite dimensional $\mathfrak{g}$-module is decomposable into the direct sum of modules built in § 2.4.

In § 2.7, Kostant's formula for the multiplicities of weights of an irreducible representation is proved and the Weyl formula for the dimension of the irreducible representation is derived from it.

In § 2.8, the Harish–Chandra theorem is proved.

2.0. Preliminaries

For the proof of the statements of this section, see [Bu], [Di], [S], [OV].

All vector spaces considered in what follows are defined over a ground field $\mathbb{K}$. Unless otherwise stated $\mathbb{K} = \mathbb{C}$, but the reader should be aware of the fact that various "practical"problems often force one to consider other possibilities: $\mathbb{K} = \mathbb{R}$, $\mathbb{Q}$, $\mathbb{F}_q$, and $\overline{\mathbb{F}_q}$, the algebraic closure of $\mathbb{F}_q$; even the case of a ring instead of the field: $\mathbb{K} = \mathbb{Z}$; some other rings, like $\mathbb{K} = \mathbb{C}[t^{-1}, t]$ or its completion, $\mathbb{C}[t^{-1}][[t]]$, are often encountered. But not in these lectures; here by default $\mathbb{K} = \mathbb{C}$.

2.0.1. Lie algebras. Let A be an associative algebra, e.g., $A = \mathrm{Mat}(n; C)$ the associative algebra of $n \times n$ matrices over a commutative algebra C. By means of the subscript L we will denote another algebra $\mathfrak{g} = A_L$ whose space is a copy of A and the product in A_L is given by the bracket $[x, y] = xy - yx$. It is subject to a direct (though somewhat boring) verification that the bracket satisfies the following identities:

$$[x, x] = 0 \quad \text{for any} \quad x \in \mathfrak{g}; \tag{2.1}$$

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0 \quad \text{for any } x, y, z \in \mathfrak{g}. \tag{2.2}$$

The identity (2.1) signifies anti-symmetry of the bracket, (2.2) is called the *Jacobi identity.* (To understand where the Jacobi identity comes from, we have to turn to Lie groups and "differentiate" the associativity law of group multiplication.) The Jacobi identity can be expressed in an equivalent form easy to remember (ad_x is a derivation):

$$[x, [y, z]] = [[x, y], z] + [y, [x, z]] \quad \text{for any } x, y, z \in \mathfrak{g}.$$

This example leads to a notion of Lie algebra which turned out to be very important. Namely, a *Lie algebra* is a vector space $\mathfrak{g}$ with multiplication $\mathfrak{g} \otimes \mathfrak{g} \longrightarrow \mathfrak{g}$ called *bracket* and is usually denoted by $[x, y]$ or $\{x, y\}$, not xy, and which satisfies (2.1) and (2.2).

Anti-symmetry and bilinearity of $[\cdot, \cdot]$ implies

$$0 = [x + y, x + y] = [x, y] + [y, x],$$

i.e., $[x, y]$ is indeed anti-symmetric, but equation (2.1) is easier to verify; besides, it holds even if the characteristic of $\mathbb{K}$ is equal to 2.

2.0.2. Examples of Lie algebras. After Bourbaki, it is now customary to denote Lie algebras by Gothic letters. The prime example is the Lie algebra $\mathfrak{gl}(n; C) = \mathrm{Mat}(n; C)_L$, the *general linear* algebra of $n \times n$ matrices over a commutative algebra C.

More generally, let $\mathrm{End}\, V$ be the associative algebra of endomorphisms of a vector space V. The *general linear algebra* of the vector space V is $\mathfrak{gl}(V) = (\mathrm{End}\, V)_L$. Having selected a basis of the n-dimensional space V, we identify $\mathfrak{gl}(V)$ with $\mathfrak{gl}(n; C)$.

Particular cases of the above passage $A \mapsto A_L$ are especially important when A is the algebra of differential operators or pseudo-differential operators.

Let $\mathfrak{g}^+$, $\mathfrak{g}^-$ and $\mathfrak{h}$ be the subsets of $\mathfrak{g} = \mathfrak{gl}(n)$ consisting of all strictly upper triangular, strictly lower triangular or diagonal matrices, respectively. Clearly, $\mathfrak{g}^+$, $\mathfrak{g}^-$ and $\mathfrak{h}$ are Lie algebras.

The main property of the *trace* on a given Lie algebra $\mathfrak{g}$ is that it vanishes on the commutators,

$$\mathrm{tr}[x, y] = 0 \quad \text{for any } x, y \in \mathfrak{g}.$$

By the main property of traces the subspace of traceless (i.e., with trace 0) elements forms a Lie subalgebra of $\mathfrak{g}$. In particular, the space of $n \times n$ matrices with trace 0 is closed under the bracket, i.e., is a Lie algebra; it is called the *special linear algebra* and denoted by $\mathfrak{sl}(n)$.

Let B be a bilinear form on a vector space V. It is easy to verify that the space $\mathfrak{aut}_B(V)$ of all operators that preserve B, i.e.,

$$\mathfrak{aut}_B(V) := \{x \in \mathfrak{gl}(V) \mid B(xu, v) + B(u, xv) = 0 \text{ for any } u, v \in V\}$$

is closed under the bracket and so $\mathfrak{aut}_B(V)$ is a Lie algebra. If B is non-degenerate, we distinguish two important subcases:

B is symmetric, then $\mathfrak{aut}_B(V)$ is called the *orthogonal Lie algebra* and denoted by $\mathfrak{o}_B(V)$.

B is anti-symmetric, then $\mathfrak{aut}_B(V)$ is called the *symplectic Lie algebra* and denoted by $\mathfrak{sp}_B(V)$.

It is well-known (see [P]) that, over $\mathbb{C}$, all non-degenerate symmetric forms on V are equivalent to each other and all non-degenerate anti-symmetric forms are equivalent to each other. So Lie algebras $\mathfrak{o}_B(V)$ and $\mathfrak{sp}_B(V)$ depend, actually, only on dim V, and we will sometimes denote them by $\mathfrak{o}(n)$ and $\mathfrak{sp}(2m)$.

The Lie algebras $\mathfrak{gl}(n)$, $\mathfrak{o}(n)$ and $\mathfrak{sp}(2m)$, as well their real forms, are called *classical.*

Let $\mathfrak{g}$ be a (not necessarily finite dimensional) Lie algebra. Having realized it by operators, we often see that, though the brackets of the operators is defined within $\mathfrak{g}$, their product seldom belongs to $\mathfrak{g}$. The desire to have an algebra, inside of which we could have the product of these operators as well, leads to the following definition.

To $\mathfrak{g}$, we assign the *associative* algebra with unit, $U(\mathfrak{g})$, called the *universal enveloping* algebra of the Lie algebra $\mathfrak{g}$. To this end, consider the tensor algebra $T(\mathfrak{g})$ of the *space* $\mathfrak{g}$, i.e.,

$$T^{\bullet}(\mathfrak{g}) = \underset{n\geq 0}{\oplus} T^n(\mathfrak{g}),$$

where $T^0(\mathfrak{g}) = \mathbb{C}$, $T^1(\mathfrak{g}) = \mathfrak{g}$, $T^n(\mathfrak{g}) = \underbrace{\mathfrak{g} \otimes \cdots \otimes \mathfrak{g}}_{n>1 \text{ factors}}$ and the (usually suppressed) dot in the superscript denotes the sum over possible values of the superscript. Set

$$U(\mathfrak{g}) = T(\mathfrak{g})/I,$$

where I is the two-sided ideal generated by

$$x \otimes y - y \otimes x - [x, y] \text{ for any } x, y \in \mathfrak{g}.$$

Recall that a *module* over an algebra A is a vector space M together with a bilinear map called *action*

$$a\colon A \otimes M \longrightarrow M, \quad \alpha(a, m) := am$$

such that for any $a, b \in A$ and $m \in M$, we have

$$\begin{array}{ll} (ab)m = a(bm) & \text{if } A \text{ is an associative algebra,} \\ [a, b]m = a(bm) - b(am) & \text{if } A \text{ is a Lie algebra.} \end{array}$$

To define the action $a\colon \mathfrak{g}\otimes M \longrightarrow M$ is the same as to define a homomorphism $\rho\colon \mathfrak{g} \longrightarrow \mathfrak{gl}(M)$ called a *representation* of $\mathfrak{g}$ in M. Instead of "the $\mathfrak{g}$-module V" we will often say "the representation (say, ρ) of $\mathfrak{g}$ in the linear space V" and instead of writing $\rho(g)v$ will write just gv for any $g \in \mathfrak{g}$ and $v \in V$.

We will identify the elements of $\mathfrak{g}$ with their images in $U(\mathfrak{g})$. Under this identification, any $\mathfrak{g}$-module may be considered as a (left, unital) $U(\mathfrak{g})$-module and, conversely, any $U(\mathfrak{g})$-module may be considered as a $\mathfrak{g}$-module. We will not distinguish the $\mathfrak{g}$-modules from the corresponding $U(\mathfrak{g})$-modules.

The algebra $U(\mathfrak{g})$ is easy to describe if $\mathfrak{g}$ is commutative: Then $U(\mathfrak{g})$ is just the symmetric algebra $S^{\bullet}(\mathfrak{g})$ of the space $\mathfrak{g}$. In the general case, when $U(\mathfrak{g})$ is not commutative, it is not clear what is the "size" of $U(\mathfrak{g})$. It is highly nontrivial that, considered as spaces, not algebras, of course, $U(\mathfrak{g})$ and $S^{\bullet}(\mathfrak{g})$ may be identified. To actually identify them, we introduce a filtration in $U(\mathfrak{g})$ by setting

$$U(\mathfrak{g})_n = \Big(\bigoplus_{i\leq n} T^i(\mathfrak{g})\Big) \mod I.$$

The associated graded algebra $\operatorname{gr} U(\mathfrak{g}) = \bigoplus\limits_{n\geq 0} (U(\mathfrak{g})_n/U(\mathfrak{g})_{n-1})$ is, clearly, commutative.

Theorem (Poincaré–Birkhoff–Witt)**.** *The natural embedding* $\mathfrak{g} \longrightarrow \operatorname{gr} U(\mathfrak{g})$ *can be extended to an isomorphism* $i\colon S^{\bullet}(\mathfrak{g}) \longrightarrow \operatorname{gr} U(\mathfrak{g})$ *of graded algebras.*

Corollary. 1) $U(\mathfrak{g})$ *is a noetherian ring without zero divisors.*

2) *Let* $\operatorname{symm}'\colon S^{\bullet}(\mathfrak{g}) \longrightarrow T^{\bullet}(\mathfrak{g})$ *be the map determined by the formula*

$$\operatorname{symm}'\colon X_1 \circ \cdots \circ X_k \mapsto \frac{1}{k!}\sum_{\sigma\in\mathfrak{S}_k} X_{\sigma(1)} \otimes \cdots \otimes X_{\sigma(k)}, \tag{2.3}$$

where $\circ$ *is the symmetric multiplication in* $S^{\bullet}(\mathfrak{g})$. *Denote by*

$$\operatorname{symm}\colon S^{\bullet}(\mathfrak{g}) \longrightarrow T^{\bullet}(\mathfrak{g}) \longrightarrow U(\mathfrak{g})$$

the composition of symm' *and the projection onto* $U(\mathfrak{g})$. *The map* symm *is an isomorphism of linear spaces (but not algebras).*

3) *If* $X_1, \dots, X_k$ *is a basis of* $\mathfrak{g}$, *then* $X_1^{n_1}\dots X_k^{n_k}$, *where the* n_i *run over the set* $\mathbb{Z}_+$ *of nonnegative integers, is a basis of* $U(\mathfrak{g})$.

Let ad denote the *adjoint* representation of the Lie algebra $\mathfrak{g}$, i.e., $\operatorname{ad}_X(Y) = [X,Y]$ for $X, Y \in \mathfrak{g}$. The *Killing* form is the bilinear form on $\mathfrak{g}$ given by the formula

$$(X,Y)_{\mathrm{ad}} = \operatorname{tr}(\operatorname{ad}_X \cdot \operatorname{ad}_Y).$$

For any irreducible representation ρ of a given Lie algebra $\mathfrak{g}$, one can similarly consider the *Killing-like* form

$$(X,Y)_\rho = \operatorname{tr}(\rho(X)\rho(Y)).$$

Theorem. *The space of invariant non-degenerate symmetric bilinear forms on any finite dimensional simple Lie algebra is of dimension* 1.

The proportionality coefficient l_ρ in the relation $(X,Y)_\rho = l_\rho(X,Y)_{\mathrm{ad}}$ is called the *Dynkin coefficient* in honor of Dynkin who computed it for all irreducible representation ρ of all finite dimensional simple Lie algebras. The Killing-like forms for some representations are more convenient for computing than he Killing form: For example, the identity representations for Lie algebras of the $\mathfrak{sl}$, $\mathfrak{o}$ and $\mathfrak{sp}$ series.

Some data on classical Lie algebras. The Lie algebra $\mathfrak{g}$ is called *semi-simple* if its *radical* is zero. For finite dimensional Lie algebras over $\mathbb{C}$, there is another characterization, often more convenient:

$\mathfrak{g}$ *is semi-simple if its Killing form is non-degenerate.*

The structure of the finite dimensional Lie algebras is as follows: Each algebra is a semidirect sum of what is called *semi-simple Lie algebra* and the radical, more precisely, the quotient modulo the radical is by definition a semi-simple Lie algebra.

Now, over fields of characteristic 0, every finite dimensional semi-simple Lie algebra is the direct sum of simple ones.

There exist commutative subalgebras $\mathfrak{t} \subset \mathfrak{g}$ that diagonally act on $\mathfrak{g}$. Such a subalgebra is called a *toral* one. A maximal nilpotent subalgebra $\mathfrak{h}$ coinciding with its normalizer is called the *Cartan subalgebra*. For finite dimensional semi-simple Lie algebras over $\mathbb{C}$, the maximal toral subalgebras are all conjugate (under the automorphism group) and each of them coincides with a Cartan subalgebra.

The nonzero weights relative a maximal toral subalgebra (or, which is the same, relative a Cartan subalgebra) in the adjoint representation are called *roots*. The finite subset of roots is usually denoted by $R \subset \mathfrak{h}^*$ (here $\mathfrak{h}^*$ is the dual space of $\mathfrak{h}$). The set R has the following properties:

1) $0 \notin R$;

For each $\gamma \in R$, there exists an element $E_\gamma \in \mathfrak{g}$ (called *root vector* or *coroot*) such that

2) $[H, E_\gamma] = \gamma(H)E_\gamma$ for any $H \in \mathfrak{h}$ and $\gamma \in R$;

3) $\mathfrak{g} = \mathfrak{h} \oplus \Big(\bigoplus\limits_{\gamma\in R} \mathbb{C}E_\gamma\Big)$.

The number $r = \dim \mathfrak{h}$ is called the *rank* of $\mathfrak{g}$.

Since the Killing form is non-degenerate on any simple finite dimensional Lie algebra $\mathfrak{g}$, the restriction of it onto $\mathfrak{h}$ is also non-degenerate, so we may use it to identify $\mathfrak{h}$ with $\mathfrak{h}^*$ and to define the bilinear form $\langle\cdot,\cdot\rangle$ on $\mathfrak{h}^*$.

For any $\gamma \in R$, let $H_\gamma \in \mathfrak{h}$ be such that, for any $\chi \in \mathfrak{h}^*$, we have

$$\chi(H_\gamma) = \frac{2\langle\chi,\gamma\rangle}{\langle\gamma,\gamma\rangle}. \tag{2.4}$$

The elements E_γ may be chosen so that

1) If $\gamma_1+\gamma_2 \in R$, then $[E_{\gamma_1}, E_{\gamma_2}] = N_{\gamma_1\gamma_2}E_{\gamma_1+\gamma_2}$, where $N_{\gamma_1\gamma_2}$ is a nonzero integer (for more, sometimes convenient, details, see [St]).

2) If $\gamma_1+\gamma_2 \notin R$ and $\gamma_1+\gamma_2 \neq 0$, then $[E_{\gamma_1}, E_{\gamma_2}] = 0$.

3) $[E_\gamma, E_{-\gamma}] = H_\gamma$.

Certain properties of the root system *R*.

i) If $\gamma \in R$, then $-\gamma \in R$ and $\lambda\gamma \notin R$ for $\lambda \neq \pm 1$.

ii) Denote by $\mathfrak{h}^*_{\mathbb{R}}$ the real subspace in $\mathfrak{h}^*$ generated by elements of R. Then $\mathfrak{h}^* = \mathfrak{h}^*_{\mathbb{R}} + i\mathfrak{h}^*_{\mathbb{R}}$ and the restriction of $\langle\cdot,\cdot\rangle$ onto $\mathfrak{h}^*_{\mathbb{R}}$ is positive definite.

iii) In R, there are subsets $R_+ \supset B = \{\alpha_1, \dots, \alpha_r\}$ such that B is a basis in $\mathfrak{h}$ and any $\gamma \in R$ can be represented in the form $\gamma = \sum\limits_{\alpha\in B} n_\alpha \alpha$, where either $n_\alpha \geq 0$ or $n_\alpha \leq 0$ for all α. Such a set B is called a *base* of R or a *system of simple roots*. Set

$$R_- = -R_+ = \{-\alpha \mid \alpha \in R_+\}.$$

Clearly, $R_+ \cup R_- = R$ and $R_+ \cap R_- = \emptyset$; each root $\gamma \in R_+$ belongs to the set (positive part of the root lattice)

$$Q_+ = \left\{\mu \in \mathfrak{h}^* \mid \mu = \sum_{\alpha\in B} n_\alpha \alpha, \text{ where } n_\alpha \in \mathbb{Z}_+\right\}. \tag{2.5}$$

If $\alpha, \beta \in B$ and $\alpha \neq \beta$, then $\langle\alpha, \beta\rangle \leq 0$; if $\alpha, \beta \in B$, then $\beta - \alpha \notin R$, hence, $[E_\beta, E_{-\alpha}] = 0$.

Note that the choice of sets R_+ is not unique. When R_+ is chosen, B is uniquely fixed: it consists of roots that can not be represented as the sum of two other roots from R_+. In what follows R_+ (hence, B) will be fixed, and we will write $\chi_1 \geq \chi_2$ whenever $\chi_1 - \chi_2 \in Q_+$.

iv) Set

$$\mathfrak{h}^*_{\mathbb{Z}} = \left\{\chi \in \mathfrak{h} \in \mathfrak{h}^* \mid \frac{2\langle\chi, \gamma\rangle}{\langle\gamma, \gamma\rangle} \in \mathbb{Z} \text{ for any } \alpha \in R\right\}. \tag{2.6}$$

Then

a) $R \subset \mathfrak{h}^*_{\mathbb{Z}} \subset \mathfrak{h}^*_{\mathbb{R}}$.

b) If $\chi \in \mathfrak{h}^*$ and $\dfrac{2\langle\chi, \gamma\rangle}{\langle\gamma, \gamma\rangle} \in \mathbb{Z}$ for any $\alpha \in B$, then $\chi \in \mathfrak{h}^*_{\mathbb{Z}}$.

c) $\mathfrak{h}^*_{\mathbb{Z}}$ is a complete lattice in $\mathfrak{h}^*_{\mathbb{R}}$.

d) Set

$$\rho = \frac{1}{2}\sum_{\gamma\in R_+} \gamma.$$

Then $\rho(H_\alpha) = 1$ for any $\alpha \in B$.

From this description of the root system R we derive the decomposition of $\mathfrak{g}$ as $\mathfrak{g} = \mathfrak{g}_- \oplus \mathfrak{h} \oplus \mathfrak{g}_+$, where $\mathfrak{g}_\pm$ is the subspace spanned by E_γ for $\gamma \in R_\pm$. Clearly,

A) $\mathfrak{g}_\pm$ is the Lie subalgebra of $\mathfrak{g}$ generated by E_α, where $\alpha \in \pm B$.

B) $[\mathfrak{h}, \mathfrak{g}_\pm] = \mathfrak{g}_\pm$.

C) The Lie algebras $\mathfrak{g}_+$ and $\mathfrak{g}_-$ are nilpotent. Moreover, if $X \in \mathfrak{g}_\pm$, then ad_X is a nilpotent operator on the whole $\mathfrak{g}$.

D) $U(\mathfrak{g}) \simeq U(\mathfrak{g}_- \oplus \mathfrak{h} \oplus \mathfrak{g}_+) \simeq U(\mathfrak{g}_-) \otimes U(\mathfrak{g}_+) \otimes U(\mathfrak{h})$.

The Weyl group of the Lie algebra $\mathfrak{g}$. For any root $\gamma \in R$, consider the linear transformation σ_γ in the space $\mathfrak{h}^*$ defined by the formula

$$\sigma_\gamma(\chi) = \chi - \frac{2\langle \gamma, \chi \rangle}{\langle \gamma, \gamma \rangle}\gamma \quad \text{for any } \chi \in \mathfrak{h}^*. \tag{2.7}$$

The transformation σ_γ is the reflection in the hyperplane defined by the equation $\langle \chi, \gamma \rangle = 0$.

The group of linear transformations of $\mathfrak{h}^*$ generated by σ_γ, where $\gamma \in R$, is said to be the *Weyl group* of $\mathfrak{g}$ and denoted by W. We will use the following properties of the Weyl group.

1) W is a finite group.
2) W is generated by σ_α, where $\alpha \in B$.
3) W preserves R, $\mathfrak{h}^*_{\mathbb{Z}}$, $\mathfrak{h}^*_{\mathbb{R}}$ and $\langle \cdot, \cdot \rangle$.
4) $\det \sigma_\gamma = -1$ and $\sigma_\gamma(\gamma) = -\gamma$ for any $\gamma \in R$.
5) If $\alpha \in B$, $\gamma \in R_+$ and $\gamma \neq \alpha$, then $\sigma_\alpha(\gamma) \in R_+$
6) Set

$$C = \{\chi \in \mathfrak{h}^*_{\mathbb{R}} \mid \langle \alpha, \chi \rangle > 0 \ \text{ for any } \alpha \in B\} \tag{2.8}$$

and denote by $\overline{C}$ the closure of C in $\mathfrak{h}^*_{\mathbb{R}}$. Then $\overline{C}$ is a *fundamental domain* for the W-action on $\mathfrak{h}^*_{\mathbb{R}}$. More precisely,

a) If $\chi \in \mathfrak{h}^*_{\mathbb{R}}$, then $w\chi \in \overline{C}$ for a certain $w \in W$.

b) If $\chi, w\chi \in \overline{C}$, then $\chi = w\chi$. If, moreover, $\chi \in C$, then $w = e$.

If $\chi_1, \chi_2 \in \mathfrak{h}^*$, then we write $\chi_1 \sim \chi_2$ whenever χ_1 and χ_2 belong to the same orbit of the Weyl group, i.e., when $\chi_1 = w\chi_2$ for a certain $w \in W$.

2.1. The Verma modules, alias modules M^χ

The aim of these lectures is the description of finite dimensional $\mathfrak{g}$-modules over simple finite dimensional Lie algebras $\mathfrak{g}$. In the 1960's it was noted that it is more natural to describe the finite dimensional modules in the framework of a wider class of $\mathfrak{g}$-modules. First, let us give several preparatory definitions.

Let V be a $\mathfrak{g}$-module. For any $\chi \in \mathfrak{h}^*$, denote by $V(\chi)$ the space of vectors $v \in V$ such that $Hv = \chi(H)v$ for any $H \in \mathfrak{h}$ and call it the *weight space* of weight χ. If $V(\chi) \neq 0$, then χ is called a *weight* of the $\mathfrak{g}$-module V.

For a subalgebra $\mathfrak{h} \subset \mathfrak{g}$, a $\mathfrak{g}$-module V is said to be $\mathfrak{h}$-*diagonalizable* if $V = \bigoplus\limits_{\chi \in \mathfrak{h}^*} V(\chi)$.

2.1.1. Lemma. *For any simple finite dimensional Lie algebra* $\mathfrak{g}$, *let* V *be a* $\mathfrak{g}$*-module. Then*

1) $E_\gamma V(\chi) \subset V(\chi+\gamma)$ *for any* $\gamma \in R$, $\chi \in \mathfrak{h}^*$.
2) $\sum\limits_{\chi\in\mathfrak{h}^*} V(\chi) = \mathop{\oplus}\limits_{\chi\in\mathfrak{h}^*} V(\chi)$, *i.e., the sum is the direct one.*
3) $\mathop{\oplus}\limits_{\chi\in\mathfrak{h}^*} V(\chi)$ *is a* $\mathfrak{g}$*-submodule of* V.

Proof. 1) Let $v \in V(\chi)$. Then

$$HE_\gamma v = E_\gamma Hv + [H, E_\gamma]v = \chi(H)E_\gamma v + \gamma(H)E_\gamma v \text{ for any } H \in \mathfrak{h}.$$

Hence, $E_\gamma v \in V(\chi+\gamma)$.

2) Let $\chi_1, \dots, \chi_k$ be distinct elements of $\mathfrak{h}^*$ and let $v_i \in V(\chi_i)$ be a nonzero vector for each i. We must show that $v_1 + \ldots + v_k = 0$ is impossible for $k > 0$. Assume that on the contrary such an equality is possible and k is the minimal number when it holds. There is an element $H \in \mathfrak{h}$ such that $\chi_1(H) = 0$ and $\chi_2(H) \neq 0$. Then

$$0 = H(v_1 + \ldots + v_k) = \chi_2(H)v_2 + \ldots + \chi_k(H)v_k$$

yielding the contradiction with the fact that k is the least number with the said property.

3) Follows immediately from 1), 2) and the decomposition

$$\mathfrak{g} = \mathfrak{h} \oplus \Big(\mathop{\oplus}\limits_{\gamma\in R} \mathbb{C}E_\gamma\Big). \qquad \square$$

Let $\mathfrak{a}$ be a Lie subalgebra of $\mathfrak{g}$ and V a $\mathfrak{g}$-module. An element $v \in V$ is called $\mathfrak{a}$-*finite* if $\dim U(\mathfrak{a})v < \infty$. The $\mathfrak{g}$-module V is called $\mathfrak{a}$-*finite* if all the elements of V are $\mathfrak{a}$-finite.

2.1.2. Lemma. *Let* $\mathfrak{a}$ *be a subalgebra of* $\mathfrak{g}$ *and* V *a* $\mathfrak{g}$*-module. Set*

$$V^{\mathfrak{a}\text{-f}} := \{v \in V \mid v \ \ \textit{is } \mathfrak{a}\textit{-finite}\}.$$

The space $V^{\mathfrak{a}\text{-f}}$ *is a* $U(\mathfrak{g})$*-submodule of* V.

Proof. Let $v_1, v_2 \in V^{\mathfrak{a}\text{-f}}$. Then $U(\mathfrak{a})(v_1+v_2) \subset U(\mathfrak{a})v_1 + U(\mathfrak{a})v_2$, and hence finite dimensional. Therefore, $v_1 + v_2 \in V^{\mathfrak{a}\text{-f}}$. Let $X \in U(\mathfrak{g})$ be such that $X \in U(\mathfrak{g})_n$ and $v \in V^{\mathfrak{a}\text{-f}}$. It suffices to show that $\dim U(\mathfrak{a})U(\mathfrak{g})_n v < \infty$. Clearly, $\dim U(\mathfrak{g})_n U(\mathfrak{a})v < \infty$, and hence the next lemma implies Lemma 2.1.2. $\square$

2.1.3. Lemma. $U(\mathfrak{g})_n U(\mathfrak{a}) \subset U(\mathfrak{a})U(\mathfrak{g})_n$.

Proof. It suffices to show that $AYZ \in U(\mathfrak{g})_n U(\mathfrak{a})$ for any $A \in \mathfrak{a}$, $Z \in U(\mathfrak{a})$ and $Y \in U(\mathfrak{g})_n$. Indeed, $AYZ = (YAZ + [A, Y]Z) \in U(\mathfrak{g})_n U(\mathfrak{a})$. $\square$

The category $\mathcal{O}$. Let us now introduce a class of $\mathfrak{g}$-modules that we will consider. The objects of the *category* $\mathcal{O}$ are the $\mathfrak{g}$-modules M satisfying the following conditions:

1) M is a finitely generated $U(\mathfrak{g})$-module;
2) M is $\mathfrak{h}$-diagonalizable;
3) M is $\mathfrak{g}_+$-finite.

Clearly, if a $\mathfrak{g}$-module M belongs to $\mathcal{O}$, then so does any submodule of M, and any quotient module of M. Clearly, if $M_1, M_2 \in \mathcal{O}$, then $M_1 \oplus M_2 \in \mathcal{O}$. Most important for us objects in the category $\mathcal{O}$ are the following modules M^χ called *Verma modules.*

Let $\chi \in \mathfrak{h}^*$. In $U(\mathfrak{g})$, consider the left ideal I^χ generated by the elements E_γ, where $\gamma \in R_+$, and by $H + (\rho - \chi)(H)$, where $H \in \mathfrak{h}$. Define the $\mathfrak{g}$-module M^χ by setting

$$M^\chi = U(\mathfrak{g})/I^\chi.$$

Denote by m_χ the natural generator of M^χ (over $\mathfrak{g}$), namely, the image of $1 \in U(\mathfrak{g})$ under the map $U(\mathfrak{g}) \longrightarrow M^\chi$.

2.1.4. Lemma. *Let $\chi \in \mathfrak{h}^*$. Then*

1) *$E_\gamma(m_\chi) = 0$ for any $\gamma \in R_+$, and m_χ is weight vector of weight $\chi - \rho$.*

2) *M^χ is a free $U(\mathfrak{g}_-)$-module with one generator m_χ.*

3) *If M is an arbitrary $\mathfrak{g}$-module and $m \in M$ is a vector of weight $\chi - \rho$ such that $E_\gamma(m) = 0$ for any $\gamma \in R_+$, then there exists the unique $\mathfrak{g}$-module homomorphism $i\colon M^\chi \longrightarrow M$ such that $i(m_\chi) = m$. If, in addition, $Xm \neq 0$ for any non-zero $X \in U(\mathfrak{g}_-)$, then i is an embedding.*

Proof. 1) follows from the definition of M^χ.

2) follows from the decomposition $U(\mathfrak{g}) = U(\mathfrak{g}_-) \otimes U(\mathfrak{h}) \otimes U(\mathfrak{g}_+)$.

3) follows immediately from 1) and 2). □

Observe that since $\mathfrak{g}_+$ is generated by E_α for $\alpha \in B$, the condition $E_\gamma m = 0$ for any $\gamma \in R$ is equivalent to the condition $E_\alpha m = 0$ for any $\alpha \in B$. The weight vector m satisfying such a condition will be called the *highest weight vector.*

2.1.5. Lemma. *Let $M \in \mathcal{O}$. Then M contains a highest weight vector.*

Proof. Let m be a non-zero weight vector in M. Consider various sequences $S = (\alpha_1, \dots, \alpha_k)$, where $\alpha_i \in B$ and α_i may enter S with multiplicities, such that $m_S = E_{\alpha_1} \dots E_{\alpha_k} m$ are nonzero. If S and S' contain different number of elements, then the weights of m_S and $m_{S'}$ are different. Since the space $U(\mathfrak{g}_+)m$ is finite dimensional, it intersects with only a finite number of weight subspaces, and therefore there is a sequence S of maximal length. Then m_S is the highest weight vector to be found. □

2.1.6. Characters. The set $\mathcal{E}$. In the study of modules from the category $\mathcal{O}$, the notion of the character of a $\mathfrak{g}$-module M is often useful. For $\mathfrak{h}$-diagonalizable $\mathfrak{g}$-modules M such that $\dim M^\chi < \infty$ for any $\chi \in \mathfrak{h}^*$, the *character* of the module M is the function π_M on $\mathfrak{h}^*$ defined by the equation

$$\pi_M(\chi) = \dim M^\chi.$$

On $\mathfrak{h}^*$, define the *Kostant function* K (sometimes simply called *partition function*) from the equation

$$K(\chi) = \text{the number of representations of the weight } \chi \text{ in the form } \chi = -\sum_{\gamma \in R_+} n_\gamma \gamma, \text{ where } n_\gamma \in \mathbb{Z}_+.$$

For any function u on $\mathfrak{h}^*$, set

$$\operatorname{supp}\ u = \{\chi \in \mathfrak{h}^* \mid u(\chi) \neq 0\}.$$

Denote:

$$\mathcal{E} = \{\text{functions } u \text{ on } \mathfrak{h}^* \mid \operatorname{supp}\ u \text{ is contained in the union of a finite number of sets of the form } \nu - Q_+, \text{ where } \nu \in \mathfrak{h}^*\}.$$

For example, supp $K = -Q_+$, hence, $K \in \mathcal{E}$.

2.1.7. Lemma. 1) $\pi_{M^\chi}(\psi) = K(\psi - \chi + \rho)$.

2) *If $M \in \mathcal{O}$, then π_M is defined and $\pi_M \in \mathcal{E}$.*

Proof. 1) Let us enumerate the elements of R_+, e.g., $\gamma_1, \dots, \gamma_s$. The elements $E_{-\gamma_1}^{n_1} \dots E_{-\gamma_s}^{n_s} m_\chi$, where $n_1, \dots, n_s \in \mathbb{Z}_+$, form, clearly, a basis in M^χ. Hence, for $\psi = -\sum n_i \gamma_i$, we have $\pi_{M^\chi}(\psi) = K(\psi - \chi + \rho)$.

2) Let $m_1, \dots, m_n$ be generators of the $\mathfrak{g}$-module M. Clearly, we may assume that the m_i are weight vectors. Since $U(\mathfrak{g}) = U(\mathfrak{g}_-)U(\mathfrak{g}_+)U(\mathfrak{h})$, we see that

$$\begin{aligned}&U(\mathfrak{g})(m_1, \dots, m_n) = U(\mathfrak{g}_-)U(\mathfrak{g}_+)U(\mathfrak{h})(m_1, \dots, m_n) =\\ &U(\mathfrak{g}_-)(U(\mathfrak{g}_+)(m_1, \dots, m_n)).\end{aligned}$$

Let $g_1, \dots, g_k$ be a basis of the finite dimensional space $U(\mathfrak{g}_+)(m_1, \dots, m_n)$ consisting of weight vectors and $\chi_1, \dots, \chi_k$ the weights of the vectors $g_1, \dots, g_k$. As in heading 1), we have

$$\dim M^\psi \leq \sum_{1 \leq i \leq k} K(\psi - \chi_i),$$

implying the lemma. $\square$

It is easy to prove the converse statement:

$$\text{If } M \text{ is a finitely generated } U(\mathfrak{g})\text{-module such that its character } \pi_M \text{ is defined and } \pi_M \in \mathcal{E}, \text{ then } M \in \mathcal{O}. \tag{2.9}$$

2.2. The representations of $\mathfrak{sl}(2)$

In this section we will describe representations of the simplest of the simple Lie algebras, i.e., of the Lie algebra $\mathfrak{sl}(2)$.

The Lie algebra $\mathfrak{sl}(2)$ consists of complex matrices $X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ such that $\operatorname{tr} X = a + d = 0$. In $\mathfrak{sl}(2)$, we select the following basis

$$E_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad E_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

The commutation relations between the elements of the basis are:

$$[H, E_+] = 2E_+, \quad [E_+, E_-] = H, \quad [H, E_-] = -2E_-.$$

Remark. Let $\mathfrak{g}$ be any (finite dimensional) simple Lie algebra and $\gamma \in R$. Then the elements E_γ, H_γ and $E_{-\gamma}$ satisfy similar relations

$$[H_\gamma, E_{-\gamma}] = -2E_{-\gamma}, \quad [H_\gamma, E_\gamma] = 2E_\gamma, \quad [E_\gamma, E_{-\gamma}] = H_\gamma.$$

Hence, the study of representations of the Lie algebra $\mathfrak{sl}(2)$ provides us with lots of information on the representations of any (semi-)simple Lie algebra $\mathfrak{g}$.

The above relations between E_-, H and E_+ and the induction yield the following relations in $U(\mathfrak{sl}(2))$:

$$[H, E_-^k] = -2kE_-^k, \quad [H, E_+^k] = 2kE_+^k, \quad [E_+, E_-^k] = kE_-^{k-1}(H - (k-1)).$$

Besides, it is easy to verify that

$$\triangle = 2\Big(E_-E_+ + E_+E_- + \frac{1}{2}H^2\Big) = 4E_-E_+ + H^2 - 2H$$

belongs to the center of $U(\mathfrak{sl}(2))$. This element is called the (second) *Casimir element*. As we will see, it generates the center of $U(\mathfrak{sl}(2))$.

2.2.1. Lemma. *Set $\mathfrak{h} = \mathbb{C}H$ (this is the Cartan subalgebra of $\mathfrak{sl}(2)$). Let V be an irreducible $\mathfrak{h}$-module, i.e., $\dim V = 1$, and $Hv = \chi v$ for some $\chi \in \mathbb{C}$ and a nonzero $v \in V$.*

Set $E_+v = 0$ and let $M^\chi = \mathbb{C}[E_-]V$. The two cases possible:

1) $\chi \notin \mathbb{Z}_+$. Then M^χ is irreducible.

2) $\chi \in \mathbb{Z}_+$. Then M^χ contains $M^{-\chi-2}$ and the quotient $L^\chi := M^\chi/M^{-\chi-2}$ is irreducible.

Proof. Set $v_k = E_-^k v$, where $k \in \mathbb{Z}_+$. Then

$$\begin{aligned} &E_+v_k = E_+E_-^k v = E_-^k E_+ v + [E_+, E_-^k]v = \\ &k(\chi - k + 1)E_-^{k-1}v = k(\chi - k + 1)v_{k-1}. \quad \square \end{aligned}$$

Let us now show that all finite dimensional irreducible representations of $\mathfrak{sl}(2)$ are of the form $L^l = M^l/M^{-l-2}$ for some $l \in \mathbb{Z}_+$. From Lemma it is clear that $\dim L^l = l+1$ and the elements $\{a_{-l}, a_{-l+2}, \dots, a_{l-2}, a_l\}$, where $a_{l-2k} = E_-^k v$, span the space L^l. The $\mathfrak{sl}(2)$-action on L^l is given by the formulas

$$\begin{aligned}
&Ha_{l-2k} = (l-2k)a_{l-2k} \text{ for } k = 0, 1, \dots, l\\
&E_-a_k = a_{k-2} \text{ for } k > -l \text{ and } E_-a_{-l} = 0,\\
&E_+a_k = k(l+1-k)a_{k+2} \text{ for } k < l \text{ and } E_+a_l = 0.
\end{aligned}$$

It is easy to verify that L^l is indeed an $\mathfrak{sl}(2)$-module. (What should one verify?)

2.2.2. Proposition. 1) *In any finite dimensional non-zero $\mathfrak{sl}(2)$-module V, there is a submodule isomorphic to one of L^l.*

2) *If $a \in L^l$, then $\triangle a = l(l-2)a$.*

3) *The modules L^l are irreducible, distinct, and exhaust all finite dimensional irreducible $\mathfrak{sl}(2)$-modules.*

Proof. 1) Let $v \in V$ be an eigenvector of H (over $\mathbb{C}$, it exists). Let $i \in \mathbb{Z}_+$ be the maximal number such that $E_+^i v \neq 0$ and set $v_0 = E_+^i v$. Such a maximal number exists because $E_+^i v$ are eigenvectors of H (verify!) with distinct eigenvalues, hence, linearly independent.

Then $Hv_0 = \chi v_0$, $E_+v_0 = 0$ and, by the same argument applied to linearly independent eigenvectors, $E_-^k v_0 = 0$ for a sufficiently large k. By Lemma 2.1, $\chi \in \mathbb{Z}_+$ and the space spanned by $E_-^r v_0$, where $r = 0, 1, \dots, \chi$, forms a submodule in V isomorphic to L^χ.

2) It is quite straightforward that $\triangle a_0 = l(l-2)a_0$. If $a \in L^l$ is a weight vector, then $a = Xa_0$ for a certain $X \in U(\mathfrak{sl}(2))$. Hence, $\triangle a = \triangle X a_0 = X \triangle a_0 = l(l-2)a$.

3) If L^l contains a non-trivial submodule V, then it contains L^k for some $k < l$; this is a contradiction to the fact that $\triangle = l(l-2)$ on L^l and $\triangle = k(k-2)$ on L^k.

Heading 1) implies that L^l, where $l \in \mathbb{Z}_+$, exhaust all irreducible $\mathfrak{sl}(2)$-modules. □

2.2.3. Proposition. *Any finite dimensional $\mathfrak{sl}(2)$-module V is isomorphic to the direct sum of modules of type L^l for different l's. In other words, finite dimensional representations of $\mathfrak{sl}(2)$ are completely reducible.*

Proof. Denote:

$$V(s_i) = \operatorname{Ker}(\triangle - s_i)^N \text{ for } N \text{ so large that this space does not grow with } N \text{ any more.}$$

As is well known from Linear Algebra (see, e.g., [P]), V is the direct sum of subspaces $V(s_i)$. Since the operators E_-, H and E_+ commute with $\triangle$, the spaces $V(s_i)$ are $\mathfrak{sl}(2)$-submodules of V. Hence, it suffices to consider the case $V = V(s)$. Let

$$\{0\} = V_0 \subset V_1 \subset \cdots \subset V_k = V \tag{2.10}$$

be submodules such that V_i/V_{i-1} is irreducible for every i (since $\dim V < \infty$, such submodules exist). The operator $\triangle$ preserves each V_i, and therefore the action of $\triangle$ on V_i/V_{i-1} is defined. Since the quotients V_i/V_{i-1} are irreducible, Proposition 2.2.2 implies that $V_i/V_{i-1} \simeq L^{l_i}$ for some $l_i \in \mathbb{Z}_+$.

Since the operator $(\triangle - s)^N$ vanishes for large N, it vanishes on all the V_i/V_{i-1}. Hence, $s = l_i(l_i - 2)$. This means that all modules V_i/V_{i-1} are isomorphic to one irreducible module L^l for a certain l.

Now, let us prove that V is isomorphic to the direct sum of several copies of L^l. The proof will be carried out by induction on the number k of submodules in the chain (2.10). Suppose that V_{k-1} is isomorphic to the direct sum of $k-1$ copies of L^l. The eigenvalues of H on L^l (hence, on V_{k-1}) are equal to $-l, -l+2, \dots, l-2, l$. Hence, H has the same eigenvalues in V_k.

Let $V(i)$ be the subspace of V corresponding to eigenvalue i, where $i = -l, -l+2, \dots, l-2, l$, of H (i.e., $V(i) = \operatorname{Ker}(H-i)^N$ for large N). Clearly,

$$E_-V(i) \subset V(i-2), \qquad E_+V(i) \subset V(i+2).$$

There exists a vector $v \in V(l)$ such that $v \notin V_{k-1}$. Indeed, the converse would mean that $V(l) \subset V_{k-1}$, and hence the operator $H - l$ in the space $V_k/V_{k-1} \simeq L^l$ would have been invertible.

Clearly, $(H-l)v \in V_{k-1} \cap V(l)$. Let us prove that $(H-l)v = 0$. Indeed,

$$E_+v \in V(l+2) = \{0\}, \qquad E_-^{l+1}v \in V(-l-2) = \{0\}.$$

Therefore,

$$0 = E_+E_-^{l+1}v - E_-^{l+1}E_+v = (l+1)E_-^l(H-l)v.$$

Since $V_{k-1} = \bigoplus\limits_{1 \leq i \leq k-1} (L^l)^i$, the operator E_-^l is without kernel on $V_{k-1} \cap V(l)$. Therefore, $(H-l)v = 0$. Hence, v is a highest weight vector in $V = V_k$ and the submodule of V generated by v is isomorphic to L^l. Therefore, $V_k = V_{k-1} \oplus L^l$. Proposition is proved. □

Corollary. 1) *If V is a finite dimensional $\mathfrak{sl}(2)$-module, then it is H-diagonalizable and the operators E_-^i and E_+^i perform an isomorphism of $V(i) \simeq V(-i)$.*

2) *If $\mathfrak{g}$ is any simple Lie algebra, and V any finite dimensional $\mathfrak{g}$-module, then $V \in \mathcal{O}$.*

Proof. It suffices to verify that V is $\mathfrak{h}$-diagonalizable. Since the operators H_α, where $\alpha \in B$, generate $\mathfrak{h}$ and commute, it suffices to verify that V is H_α-diagonalizable.

This, in turn, follows from Proposition 2.2.3, since V is the finite dimensional $\mathfrak{sl}(2)$-module for $\mathfrak{sl}(2)$ generated by $E_{-\alpha}$, H_α and E_α. □

2.3. The modules L^χ

In this section, we provide with a supply of finite dimensional $\mathfrak{g}$-modules. In what follows, we will show that these modules exhaust all irreducible finite dimensional $\mathfrak{g}$-modules.

2.3.1. Lemma. *Suppose $\chi \in \mathfrak{h}^*$ and $\alpha \in B$ are such that $\chi \geq \sigma_\alpha(\chi)$. Then there exists an embedding $M^{\sigma_\alpha(\chi)} \longrightarrow M^\chi$ that transforms $m_{\sigma_\alpha(\chi)}$ into $m' = E_{-\alpha}^k m_\chi$, where $k = \dfrac{2\langle\chi,\alpha\rangle}{\langle\alpha,\alpha\rangle}$.*

Proof. Clearly, the weight of m' is equal to $\chi - \dfrac{2\langle\chi,\alpha\rangle}{\langle\alpha,\alpha\rangle}\alpha = \sigma_\alpha(\chi)$. By Lemma 2.1.4it suffices to show that $E_\beta m' = 0$ for $\beta \in B$. If $\beta \neq \alpha$, then

$$E_\beta m' = E_\beta E_{-\alpha}^k m_\chi = E_{-\alpha}^k E_\beta m_\chi = 0,$$

because $[E_\beta, E_{-\alpha}] = 0$. Further,

$$E_\alpha m' = E_\alpha E_{-\alpha}^k m_\chi = E_{-\alpha}^k E_\alpha m_\chi + kE_{-\alpha}^{k-1}(H_\alpha - (k-1))m_\chi = 0 + 0 = 0,$$

since $H_\alpha m_\chi = (\chi - \rho)(H_\alpha)m_\chi = (k-1)m_\chi$. □

2.3.2. Lemma. *Suppose $\chi \in \mathfrak{h}^*$ and $\alpha \in B$ is such that $\sigma_\alpha(\chi) \leq \chi$. Let M be a submodule in M^χ containing $M^{\sigma_\alpha(\chi)}$. Let $M' = M^\chi/M$. Then $\pi_{M'}(\psi) = \pi_{M'}(\sigma_\alpha(\psi))$ for any $\psi \in \mathfrak{h}^*$.*

Proof. Let $\mathfrak{a}_\alpha$ be a Lie subalgebra of $\mathfrak{g}$ generated by $E_{-\alpha}$, H_α, E_α (a copy of $\mathfrak{sl}(2)$). By Lemma 2.3.1 the image of m_χ in M' is $\mathfrak{a}_\alpha$-finite. Lemma 2.1.2 implies that M' is $\mathfrak{a}_\alpha$-finite. Let $\psi \in \mathfrak{h}^*$. Let us consider a finite dimensional $\mathfrak{a}_\alpha$-submodule V in M' spanned by $M'(\psi)$ and $M'(\sigma_\alpha(\psi))$. (Recall that $M'(\psi)$ is the subspace of M' consisting of vectors of weight ψ.)

We may assume that $\dfrac{2\langle\alpha,\psi\rangle}{\langle\alpha,\alpha\rangle} = k$ is an integer, otherwise

$$M'(\psi) = M'(\sigma_\alpha(\psi)) = 0.$$

Permuting, if necessary, ψ and $\sigma_\alpha(\psi)$, we may assume that $k \geq 0$. Let

$$V(i) = \{v \in V \mid H_\alpha(v) = iv\}.$$

Corollary 2.2.3 implies that $E_{-\alpha}^k\colon V(k) \longrightarrow V(-k)$ is an isomorphism. This isomorphism transforms $M'(\psi)$ into $M'(\sigma_\alpha(\psi))$, and therefore

$$\dim M'(\psi) \leq \dim M'(\sigma_\alpha(\psi)).$$

Similarly, considering the isomorphism $E_\alpha^k\colon V(-k) \longrightarrow V(k)$ we see that

$$\dim\ M'(\sigma_\alpha(\psi)) \leq \dim\ M'(\psi). \quad \square$$

Hence, $\pi_{M'}(\psi) = \pi_{M'}(\sigma_\alpha(\psi)$.
Set

$$\overline{\mathcal{D}} = \mathfrak{h}^*_{\mathbb{Z}} \cap \overline{C}. \tag{2.11}$$

Recall that $\sigma_\alpha(\varphi+\rho) < \varphi+\rho$ and, for any $\varphi \in \overline{\mathcal{D}}$, define the $\mathfrak{g}$-module

$$L^\varphi := M^{\varphi+\rho} / \bigcup_{\alpha\in B} M^{\sigma_\alpha(\varphi+\rho)}.$$

2.3.3. Theorem. 1) $\pi_{L^\varphi}(\varphi) = 1$.
2) $\pi_{L^\varphi}(w\psi) = \pi_{L^\varphi}(\psi)$ *for any* $w \in W$ *and* $\psi \in \mathfrak{h}^*$.
3) *If* ψ *is a weight of* L^φ, *then either* $|\psi| < |\varphi|$, *where* $|\psi|^2 = \langle\psi,\psi\rangle$, *or* $\psi \sim \varphi$.
4) $\dim L^\varphi < \infty$.

Proof. 1) The modules $M^{\sigma_\alpha(\varphi+\rho)}$ do not contain vectors of weight φ, and hence these modules belong to $\sum\limits_{\psi\in\mathfrak{h}^*\setminus\{\varphi\}} M^{\varphi+\rho}(\psi)$. Therefore,

$$\dim L^\varphi(\varphi) = \dim M^{\varphi+\rho}(\varphi) = 1.$$

2) If $w = \sigma_\alpha$, where $\alpha \in B$, then heading 2) of Theorem follows from Lemma 2.3.2. Since W is generated by σ_α, where $\alpha \in B$, heading 2) holds for any $w \in W$.

3) It follows from Lemma 2.1.6 that

$$\text{supp }\ \pi_{L^\varphi} = \text{supp }\ \pi_{M^{\varphi+\rho}} = \varphi - Q_+.$$

Let $\pi_{L^\varphi}(\psi) \neq 0$. By replacing ψ by a W-equivalent element ψ' that belongs to $\overline{C}$, we see that $\pi_{L^\varphi}(\psi') = \pi_{L^\varphi}(\psi) \neq 0$. Hence, $\varphi = \psi' + \lambda$, where $\lambda \in Q_+$. Further on

$$|\varphi|^2 = |\psi|^2 + |\lambda|^2 + 2\langle\psi',\ \lambda\rangle \geq |\psi'|^2 + |\lambda|^2.$$

Hence, either $|\varphi| > |\psi'| = |\psi|$ or $|\lambda| = 0$ and then $\varphi = \psi' \sim \psi$.

4) Since supp π_{L^φ} is contained in the intersection of the lattice $\varphi - Q_+$ with the ball $\{\psi \mid |\psi| \leq |\varphi|\}$, it follows that supp π_{L^φ} is finite. Hence $\dim L^\varphi < \infty$. $\square$

2.4. The Harish–Chandra theorem

In this section we formulate the Harish–Chandra theorem that describes the center of $U(\mathfrak{g})$. Proof of this theorem is carried out in § 2.8.

Denote by $Z(U)$ the center of $U(\mathfrak{g})$. Let $S(\mathfrak{h})$ be the ring of polynomial functions on $\mathfrak{h}^*$.

2.4.1. Lemma. *Let* $\chi \in \mathfrak{h}^*$ *and* $z \in Z(U)$. *Then*
1) *The element* z *acts on* M^χ *is the operator of multiplication by a constant* $\theta_\chi(z)$, *i.e.,*

$$zm = \theta_\chi(z)m \quad \text{for any } m \in M^\chi.$$

2) *The function $\theta_\chi(z)$ is polynomial in χ.* (Denote by j_z the corresponding element in $S(\mathfrak{h})$.)

Proof. Let us arbitrarily enumerate the roots $\gamma \in R_+$; in $\mathfrak{g}$, consider the basis

$$\{E_{-\gamma_1}, \dots, E_{-\gamma_s}, H_1, \dots, H_r, E_{\gamma_s}, \dots, E_{\gamma_1}\}$$

where $\{H_1, \dots, H_r\}$ is a basis in $\mathfrak{h}$. For any set

$$N = \{n_1, \dots, n_s, m_1, \dots, m_r, n'_s, \dots, n'_1 \mid n_i, n'_i, m_j \in \mathbb{Z}_+\},$$

consider the element

$$X_N = E_{-\gamma_1}^{n_1} \dots E_{-\gamma_s}^{n_s} H_1^{m_1} \dots H_r^{m_r} E_{\gamma_s}^{n'_s} \dots E_{\gamma_1}^{n'_1}.$$

By the PBW theorem the elements X_N form a basis in $U(\mathfrak{g})$. If $H \in \mathfrak{h}$, then

$$[H, X_N] = \Big(\sum (n'_i - n_i)\gamma_i(H)\Big) \cdot X_N,$$

where $\sum (n'_i - n_i)\gamma_i$ is the weight of X_N. We have $[H, z] = 0$ for any $z \in Z(U)$ and $H \in \mathfrak{h}$. Hence, $z \equiv \mu(z) \mod (U(\mathfrak{g})\mathfrak{g}_+)$, where $\mu(z)$ is a polynomial in the H_i.

Let m_χ be a generator of the module M^χ. Then $z(m_\chi) = (\mu(z)(\chi - \rho))m_\chi$, where $\mu(z)$ is considered as the polynomial function on $\mathfrak{h}^*$. Since m_χ generates the module M^χ, it follows that

$$z(m) = \mu(z)(\chi - \rho)m \quad \text{for any } m \in M^\chi.$$

Setting

$$j_z(\chi) = \mu(z)(\chi - \rho),$$

we are done. □

2.4.2. Lemma. *We have $j_z(\chi) = j_z(w\chi)$ for any $z \in Z(U)$ and $w \in W$.*

Proof. It suffices to consider the case where $w = \sigma_\alpha$ for $\alpha \in B$. Since $j_z(\chi)$ and $j_z(\sigma_\alpha(\chi))$ are polynomial functions in χ, it suffices to prove the equality for $\chi \in \overline{\mathfrak{D}}$, see eq. (2.11). But, in this case, $M^{\sigma_\alpha(\chi)} \subset M^\chi$, and hence

$$zm = j_z(\chi)m = j_z(\sigma_\alpha(\chi))m,$$

i.e., $j_z(\chi) = j_z(\sigma_\alpha(\chi))$ for any $m \in M^{\sigma_\alpha(\chi)}$. □

2.4.3. The Harish–Chandra theorem. Define the W-action in $S(\mathfrak{h})$ by the formula

$$wP(\chi) = P(w^{-1}\chi) \quad \text{for any } w \in W,\ P \in S(\mathfrak{h}),\ \chi \in \mathfrak{h}^*.$$

Let $S(\mathfrak{h})^W$ be the subring of W-invariant functions.

Theorem. *The correspondence $z \mapsto j_z$ defines a ring homomorphism $j\colon Z(U) \longrightarrow S(\mathfrak{h})^W$. The map j is an isomorphism.*

2.5. Corollaries of the Harish–Chandra theorem. Central characters

For any $\chi \in \mathfrak{h}^*$, we have the homomorphism $\theta_\chi\colon Z(U) \longrightarrow \mathbb{C}$. These homomorphisms, and sometimes their kernels, are called the *central characters.* The following lemma describes the relation between such homomorphisms for different χ's.

2.5.1. Lemma. $\theta_{\chi_1} = \theta_{\chi_2}$ *if and only if* $\chi_1 \sim \chi_2$.

Proof. If $\chi_1 \sim \chi_2$, then by Lemma 2.4.2 $\theta_{\chi_1} = \theta_{\chi_2}$.

Let $\chi_1 \not\sim \chi_2$. Let us construct a polynomial $T \in S(\mathfrak{h})^W$ such that $T(\chi_1) = 0$, while $T(\chi_2) \neq 0$. For this, take a polynomial $T' \in S(\mathfrak{h})$ such that $T'(\chi_1) = 0$ and $T'(w\chi_2) = 1$ for any $w \in W$ and set

$$T(\chi) = \sum_{w\in W} T'(w\chi).$$

As follows from the Harish–Chandra theorem, there is an element $z \in Z(U)$ such that $j_z = T$. But then

$$j_z(\chi_1) = \theta_{\chi_1} \neq \theta_{\chi_2}(z) = j_z(\chi_2). \qquad \square$$

Denote by Θ the set of all homomorphisms $\theta\colon Z(U) \longrightarrow \mathbb{C}$. Let M be a $\mathfrak{g}$-module. Let M contain an eigenvector m of all operators $z \in Z(U)$. Let $\Theta(M)$ denote the set of all homomorphisms

$$\theta\colon Z(U) \longrightarrow \mathbb{C}, \quad zm = \theta(z)m.$$

Remark. One can show that any homomorphism $\theta\colon Z(U) \longrightarrow \mathbb{C}$ is of the form θ_χ for a certain $\chi \in \mathfrak{h}^*$.

2.5.2. Theorem. *Let* $M \in \mathcal{O}$. *Then*

1) *The set* $\Theta(M)$ *is finite.*

2) *If* $\theta \in \Theta\backslash\Theta(M)$, *then there exists an element* $z \in \operatorname{Ker}\ \theta \subset Z(U)$ *such that* $zm = m$ *for any* $m \in M$.

3) *Let* $\Theta(M) = \{\theta_1, \dots, \theta_k\}$. *Set* $I_{\theta_i} = \operatorname{Ker}\ \theta_i$ *and*

$$M^{\theta_i,n} = \{m \in M \mid (I_{\theta_i})^n m = 0\}.$$

For a fixed θ_i, *the modules* $M^{\theta_i,n}$ *coincide for* n *sufficiently large.* The module to which all these coinciding modules are isomorphic will be denoted by M^{θ_i}.

4) $\Theta(M^{\theta_i}) = \{\theta_i\}$.

5) $M = M^{\theta_1} \oplus \dots \oplus M^{\theta_k}$.

Proof. The module M is generated by the finite dimensional space $V = \bigoplus\limits_{\chi\in\Xi} M(\Xi)$, where Ξ is a finite subset in $\mathfrak{h}^*$. Since the elements of $Z(U)$ commute with the elements of $\mathfrak{h}$, we have $zV \subset V$ for any $z \in Z(U)$. Set

$I = \text{Ker } (Z(U) \longrightarrow \text{End } V)$; then $\dim\ Z(U)/I < \infty$. Since V generates M and $IV = 0$, it follows that $IM = 0$, i.e., the finite dimensional commutative algebra $A = Z(U)/I$ acts on M.

Now, Theorem 2.5.2 is the direct corollary of the following standard proposition (see [W], § 98, Th. 12).

2.5.2.1. Proposition. *Let A be a finite dimensional over $\mathbb{C}$ commutative algebra with unit. Then*

1) *In A, there is a finite number of maximal ideals $\mathfrak{m}_i$, where $i = 1, \dots, k$.*
2) *There are elements $e_i \in A$, where $i = 1, \dots, k$, such that*

$$\begin{aligned} &e_i e_j = 0 \textit{ for } i \neq j \textit{ and } e_i^2 = e_i;\\ &e_1 + e_2 + \ldots + e_k = 1;\\ &e_i \notin \mathfrak{m}_j \textit{ for } i \neq j;\\ &e_i \mathfrak{m}_i^n = 0 \textit{ for } n > \dim A. \end{aligned}$$

Theorem 2.5.2 is proved. □

Note that the decomposition $M = \bigoplus\limits_{1 \leq i \leq k} M^{\theta_i}$ defined in Theorem 2.5.2 is *functorial*. This means that if $M, M' \in \mathcal{O}$ and $\tau \colon M \longrightarrow M'$ is a $\mathfrak{g}$-module homomorphism, then

$$\tau(M^\theta) \subset M'^\theta \quad \text{for any } \theta \in \Theta;$$

if $\theta \notin \Theta(M)$, then we set $M^\theta = 0$.

Moreover, if $M_1 \subset M_2 \subset M$, then $\Theta(M_2/M_1) \subset \Theta(M)$.

2.5.3. Lemma. *Let $M \in \mathcal{O}$. Then every element $\theta \in \Theta(M)$ is of the form $\theta = \theta_\xi$, where $\xi - \rho$ is a weight of M.*

Proof. We may assume that $M = M^\theta$, where $\theta \in \Theta$. Let m be a highest weight vector in M^θ of weight $\xi - \rho$. Then

$$zm = \theta_\xi(z)m \quad \text{for any } z \in Z(U),$$

i.e., $\theta_\xi \in \Theta(M^\theta) = \{\theta\}$. Hence, $\theta = \theta_\xi$. □

2.5.4. Jordan–Hölder series. Recall that the *Jordan–Hölder series* of the module M is the sequence of submodules

$$\{0\} = M_0 \subset M_1 \subset \ldots \subset M_k = M$$

such that quotient modules $L_i = M_i/M_{i-1}$ are irreducible.

Proposition. *Let $M \in \mathcal{O}$. Then the Jordan–Hölder series of M is finite.*

The Jordan–Hölder theorem [W], § 51, claims that the modules L_i are defined by M uniquely up to the order. The collection of the L_i will be called the *Jordan–Hölder decomposition* of M.

Proof. Set $\Theta = \{\xi \in \mathfrak{h}^* \mid \theta_{\xi+\rho} \in \Theta(M)\}$ and $V_M = \bigoplus\limits_{\xi\in\Theta} M^\xi$. Since the set Θ is finite, $\dim V_M < \infty$. By Lemma 2.5.3,$V_M \neq \{0\}$. Let M' be a submodule in M such that $V_{M'} = V_M \cap M' \neq V_M$ and the dimension of $V_{M'}$ is maximal. Let us prove that M' is a maximal proper submodule of M, i.e., let us prove that M/M' is irreducible.

Indeed, suppose $M' \subset M'' \subset M$. By the choice of M' we have $V_{M'} = V_{M''}$. But $\Theta(M''/M') \subset \Theta(M)$. Hence, Lemma 2.5.3 implies that there is a weight in M''/M' that belongs to Θ, contradicting the equality $V_{M'} = V_{M''}$. Thus, M/M' is irreducible. Let us construct a sequence of submodules $M \supset M_1 \supset \dots$ setting $M_1 = M'$ and $M_i = (M_{i-1})'$. Since $\dim V_{M_i'}$ decreases, it follows that $V_{M_{i_0}} = 0$ for a certain i_0. But then by Lemma 2.5.3, $M_{i_0} = 0$. The sequence

$$M \supset M_1 \supset \dots \supset M_{i_0} = \{0\}$$

is a finite Jordan–Hölder series. □

2.5.5. A description of the irreducible modules in the category $\mathcal{O}$.

Lemma. 1) *Let $\xi \in \mathfrak{h}^*$ and M be the union of all proper submodules of M^ξ. Then M is the proper submodule of M^ξ and the quotient module $L^\xi = M^\xi/M$ is irreducible.*

2) *Any irreducible module $L \in \mathcal{O}$ is of the form $L = L^\xi$ for a certain $\xi \in \mathfrak{h}^*$.*

3) *Let $\{L_i\}_{i=1}^k$ be the Jordan–Hölder decomposition of M^ξ. Then $L_i = L^{\xi_i}$, where $\xi_i \sim \xi$ and $\xi_i \leq \xi$. There is only one L^ξ among the L_i.*

Proof. 1) Any proper submodule M' of M^ξ is contained in $\bigoplus\limits_{\varphi\neq\xi-\rho} M^\xi(\varphi)$. Hence, M is a proper submodule of M^ξ. The module $L^\xi = M^\xi/M$ is irreducible, since M is maximal proper submodule of M^ξ.

2) Let $L \in \mathcal{O}$ be an irreducible module and $l \in L$ be the highest weight vector. Then there is a map $\tau\colon M^\xi \longrightarrow L$, where $\xi - \rho$ is the weight of l, such that $\tau(m_\xi) = l$. Clearly, τ induces the isomorphism of L^ξ and L.

3) Let $L_i = L^\xi$. Then $\theta_\psi = \theta_\xi$ and by Lemma 2.5.1 $\psi \sim \xi$. We have $M^{\psi-\rho}(\xi) \neq \{0\}$; hence, $\psi \leq \xi$. Since $\dim M^{\xi-\rho}(\xi) = 1$, the module L^ξ enters the Jordan–Hölder decomposition of M^ξ only once. □

Denote by $C(\mathcal{O})$ the free Abelian group generated by expressions $[M]$, where M runs the objects of $\mathcal{O}$ and by $C'(\mathcal{O})$ the subgroup of $C(\mathcal{O})$ generated by the expressions $[M_1] + [M_2] - [M]$ for all exact sequences

$$0 \longrightarrow M_1 \longrightarrow M \longrightarrow M_2 \longrightarrow 0.$$

The quotient group $K(\mathcal{O}) = C(\mathcal{O})/C'(\mathcal{O})$ is called the *Grothendieck group* of the category $\mathcal{O}$.

2.5.6. Lemma. *The map $\pi\colon [M] \longrightarrow \pi_M$ can be extended to an additive map $K(\mathcal{O}) \longrightarrow \mathcal{E}$.*

Proof follows immediately from the fact that $\pi_M = \pi_{M_1} + \pi_{M_2}$ for any exact sequence

$$0 \longrightarrow M_1 \longrightarrow M \longrightarrow M_2 \longrightarrow 0, \text{ where } M_1, M_2, M \in \mathcal{O}. \quad \square$$

2.5.7. Lemma. *The elements $[M^\psi]$, where $\psi \in \mathfrak{h}^*$, generate the group $K(\mathcal{O})$. More precisely, if $M \in \mathcal{O}$, then $[M]$ belongs to the subgroup generated by the finite set*

$$\{[M^\psi] \mid \psi \in \mathfrak{h}^* \textit{ is such that } \theta_\psi \in \Theta(M)\}.$$

Proof. Let

$$\{0\} = M_0 \subset M_1 \subset \ldots \subset M_k = M$$

be the Jordan–Hölder series of M and $L_i = M_i/M_{i-1}$. Clearly,

$$[M] = [L_1] + \ldots + [L_k] \quad \text{and} \quad \Theta(L_i) \subset \Theta(M).$$

Therefore, it suffices to verify the lemma for a given irreducible module $M = L^\xi$. Let us prove that

$$[L^\xi] = [M^\xi] + \sum_{\varphi \leq \xi,\ \varphi \sim \xi} c_\varphi [M^\varphi], \text{ where } c_\varphi \in \mathbb{Z}. \tag{2.12}$$

From Lemma 2.5.5 it follows that

$$[M^\psi] = [L^\psi] + \sum_{\varphi \leq \psi,\ \varphi \sim \psi} n_\varphi [L^\varphi],$$

where $n_\varphi \in \mathbb{Z}$ for any $\psi \in \mathfrak{h}^*$. Therefore, eq. (2.12) for $\psi = \xi$ follows from similar formulas for ψ such that $\psi \leq \xi$ and $\psi \sim \xi$. Proof of Lemma is completed by the trivial induction on the finite set $\{w\xi \mid w \in W\}$. $\square$

2.6. Description of finite dimensional representations

In this section we will describe all finite-dimensional representations of a semisimple Lie algebra $\mathfrak{g}$. As was shown in § 2.2, all the highest weights of such representations belong to Θ. Recall that in § 2.3 we have constructed a finite dimensional $\mathfrak{g}$-module L^φ for any $\varphi \in \overline{\mathcal{D}}$.

2.6.1. Theorem. 1) *Let M be a finite dimensional $\mathfrak{g}$-module. Then M is isomorphic to the direct sum of modules of the form L^φ for $\varphi \in \overline{\mathcal{D}}$.*

2) *All the modules L^φ, where $\varphi \in \overline{\mathcal{D}}$, are irreducible.*

Proof. 1) We may assume that $M = M^\theta$, where $\theta \in \Theta$. Let m be any highest weight vector of M and φ its weight. Then $\theta = \theta_{\varphi+\rho}$. Besides, since $E_\alpha^k m = 0$ for k large and for $\alpha \in B$, then Lemma 2.2.1 implies that $\sigma_\alpha(\varphi + \rho) < \varphi + \rho$. Therefore, $\varphi + \rho \in \mathfrak{h}^*_{\mathbb{Z}} \cap \overline{C}$, i.e., $\varphi \in \overline{\mathcal{D}}$.

Thus, φ is uniquely recovered from the element θ by conditions

$$\theta_{\rho+\varphi} = \theta \quad \text{and} \quad \varphi + \rho \in C.$$

Let $m_1, \dots, m_l$ be a basis of $M(\varphi)$. Let us construct the map

$$\tau\colon \bigoplus_{1 \le i \le l} (M^{\varphi+\rho})_i \longrightarrow M$$

so that each generator $(m_{\varphi+\rho})_i$ for $i = 1, 2, \dots, l$ is mapped to m. As follows from Lemma 2.1, for any $\alpha \in B$, we have

$$E_{-\alpha}^{k_\alpha} m_i = 0, \quad \text{where } k_\alpha = \frac{2\langle \varphi + \rho, \alpha \rangle}{\langle \alpha, \alpha \rangle}.$$

Hence, τ may be considered as the map

$$\tau\colon \bigoplus_{1 \le i \le l} (L^\varphi)_i \longrightarrow M.$$

Let L_1 and L_2 be the kernel and cokernel (i.e., $L_2 = M/\operatorname{Im}\tau$) of this map. Then $\Theta(L_i) = \{\theta\}$, and hence $L_i(\varphi) = 0$, where $i = 1, 2$. As was shown above, $L_1 = L_2 = 0$, i.e., $M = \bigoplus\limits_{1 \le i \le l} (L^\varphi)_i$.

2) Let M be a non-trivial submodule of L^φ. Then $\Theta(M) = \theta_{\varphi+\rho}$, hence, $M(\varphi) \neq 0$, i.e., M contains an element of the form $m_{\varphi+\rho}$. But then $M = L^\varphi$. Thus, the module L^φ is irreducible. □

Corollary. $L^\varphi = L^{\varphi+\rho}$, *where* $\varphi \in \overline{\mathcal{D}}$.

2.7. The Kostant formula for the multiplicity of the weight

In $\mathcal{E}$ (see subsect. 2.1.6), introduce the convolution by setting

$$(u * v)(\xi) = \sum_{\varphi \in \mathfrak{h}^*} u(\varphi) v(\xi - \varphi) \text{ for any } u, v \in \mathcal{E}. \tag{2.13}$$

(Observe, that this sum only contains a finite number of non-zero terms.) Clearly, $u * v \in \mathcal{E}$. The convolution endows $\mathcal{E}$ with a commutative algebra structure.

Define the W-action in the space of functions on $\mathfrak{h}^*$ by setting

$$(wu)(\xi) = u(w^{-1}\xi) \text{ for any } w \in W,\ \xi \in \mathfrak{h}^*,\ u \in S(\mathfrak{h}).$$

For any $\xi \in \mathfrak{h}^*$, define $\delta_\xi \in \mathcal{E}$ by setting

$$\delta_\xi(\varphi) = 0 \quad \text{for any } \varphi \neq \xi \text{ and } \delta_\xi(\xi) = 1.$$

Set

$$D = \prod_{\gamma \in R_+} (\delta_{\gamma/2} - \delta_{-\gamma/2}), \quad \text{where } \prod \text{ is the product in } \mathcal{E}, \text{ not convolution.}$$

Clearly, δ_0 is the unit of $\mathcal{E}$.

2.7.1. Lemma. *Let K be the Kostant function. Then*

$$K * \delta_{-\rho} * D = \delta_0. \tag{2.14}$$

Proof. For any $\gamma \in R_+$, set $a_\gamma = \delta_0 + \delta_{-\gamma} + \ldots + \delta_{-n\gamma} + \ldots$. The definition of K implies that

$$K = \prod_{\gamma \in R_+} a_\gamma.$$

Further, $(\delta_0 - \delta_{-\gamma})a_\gamma = \delta_0$. Since D can be represented as $\prod\limits_{\gamma \in R_+} (\delta_0 - \delta_{-\gamma})\delta_\rho$, we are done. □

Lemmas 2.7.1 and 2.1.6 imply the following lemma.

2.7.2. Lemma.

$$D * \pi_{M^\xi} = \delta_\xi \text{ for any } \xi \in \mathfrak{h}^*. \tag{2.15}$$

2.7.3. Lemma. $wD = \det w \cdot D$ *for any* $w \in W$.

Proof. It suffices to verify that $\delta_\alpha D = -D$ for $\alpha \in B$. Since δ_α permutes the elements of the set $R_+ \setminus \{\alpha\}$ and transforms α into $-\alpha$, it follows that

$$\delta_\alpha D(\delta_{-\alpha/2} - \delta_{\alpha/2}) \prod_{\gamma \in R_+ \setminus \{\alpha\}} (\delta_{\gamma/2} - \delta_{-\gamma/2}) = -D. \quad \square$$

Denoted the ring of $\mathbb{Z}$-valued functions on $\mathfrak{h}^*_{\mathbb{Z}}$ by $\mathcal{E}_{\mathbb{Z}}$.

2.7.4. Theorem. *The map $\tau\colon K(\mathcal{O}) \longrightarrow \mathcal{E}_{\mathbb{Z}}$ defined by the formula $\tau([M]) = D * \pi_M$, where $M \in \mathcal{O}$, is an isomorphism.*

Proof. By Lemmas 2.7.2 and 2.5.6 we have $D * \pi_M \in \mathcal{E}_{\mathbb{Z}}$. Define the map $\eta\colon \mathcal{E}_{\mathbb{Z}} \longrightarrow K(\mathcal{O})$ by setting

$$\eta(u) = \sum_{\xi \in \mathfrak{h}^*_{\mathbb{Z}}} u(\xi)[M^\xi]. \tag{2.16}$$

By Lemma 2.5.6 the map η is an epimorphism. Lemma 2.7.2 implies that η is the identity map. Hence, τ is an isomorphism. □

2.7.5. Theorem. $D * \pi_{L^\varphi} = \sum\limits_{w \in W} \det w \cdot \delta_{w(\varphi+\rho)}$.

Proof. Since $L^\varphi = L^{\varphi+\rho}$, it follows that eq. (2.12) implies that

$$[L^\varphi] = [M^{\varphi+\rho}] + \sum_{\psi \sim \varphi+\rho,\ \psi < \varphi+\rho} c_\psi [M^\psi], \text{ where } c_\psi \in \mathbb{Z}.$$

Hence, Lemma 2.7.2 shows that

$$D * \pi_{L^\varphi} = \delta_{\varphi+\rho} + \sum_{\psi \sim \varphi+\rho,\ \psi < \varphi+\rho} c_\psi \delta_\psi.$$

Since $w\pi_{L^\varphi} = \pi_{L^\varphi}$ and $wD = \det w \cdot D$ for any $w \in W$, we see that

$$w(D * \pi_{L^\varphi}) = \det w \cdot D * \pi_{L^\varphi}.$$

Hence,

$$D * \pi_{L^\varphi} = \sum_{w\in W} \det w \cdot \delta_{w(\varphi+\rho)}. \quad \square \tag{2.17}$$

Corollary. 1)

$$[L^\varphi] = \sum_{w\in W} (\det w)[M^{w(\varphi+\rho)}] \textit{ for any } \varphi \in \overline{\mathcal{D}}. \tag{2.18}$$

2) (the Kostant formula for the multiplicity of the weight)

$$\pi_{L^\varphi}(\psi) = \sum_{w\in W} \det w \cdot K(\psi + \rho - w(\varphi+\rho)) \textit{ for any } \psi \in \mathfrak{h}^*. \tag{2.19}$$

Proof. Let us multiply both parts of (2.17) by $K * \delta_{-\rho}$ and apply Lemma 2.7.1. $\square$

Denote by $S[[\mathfrak{h}^*]]$ the ring of formal power series in the elements of $\mathfrak{h}^*$ with complex coefficients. For any $\xi \in \mathfrak{h}^*$, set $e^\xi = \sum\limits_{i\geq 0} \dfrac{\xi^i}{i!}$.

Clearly, $e^\xi \in S[[\mathfrak{h}^*]]$ and $e^{\xi+\psi} = e^\xi e^\psi$ for any $\xi, \psi \in \mathfrak{h}^*$. Let M be a finite dimensional $\mathfrak{g}$-module. Define the *character* $\mathrm{ch}_M \in S[[\mathfrak{h}^*]]$ of M by the formula

$$\mathrm{ch}_M = \sum_{\xi\in\mathfrak{h}^*_{\mathbb{Z}}} \pi_M(\xi) e^\xi. \tag{2.20}$$

2.7.6. Theorem. *Set*

$$\tilde{D} = \sum_{w\in W} (\det w) e^{w\rho}.$$

For L^φ, where $\varphi \in \overline{\mathcal{D}}$, we have

$$\tilde{D}\,\mathrm{ch}_{L^\varphi} = \sum_{w\in W} (\det w) e^{w(\varphi+\rho)}. \tag{2.21}$$

Proof. The map $j\colon \mathcal{E}_{\mathbb{Z}} \longrightarrow S[[\mathfrak{h}^*]]$ defined by the formula $j(u) = \sum\limits_{\xi\in\mathfrak{h}^*_{\mathbb{Z}}} u(\xi)e^\xi$ is a ring homomorphism. Inserting $\varphi = 0$ in (2.17) we obtain

$$\sum_{w\in W} \det w \cdot \delta_{w\rho} = D * \pi_{L^0} = D * \delta_0 = D. \tag{2.22}$$

Hence, $j(D) = \tilde{D}$. Theorem 2.7.6 now easily follows from Theorem 2.7.5. $\square$

Remark. 1) All the power series of Theorem 2.7.6 converge and define analytic functions on $\mathfrak{h}$. Theorem 2.7.6 claims the equality of two such functions.

2) Let $\mathcal{G}$ be a complex semisimple Lie group with Lie algebra $\mathfrak{g}$ and $\mathcal{H} \subset \mathcal{G}$ the Cartan subgroup corresponding to the Lie subalgebra $\mathfrak{h}$. Consider the finite dimensional representation T of $\mathcal{G}$ corresponding to the $\mathfrak{g}$-action in L^φ. Let $h \in \mathcal{H}$. Then $h = \exp(H)$, where $H \in \mathfrak{h}$. It is easy to derive from Theorem 2.7.6 that

$$\operatorname{tr}\ T(h) = \frac{\sum\limits_{w\in W} \det w \cdot e^{w(\varphi+\rho)(H)}}{\sum\limits_{w\in W} \det w \cdot e^{w\rho(H)}}. \tag{2.23}$$

This is the well-known *H. Weyl's formula for characters of irreducible representations of semi-simple complex Lie groups.*

2.7.7. Theorem. *Let $\varphi \in \overline{\mathcal{D}}$. Then*

$$\dim L^\varphi = \prod_{\gamma\in R_+} \frac{\langle \varphi+\rho, \gamma\rangle}{\langle \rho, \gamma\rangle}. \tag{2.24}$$

Proof. Set

$$F_\xi = \sum_{w\in W} \det w \cdot e^{w\xi} \text{ for any } \xi \in \mathfrak{h}^*. \tag{2.25}$$

Clearly, $F_\rho = \tilde{D} = \prod\limits_{\gamma\in R_+} (e^{\gamma/2} - e^{-\gamma/2})$. For any $\xi \in \mathfrak{h}^*$ and $H \in \mathfrak{h}$, we may consider $F_\xi(tH)$ as the formal power series in one indeterminate t.

Let h_ρ and h_φ be elements of $\mathfrak{h}$ corresponding to ρ and φ, respectively, after the identification of $\mathfrak{h}$ with $\mathfrak{h}^*$ by means of the Killing form. Then

$$\dim L^\varphi = \operatorname{ch} L^\varphi(0) = \frac{F_{\varphi+\rho}(th_\rho)}{F_\rho(th_\rho)}\Big|_{t=0}. \tag{2.26}$$

Observe that

$$F_{\varphi+\rho}(th_\rho) = \sum_{w\in W} \det w \cdot e^{t\langle\varphi+\rho,\rho\rangle} = F_\rho(t(h_\varphi + h_\rho)). \tag{2.27}$$

Hence,

$$\begin{gathered}\dim L^\varphi = \frac{\tilde{D}(t(h_\varphi+h_\rho))}{\tilde{D}(t\rho)}\Big|_{t=0} = \\ \prod_{\gamma\in R_+}\left(\frac{e^{\frac{t}{2}(\gamma(h_\varphi+h_\rho))} - e^{-\frac{t}{2}(\gamma(h_\varphi+h_\rho))}}{e^{\frac{t}{2}(\gamma(h_\rho))} - e^{-\frac{t}{2}(\gamma(h_\rho))}}\Big|_{t=0}\right) = \prod_{\gamma\in R_+} \frac{\langle\gamma, \varphi+\rho\rangle}{\langle\gamma,\rho\rangle}. \quad \square\end{gathered} \tag{2.28}$$

7.8. Summary. 1) For any weight $\varphi \in \overline{\mathfrak{D}}$, there exists a finite dimensional irreducible $\mathfrak{g}$-module L^φ. All such modules are non-isomorphic. Any finite dimensional irreducible $\mathfrak{g}$-module is isomorphic to one of L^φ, where $\varphi \in \overline{\mathfrak{D}}$.

2) Any finite dimensional $\mathfrak{g}$-module M is isomorphic to the direct sum of the modules L^φ.

3) The module L^φ is $\mathfrak{h}$-diagonalizable and has a unique (up to a factor) highest weight vector l_φ. The weight of l_φ is equal to φ.

4) (*The Harish–Chandra theorem on ideal.*) The module L^φ is generated as $U(\mathfrak{g}_-)$-module by the vector l_φ (in particular, all the weights of L^φ are $\leq \varphi$). The ideal of relations $I = \{X \in U(\mathfrak{g}_-) \mid Xl_\varphi = 0\}$ is generated by the elements

$$E_{-\alpha}^{m_\alpha+1}, \text{ and } H_\alpha - m_\alpha, \text{ where } m_\alpha = \frac{2\langle\varphi,\alpha\rangle}{\langle\alpha,\alpha\rangle} = \varphi(H_\alpha) \text{ and where } \alpha \in B.$$

5) The function π_{L^φ} is W-invariant.

6) If ψ is a weight of L^φ, then either $\varphi \sim \psi$ or $|\psi| < |\varphi|$.

7) For any $a \in L^\varphi$ and $z \in Z(U)$, we have $z_a = \theta_{\varphi+\rho}(z)a$. If $\varphi_1, \varphi_2 \in \overline{\mathfrak{D}}$ and $\varphi_1 \neq \varphi_2$, then the homomorphisms $\theta_{\varphi_1+\rho}$ and $\theta_{\varphi_2+\rho}$ are distinct.

8) (The *Kostant formula.*)

$$\pi_{L^\varphi}(\mu) = \sum_{w\in W} (\det w) K(\mu + \rho - w(\varphi+\rho)).$$

9) $\dim L^\varphi = \prod\limits_{\gamma\in R_+} \dfrac{\langle\gamma,\varphi+\rho\rangle}{\langle\gamma,\rho\rangle}$.

2.8. Proof of the Harish–Chandra theorem

In this section $\mathfrak{g}$ is a simple (finite dimensional) Lie algebra, and hence $\mathfrak{g} \simeq \mathfrak{g}^*$. Clearly, the map $\mathrm{symm}\colon S(\mathfrak{g}^*)^{\mathfrak{g}} \longrightarrow Z(U)$, where $S(\mathfrak{g}^*)^{\mathfrak{g}}$ is the subring of $\mathfrak{g}$-invariant polynomials, is an isomorphism. Therefore, it suffices to verify that

$$\tau = j \cdot \mathrm{symm}\colon S(\mathfrak{g}^*)^{\mathfrak{g}} \longrightarrow S(\mathfrak{h})^W$$

is an isomorphism. Consider the embedding $i\colon \mathfrak{h}^* \longrightarrow \mathfrak{g}^*$ defined by the formulas

$$i(\xi)(E_\gamma) = 0 \text{ and } i(\xi)(H) = \xi(H) \text{ for any } \gamma \in R \text{ nd } H \in \mathfrak{h}.$$

Let $\tau'\colon S(\mathfrak{g}) \longrightarrow S(\mathfrak{h})$ be the corresponding restriction map

$$\tau'(H) = H \text{ for any } H \in \mathfrak{h} \text{ and and } \tau'(E_\gamma) = 0 \text{ for any } \gamma \in R.$$

We have constructed maps $\tau\colon S(\mathfrak{g}^*)^{\mathfrak{g}} \longrightarrow S(\mathfrak{h})$ and $\tau'\colon S(\mathfrak{g}) \longrightarrow S(\mathfrak{h})$. Since

$$S(\mathfrak{g}^*)^{\mathfrak{g}} \subset S(\mathfrak{g}^*) \cong S(\mathfrak{g})$$

(due to non-degeneracy of the Killing form), we may consider τ' as the map $S(\mathfrak{g}^*)^{\mathfrak{g}} \longrightarrow S(\mathfrak{h})$.

Let us prove that the values of τ and τ' coincide on the leading terms. More precisely, we will prove that if $X \subset S(\mathfrak{g}^*)^{\mathfrak{g}}$ is of degree n, then $\deg(\tau(X) - \tau'(X)) < n$. For any $n \in \mathbb{Z}$ and any (finite dimensional) vector space V, let $S_n(V)$ be the subspace in $S(V)$ consisting of all polynomials of degree $\leq n$.

The following facts are straightforward:

a) symm $Y_N = X_N \mod U(\mathfrak{g})_{|N|-1}$, where

$$|N| = n_1 + \ldots + n_s + m_1 + \ldots + m_r + n'_1 + \ldots + n'_s.$$

b) Let $X = \sum\limits_{n \leq |N|} C_N X_N$. Then

$$\mu(z) = \sum_{N=(0,\ldots,0,m_1,\ldots,m_r,0,\ldots,0)} C_N H_1^{m_1} \ldots H_r^{m_r} \mod S_{n-1}(\mathfrak{h}).$$

c) If $\mu(z) \in S_n(\mathfrak{h})$, then $\mu(z) - j_z \in S_{n-1}(\mathfrak{h})$.

d) $\tau'(Y_N) = 0$ if at least one of n_i, n'_i is non-zero and

$$\tau'(Y_N) = H_1^{m_1} \ldots H_r^{m_r} \quad \text{if } N = (0, \ldots, 0, m_1, \ldots, m_r, 0 \ldots, 0).$$

Hence, the values of τ and τ' coincide on the leading terms.

Let us prove that $\tau'(S(\mathfrak{g}^*)^{\mathfrak{g}}) \subset S(\mathfrak{h})^W$. Let $Y \in S(\mathfrak{g}^*)^{\mathfrak{g}}$. We may assume that Y is a homogeneous polynomial of degree n. Then $\tau'(Y)$ is also a homogeneous polynomial of degree n. Since $\tau(Y) \in S(\mathfrak{h})^W$ and homogeneous components of degree n of $\tau(Y)$ and $\tau'(Y)$ coincide, it follows that $\tau'(Y) \in S(\mathfrak{h})^W$.

Thus, we have two linear maps $\tau, \tau' \colon S(\mathfrak{g}^*)^{\mathfrak{g}} \longrightarrow S(\mathfrak{h})^W$ that coincide on the leading terms. Routine considerations show that if one of these maps is an isomorphism, then so is the other one. Therefore, to prove the Harish–Chandra theorem, it only suffices to show that $\tau' \colon S(\mathfrak{g}^*)^{\mathfrak{g}} \longrightarrow S(\mathfrak{h})^W$ is an isomorphism.

Let us identify $\mathfrak{g}$ with $\mathfrak{g}^*$ and $\mathfrak{h}$ with $\mathfrak{h}^*$ by means of the Killing form. Hence, $S(\mathfrak{g}^*)$ is identified with $S(\mathfrak{g})$, and $S(\mathfrak{h}^*)$ with $S(\mathfrak{h})$. With the help of τ', the restriction homomorphism $\eta \colon S(\mathfrak{g}^*) \longrightarrow S(\mathfrak{h}^*)$ and the representation of $\mathfrak{g}$ in $S(\mathfrak{g})$, we can describe the coadjoint representation of $\mathfrak{g}$ in $S(\mathfrak{g}^*)$ as follows:

a) $X(F)(Y) = F([X, Y])$ for any $F \in \mathfrak{g}^* \subset S(\mathfrak{g}^*)$ and $X, Y \in \mathfrak{g}$. (Here we use the invariance of the Killing form B, i.e., the identity

$$B([X, Y], Z) + B(Y, [X, Z]) = 0$$

valid for any $X, Y, Z \in \mathfrak{g}$.)

b) $X(T) = \sum\limits_{1\leq i\leq k} F_1 \dots F_{i-1} X(F_i) F_{i+1} \dots F_k$ for any monomial

$$T = F_1 \dots F_k \in S^k(\mathfrak{g}^*),$$

where $F_i \in \mathfrak{g}^*$ for all i, and any $X \in \mathfrak{g}$. Clearly, under the identification $\mathfrak{g} \simeq \mathfrak{g}^*$ the subring $S(\mathfrak{g}^*)^{\mathfrak{g}} \subset S(\mathfrak{g}^*)$ of $\mathfrak{g}$-invariant polynomials on $\mathfrak{g}$ turns into the subring $S(\mathfrak{g})^{\mathfrak{g}} \subset S(\mathfrak{g})$ of $\mathfrak{g}$-invariant polynomials on $\mathfrak{g}^*$. Hence, to prove that $\tau'\colon S(\mathfrak{g}^*)^{\mathfrak{g}} \longrightarrow S(\mathfrak{h})^W$ is an isomorphism, it only suffices to prove the following theorem.

2.8.1. Theorem (The Chevalley theorem)**.** *Let $S(\mathfrak{g}^*)$ and $S(\mathfrak{h}^*)$ be polynomial rings on $\mathfrak{g}$ and $\mathfrak{h}$, respectively, and $\eta\colon S(\mathfrak{g}^*) \longrightarrow S(\mathfrak{h}^*)$ the restriction homomorphism. Then $S(\mathfrak{g}^*)^{\mathfrak{g}} \longrightarrow S^W(\mathfrak{h}^*)$ is isomorphism.*

Proof. While proving the Harish–Chandra theorem we have already established that $\eta(S(\mathfrak{g}^*)^{\mathfrak{g}}) \subset S^W(\mathfrak{h}^*)$. Let us prove that $\operatorname{Ker}(\eta|_{S(\mathfrak{g}^*)^{\mathfrak{g}}}) = 0$.

2.8.1.1. Lemma. *Let $X, Y \in \mathfrak{g}$ and $T \in S(\mathfrak{g}^*)$. Let ad_X be nilpotent. Then*

$$e^X(T)(Y) = T(e^{-\operatorname{ad}_X}(Y)).$$

(Here $e^X := \sum\limits_{n\geq 0} \frac{X^n}{n!}$ and the sum is well-defined since it only has a finite number of non-zero terms.)

Proof. It suffices to consider the case where $T = F_1 \dots F_k$ for any $F_i \in \mathfrak{g}^*$. Then

$$\Big(\sum \frac{X^n}{n!}(T)\Big)(Y) = \sum_{m_1,\dots,m_k \in \mathbb{Z}_+} \frac{1}{m_1! \dots m_k!} F_1[(-\operatorname{ad}_X)^{m_1}(Y)] \dots$$
$$\dots F_k[(-\operatorname{ad}_X)^{m_k}(Y)] = \prod_{1\leq i\leq k} F_i(e^{-\operatorname{ad}_X}(Y)) = T(e^{-\operatorname{ad}_X}(Y)). \qquad \square$$

2.8.1.2. Lemma. *Let $T \in S(\mathfrak{g}^*)$. Then $T \in S(\mathfrak{g}^*)^{\mathfrak{g}}$ if and only if*

$$T(Y) = T(e^{\operatorname{ad}_X}(Y)) \text{ for any } X, Y \text{ such that } \operatorname{ad}_X \text{ is nilpotent.} \tag{2.29}$$

Proof. If $T \in S(\mathfrak{g}^*)^{\mathfrak{g}}$ then eq. (2.29) follows from Lemma 2.8.1.1. Conversely, let $X \in \mathfrak{g}$ and ad_X be nilpotent. Then, for any $Y \in \mathfrak{g}$, we have

$$T(Y) = T(e^{\operatorname{ad}_{tX}}(Y)) = (e^{tX}T)(Y) \text{ for any } t \in \mathbb{C}.$$

Comparing the coefficients of the first degree in t in this equality we see that $X(T) = 0$. Hence, $E_\gamma(T) = 0$ and $H_\gamma(T) = [E_{-\gamma}, E_\gamma](T) = 0$ for any $\gamma \in R$. $\square$

Let $T \in S(\mathfrak{g}^*)^{\mathfrak{g}}$ and $\tau(T) = 0$, i.e., $T|_{\mathfrak{h}} = 0$. Let us prove that $T = 0$.

Consider the map $\mathfrak{g}_- \oplus \mathfrak{h} \oplus \mathfrak{g}_+ \longrightarrow \mathfrak{g}$ defined by the formula $(X, H, Y) \mapsto e^{\operatorname{ad}_X} e^{\operatorname{ad}_Y}(H)$. The polynomial T vanishes identically on the image of this map. The value of the Jacobian of this map at the point $\Bigl(0, \sum_{\gamma\in R_+} H_\gamma, 0\Bigr)$ is non-zero, and therefore the image of this map is open in $\mathfrak{g}$. Hence, $T = 0$.

Set $S = \eta(S(\mathfrak{g}^*)^{\mathfrak{g}})$. We must show that $S = S^W(\mathfrak{h}^*)$.

Let $S' \subset S[[\mathfrak{h}^*]]$ be the subring in the ring of formal power series on $\mathfrak{h}$ consisting of series such that all their homogeneous components belong to S.

2.8.1.3. Lemma. *For any $\xi \in \mathfrak{h}^*$, set $e^\xi = \sum_{n\geq 0} \frac{\xi^n}{n!}$. Then any W-invariant finite sum $\sum_{\xi\in\mathfrak{h}^*_{\mathbb{Z}}} c_\xi e^\xi$ belongs to S'.*

Proof. First, let us show that, for any $\varphi \in \overline{\mathcal{D}}$, the series

$$\operatorname{ch} L^\varphi = \sum_{\xi\in\mathfrak{h}^*_{\mathbb{Z}}} \pi_{L^\varphi}(\xi) e^\xi$$

belongs to S'. Indeed,

$$\sum_{\xi\in\mathfrak{h}^*_{\mathbb{Z}}} \pi_{L^\varphi}(\xi) e^\xi = \sum_{n\geq 0} \frac{1}{n!} \sum_{\xi\in\mathfrak{h}^*_{\mathbb{Z}}} \pi_{L^\varphi}(\xi)\xi^n.$$

But $\sum_{\xi} \pi_{L^\varphi}(\xi)\xi^n$ is the restriction onto $\mathfrak{h}$ of the function $\xi'_{\varphi,n}(X) = \operatorname{tr}\ \xi_\varphi(X)^n$, where $X \in \mathfrak{g}$ and ξ_φ is the representation of $\mathfrak{g}$ in L^φ. The function $\xi_{\varphi,n}$ is $\mathfrak{g}$-invariant and by Lemma 8.2 belongs to $S(\mathfrak{g}^*)^{\mathfrak{g}}$. Hence,

$$\sum_{\xi\in\mathfrak{h}^*} \pi_{L^\varphi}(\xi) e^\xi \in S'.$$

Let $s = \sum_{\xi\in\mathfrak{h}^*_{\mathbb{Z}}} c_\xi e^\xi$ be a finite W-invariant sum and $r_s = \max_{c_\xi\neq 0} |\xi|$. The value of r_s may only belong to a discrete set of non-negative numbers. If $r_s = 0$, then $s = c_0 \in P'$.

Let Lemma 2.8.1.3 be proved for any s' such that $r_{s'} < r_s$. Set

$$s' = s - \sum_{\xi\in\mathfrak{h}^*_{\mathbb{Z}}\cap\overline{C}} c_\xi \operatorname{ch}_\xi .$$

Since any weight of L^φ is either equivalent to φ or its "height" is less than that of φ (Theorem 2.3.3), we see that $r_{s'} < r_s$. Therefore, Lemma 2.8.1.3 is proved for $s = s' - \sum c_\xi \operatorname{ch}_\xi$. □

Now we can complete the proof of the Chevalley theorem. Let T be a homogeneous polynomial $S^W(\mathfrak{h}^*)$ of degree n. We may consider T as the polynomial in α_i, where $\alpha_i \in B$. Consider the formal power series

$$T' = \frac{1}{\mid W\mid} \sum_{w\in W} wT(e^{\alpha_i} - 1).$$

By Lemma 2.8.1.3, $T' \in S'$. But clearly, all homogeneous components of T' of degree less than n vanish and the homogeneous term of degree n coincides with T. Hence, $T \in S$ and the Chevalley theorem is proved together with the Harish–Chandra theorem. □

References

[Bu] Bourbaki N., *Groupes et algébres de Lie*. Hermann, Paris, Ch. II–III: 1972; Ch. IV–VI: 1968; Ch. VII–VIII: 1974

[Di] Dixmier J., *Enveloping algebras*, AMS, 1996

[FH] Fulton W., Harris J., *Representation Theory. A First Course* Corr. 5th printing, Springer, New York, 1999, XV, 551 p.

[OV] Onishchik A., Vinberg E., *Lie groups and Algebraic Groups*, Springer, 1987.

[P] Prasolov V., *Problems and theorems in linear algebra*, AMS, 1991, 226 pp.

[S] Serre J.-P. *Lie algebras and Lie groups*. N.Y, W. A. Benjamin, 1965; *Algébres de Lie semi-simples complexes*. N.Y, W. A. Benjamin, 1966

[St] Steinberg, R. *Lectures on Chevalley groups*. Notes prepared by John Faulkner and Robert Wilson. Yale University, New Haven, Conn., 1968. iii+277 pp.

[W] van der Waerden, B. L. *Algebra*. Vol. I. Based in part on lectures by E. Artin and E. Noether. Translated from the seventh German edition by Fred Blum and John R. Schulenberger. Springer, New York, 1991. xiv+265 pp. ; Vol. II., 1991. xii+284 pp.

Chapter 3

On evolution of magnetic field under translation and diffusion (V. I. Arnold)

3.1. A problem of a stationary magnetic dynamo

The evolution equation of the divergence-free magnetic field H translated by a divergence-free flow of velocity v and with diffusion coefficient μ is of the form

$$\dot{H} = \{v, H\} + \mu \Delta H, \tag{3.1}$$

where $\Delta = \operatorname{rot}\operatorname{rot}$ is the *Laplacian* and $\{\cdot\,,\cdot\}$ is the *Poisson bracket.* We consider this equation with periodic boundary conditions (H does not change under translations by integer multiples of 2π along any of the three coordinate axes). The field v is supposed to be fixed and of the same periodicity, namely

$$v = (\cos y + \sin z)\frac{\partial}{\partial x} + (\cos z + \sin x)\frac{\partial}{\partial y} + (\cos x + \sin y)\frac{\partial}{\partial z}. \tag{3.2}$$

We are interested to find out how does the increment $\gamma = \operatorname{Re}\lambda$ of the mode of fastest growth $H = e^{\lambda t} H_0(x, y, z)$ depend on the magnetic Reynolds number $R := \dfrac{1}{\mu}$. For a fixed Reynolds number, the field v is called a *dynamo* if there exists a growing mode ($\gamma > 0$). The dynamo is said to be *strong* if the increment remains bounded below by a positive constant as the magnetic viscosity decays, i.e., if

$$\lim_{\mu \to 0} \gamma(\mu) > 0. \tag{3.3}$$

Nobody knows if a strong dynamo exists.

No two-dimensional flow can be a strong dynamo since in such a flow the particles can not scatter exponentially.

The particular form of the flow (3.2) was selected for the following reasons. At $\mu = 0$, eq. (3.1) expresses the fact that the field H is frozen in the liquid. If the field v exponentially scatter the particles of the liquid (i.e., if the greatest characteristic Lyapunov exponent [1)] is positive on a set of positive measure), then the frozen field H grows exponentially (is expanded by the flow).

[1] http://en.wikipedia.org/wiki/Lyapunov_exponent.

From numerical experiments, see [1], we conclude that the field (3.2) is one of the simplest fields with an exponential scattering of the particles. Conjecturally, the scattering of the particles under the influence of the "generic" field follows approximately the same pattern as for the flow (3.2).

Since the flow scatters the particles inhomogeneously, the growing field H becomes jagged. The viscosity quickly dampens the higher harmonics. Will the dynamo work (will it be that $\gamma > 0$) depends on which of the processes is faster: the growth of the frozen field or the viscous damping.

3.2. The results of a numerical experiment

E. I. Korkina studied the dependence of the increment γ on the magnetic Reynolds number $R := \frac{1}{\mu}$ be means of Galerkin [2)] approximations. She was looking for an eigenvalue of the operator

$$A_R(H) := R\{v, H\} + \Delta H \tag{3.4}$$

with the greatest real part. She expanded the eigenvector H into the Fourier series and ignored the higher harmonics.

The computations were performed up to $R = 19$; this required to take into account the harmonics $e^{(k,r)}$ with $k^2 \leq 169$. Therefore the size of the matrix whose eigenvalue was being sought was around 20,000.

Korkina controlled the results of her computations by several methods: By changing the order of the harmonics taken into account, by preserving the divergence-free property of the fields under evolution and by keeping the symmetry conditions I will retell further on.

It turned out that one has to take into account the harmonics with $k^2 \leq 9R$.

Korkina started her computations with small Reynolds numbers in order to use the found eigenvector as a first approximation for computations with the bigger Reynolds numbers. It turned out, however, that for small Reynolds numbers R, the eigenvalue λ of A_R does not depend on R: It is equal to 0.

The reason for this phenomenon is that for high viscosity, the solution of eq. (3.1) becomes stationary as $t \longrightarrow \infty$ and is only determined by the cohomology class of the initial field H (i.e., by the vector of the mean values of the field). The passage to the stationary regime is discussed in detail in [2] in a more general setting of the evolution of k-forms on an n-dimensional manifold.

To get rid of this effect, we confined us to the fields with zero mean. Still, even in this case, the computations gave the eigenvalue $\lambda = -1$ of A_R; this eigenvalue does not depend on R as long as R remains smaller than a critical value, $R_1 \approx 2.3$.

[2] http://en.wikipedia.org/wiki/Galerkin_method.

The reason for this is that $\{v, v\} = 0$, so due to eq. (3.2) we have $\Delta v = -v$, so v itself is the eigenvector of A_R with eigenvalue $\lambda = -1$.

As the Reynolds number grows, there appears a pair of complex eigenvalues with $\gamma > -1$. This pair moves to the right and intersects the boundary of dynamo (i.e., $\gamma = 0$) at $R_2 \approx 9.0$. The largest value of γ on this branch is attained at $R_3 \approx 12.5$ and is equal to ≈ 0.096. Then the increment decays and at $R_4 \approx 17.5$ it becomes negative again.

Therefore, the field v is a dynamo if $R_2 < R < R_4$. The problem *if this field is dynamo as $R \longrightarrow \infty$ remains open.* D. Galloway [3)] found a dynamo for $30 < R < 100$.

The distribution of energy of the eigenvector found along the harmonics reveals a strange anomaly: For certain “empty” values of k^2 we see that $\sum |H_k|^2$ is exactly equal to 0, whereas for certain “wonderful” values of k^2 the sum is equal to 0 up to accuracy of computations.

The “empty” values are:

$$k^2 = 7,\ 15,\ 23,\ 28,\ 31,\ \ldots \tag{3.5}$$

The “wonderful” values are

$$k^2 = 3,\ 4,\ 12,\ 16,\ 48,\ 64,\ \ldots \tag{3.6}$$

The “empty” values are easy to explain: They are the numbers of the form $4^a(8b+7)$. As is known, these and only these integers can not be represented as the sums of three squares.

The “wonderful” values are of the form 7^a and $3 \cdot 4^a$. In what follows we will explain their appearance from symmetry considerations: They are related with the decomposition of the representation of the group of rotations of the cube in the space of vector-valued trigonometric polynomials into irreducibles.

These considerations allow one to hasten computations tenfold and even hundredfold. For example, we (Korkina and me) were able to find the eigenvector of the Galerkin approximation (with $k^2 \leq 5$ being taken into account) explicitly, whereas without symmetry considerations we would have had to consider a 112×112 matrix. We were even able to find the exact value of the first harmonic of the true eigenvector of the mode under the study for any Reynolds numbers:

$$H_1 = (\cos y - \sin z)\frac{\partial}{\partial x} + (\cos z - \sin x)\frac{\partial}{\partial y} + (\cos x - \sin y)\frac{\partial}{\partial z}. \tag{3.7}$$

---

[3] Galloway D.J., Hollerbach R., Proctor M.R.E. Fine structures in fast dynamo computations, Small scale structures in three-dimensional hydrodynamic and magnetohydrodynamic turbulence, Small scale structures in three-dimensional hydrodynamic and magnetohydrodynamic turbulence, Maurice Meneguzzi, Annick Pouquet and Pierre-Louis Sulem (eds.), Lecture Notes in Physics 462, Springer-Verlag, Berlin, (1995), 341–346.

Galloway D.J. Fast dynamos, Advances in nonlinear dynamos, The fluid mechanics of astrophysics and geophysics, Taylor & Francis, London, (2003), 37–59.

The fact that it is this very mode that grows fastest (has the greatest increment) is, however, obtained only as a result of a computer experiment and only for $R \leq 19$.

3.3. A symmetry of the velocity field

It is clear from eq. (3.2) that the cyclic permutation of coordinates sends v into itself. It is easy to prove the following

3.3.1. Theorem. *The symmetry group of the field v preserving Δ contains 24 elements and is isomorphic to the group G of rotations of the cube.*

Indeed, it is easy to check that the field v turns into itself under the transformations

$$g_4\colon (x, y, z) \mapsto (x + \frac{5}{2}, z - \frac{\pi}{2}, \frac{\pi}{2} - y). \tag{3.8}$$

The order of g_4 is equal to 4, i.e., $g_4^4 = \mathrm{id}$. Together with the cyclic permutation of coordinates

$$g_3\colon (x, y, z) \mapsto (y, z, x) \tag{3.9}$$

the transformation g_4 generates the group G isomorphic to the group of rotations of the cube.

Note that there is a transformation

$$h\colon (x, y, z) \mapsto (x + \pi, y + \pi, z + \pi). \tag{3.10}$$

which preserves Δ but changes the sign of v.

The group G of rotations of the cube has 5 irreducible representations:

- The trivial 1-dimensional representation; denote it (1);
- The non-trivial 1-dimensional representation; denote it (-1); it is the permutation of the two inscribed tetrahedra;
- The 2-dimensional representation; denote it (2); it is the permutation of the three coordinate axes;
- The 3-dimensional representation; denote it (3);
- The twisted 3-dimensional representation; denote it (-3); we have $(-3) \simeq (3) \otimes (-1)$.

The group G acts also on the space of divergence-free vector fields on the torus.

The G-action commutes with the action of the operators Δ and $\{v, \cdot\}$, and hence with A_R. Therefore

1. G acts on the eigenspaces of Δ;
2. The operator A_R is the direct sum of the five operators each acting in the direct sum of several copies of the same irreducible representation of G;

3. The eigenvector "almost certainly" belongs to one of the five spaces described in item 2: The coincidence of eigenvalues of any two of the five operators described in item 2 is "improbable";
4. The operators A_{-R} and A_R are equivalent, and hence their eigenvalues coincide, so the characteristic equation does not vary under the change $R \longrightarrow -R$.

The property 4 follows from the fact that h sends A_R to A_{-R}.

3.4. Decomposition of representations into irreducibles

Consider the representation of the group G of rotations of the cube in the eigenspace of the Laplace operator on the torus, acting in the space of divergence-free vector fields with mean value 0. The eigenspaces consist of trigonometric polynomials with vector-valued coefficients, these polynomials being orthogonal to the wave vectors

$$\sum H_k e^{i(k,r)}, \quad (k, H_k) = 0, \tag{3.11}$$

with a fixed sum of the squares of components of the wave vector (recall that the eigenvalue of the Laplace operator is $-k^2$). The dimensions of these spaces are twice the number of the integer points in the respective spheres, e.g.,

$$\begin{aligned} &d(1) = 12,\, d(2) = 24,\, d(3) = 16,\, d(4) = 12,\\ &d(5) = 48,\, d(6) = 48,\, d(7) = 0, \quad d(14) = 96. \end{aligned} \tag{3.12}$$

The given in subsect. 3.3.1 expressions for the action of the generators g_3 and g_4 on the space of vector fields on the torus allow one to find characters of the representations of the group G in each eigenspace of the Laplace operator.

Knowing the characters we can decompose the representation into irreducibles.

The results of (considerably long) calculations are as follows.

We separate the integer points (i.e., the points with integer coordinates) into orbits of the group PS generated by coordinate permutations and changes of coordinates's signs. For each of the seven of such orbits indicated in the first column of the following table, the next columns give: The dimension of the corresponding eigenspace of the Laplace operator (i.e., the doubled number of the cardinality of the orbit), the remaining characters, then the multiplicities of the respective irreducibles, and lastly, the first three values of k^2 for which such an orbit can be encountered:

Orbit	characters					multiplicities of the irreps					the first three values of k^2		
$(2r, 0, 0)$	12	0	0	-4	0	0	0	0	2	2	4	16	36
$(2r+1, 0, 0)$	12	0	0	4	0	1	1	2	1	1	1	9	25
$(a, b, 0)$	48	0	0	0	0	2	2	4	6	6	5	10	13
$(a, a, 0)$	24	0	0	0	-4	0	2	2	4	2	2	8	18
(a, b, c)	96	0	0	0	0	4	4	8	12	12	14	21	26
(a, a, c)	48	0	0	0	0	2	2	4	6	6	6	27	36
(a, a, a)	16	-2	0	0	0	0	0	2	2	2	3	12	27

For an illustration, let us describe the representations in the eigenspaces of the Laplace operator with small eigenvalues:

$$\begin{aligned}
&k^2 = 1 : (1) \oplus 2(2) \oplus (3) \oplus (-3),\\
&k^2 = 2 : (2)(-1) \oplus 2(2) \oplus 4(3) \oplus 2(-3),\\
&k^2 = 3 : 2(2) \oplus 2(3) \oplus 2(-3),\\
&k^2 = 4 : 2(3) \oplus 2(-3),\\
&k^2 = 5 : 2(1) \oplus 2(-1) \oplus 4(2) \oplus 6(3) \oplus 6(-3),
\end{aligned} \tag{3.13}$$

The above tables imply the following

3.4.1. Theorem. *The decomposition of the G-action in the space of vector fields on the torus with a given k^2 into irreducibles is of the form*

$$\begin{aligned}
&A[(1) \oplus (-1) \oplus 2(2) \oplus 3(3) \oplus 3(-3)] \bigoplus B[(1) \oplus (-1) \oplus 2(2) \oplus 3(-3)] \bigoplus\\
&C[(3) \oplus (-3)] \bigoplus D[(-1) \oplus (2) \oplus (3) \oplus 2(-3)] \bigoplus E[(2) \oplus (3) \oplus (-3)],
\end{aligned} \tag{3.14}$$

where the coefficients are expressed in terms of the number N of integer points on the sphere as follows:

k^2	A	B	C	D	E
$\neq n^2,\ 2n^2,\ 3n^2$	$\frac{N}{12}$	0	0	0	0
$(2r+1)^2$	$\frac{N-6}{12}$	1	0	0	0
$4r^2$	$\frac{N-6}{12}$	0	2	0	0
$2n^2$	$\frac{N-12}{12}$	0	0	2	0
$3n^2$	$\frac{N-8}{12}$	0	0	0	2

(3.15)

3.4.1.1. Remark. A. B. Givental pointed out that this theorem can be obtained in a simpler way with the help of the Frobenius reciprocity applied to the pair (G, the stationary subgroup of the wave vector). For us only the above decomposition table are essential, so we will not dwell on this.

3.5. Symmetries of the growing mode

Let us compare the "wonderful" values of k^2 from § 3.2, property 3 of G from § 3.3, and the decomposition table (3.14). The following corollary is obvious.

3.5.1. Theorem. *The amplitudes of the harmonics of the mode to be found with a prescribed k^2 are zero if the representation to which the mode belongs does not enter in the decomposition of the space of vector fields with given k^2 into irreducibles.*

The "wonderful" values of k^2 are therefore certain analogs of fingerprints that allow one to find if not the mode itself then, at least, its symmetry.

The tables of §4 yield that for

$$k^2 = 3,\ \ 4,\ \ 12,\ \ 16,\ \ 48,\ \ 64,\ \ \dots \tag{3.16}$$

only the non-trivial 1-dimensional representation (and only this representation) does not enter the decomposition. Therefore, the mode belongs to the space of this representation. The fields of this space satisfy the condition

$$g_3 H = H, \quad g_4 H = -H. \tag{3.17}$$

This space decomposes into the direct sum of its intersections with the eigenspaces of the Laplace operator. The dimensions of these representations are, according to §4, as follows:

k^2	1	2	3	4	5	6	8	9	10	11	12	13	14	16
dim	1	2	0	0	2	2	2	3	2	2	0	2	4	0

(3.18)

For the "wonderful" values of k^2 of the form 4^a and $3 \cdot 4^a$, the dimension of the intersection is equal to 0. For the remaining values of k^2, the dimension is positive as follows from the following statements of the Number theory known already to Gauss (the author is thankful to J.-P. Serre for this information):

1. The square of each prime can be non-trivially represented as the sum of three perfect squares.
2. Three times square of each prime can be represented as the sum of distinct squares.

3.6. The algebra of even and odd fields

To find the fields that transform along the representation (-1), it is convenient to proceed as follows. We say that a given field H is

$$\begin{cases} \text{even} & \text{if } g_3 H = H = g_4 H; \\ \text{odd} & \text{if } g_3 H = H = -g_4 H. \end{cases} \tag{3.19}$$

The sums of even and odd fields are invariant with respect to the operators g_3 and g_4^2; they form a Lie algebra.

Our initial field

$$v = (\cos y + \sin z)\frac{\partial}{\partial x} + \dots, \tag{3.20}$$

hereafter $\dots$ stands for the cyclic permutation of the written terms relative the coordinates, is even. It is not difficult to verify that the field

$$H_1 = (\cos y - \sin z)\frac{\partial}{\partial x} + \dots \tag{3.21}$$

is odd. The repeated Poisson bracket of even and odd fields is even (odd) depending on the parity of the odd fields involved; in other words, the Lie algebra of the sums of even and odd fields is $\mathbb{Z}/2$-graded. In particular, all the fields

$$\{v, H_1\}, \quad \{v, \{v, H_1\}\}, \quad \dots \tag{3.22}$$

are odd.

If the field is even (odd), then its projection onto each eigenspace of the Laplace operator is also even (odd). Let $(H)_{k^2}$ denote the projection of the field H onto the subspace with the eigenvalue $-k^2$, i.e., the sum of harmonics of H with the wave vector whose squared length is equal to k^2.

The linear combinations of the Poisson brackets and projections allow one to obtain from v and H_1 [4)] plenty of odd fields (perhaps, all of them?).

The Lie algebra obtained is rather involved. Its beginning segment is as follows:

3.6.1. Theorem. *We have*

$$\begin{aligned} &\{v, H_1\} = -2H_2, \quad \{v, H_2\} = H_5, \quad (\{v, H_5\})_2 = -\frac{1}{2}H_2 + H_2', \\ &\{v, H_2'\} = \frac{1}{2}H_1 + \frac{1}{2}H_5', \quad (\{v, H_5'\})_2 = -H_2', \end{aligned} \tag{3.23}$$

where the explicit forms of the basis fields are

[4] Do not confuse with $(H)_1$!

$$\begin{aligned}
H_2 &= (\cos x\cos z + \sin x\sin y)\frac{\partial}{\partial x} + \dots,\\
H_2' &= \sin y\cos z\frac{\partial}{\partial x} + \dots,\\
H_5 &= (\cos x\sin 2y - \sin x\sin 2z - \frac{1}{2}\cos 2x\cos y - \frac{1}{2}\cos 2x\sin z - \\
&\quad \frac{1}{2}\cos 2y\sin z - \frac{1}{2}\cos y\sin 2z)\frac{\partial}{\partial x} + \dots,\\
H_5' &= (\cos y\cos 2z + \cos 2y\sin z)\frac{\partial}{\partial x} + \dots\ .
\end{aligned} \tag{3.24}$$

3.7. The representation of the quiver generated by bracketing with v

A vector field on the torus will be said to be *polynomial* if its components are trigonometric polynomials. A polynomial field is said to be *homogeneous* if it is an eigenvector of the Laplace operator, i.e., if for all its harmonics $\mathbf{H}e^{i(\mathbf{k},\mathbf{r})}$ the value of $\mathbf{k}^2$ is the same. We call this common value k^2 the *degree* of the homogeneous vector field.

The spaces of the even and odd polynomial fields are the direct sums of their homogeneous components. The operator $\{v,\cdot\}$ of Poisson bracketing with the field (3.2) intermixes the components in a certain not arbitrary way. Namely, we call the harmonics $e^{i(\mathbf{k},\mathbf{r})}$ and $e^{i(\mathbf{k'},\mathbf{r})}$ *neighboring* if the integer vectors $\mathbf{k}$ and $\mathbf{k'}$ differ only by 1 in one of the three components. (Thus, each harmonic has 6 neighbors.)

Now, let us form a graph whose nodes are the degrees of neighboring odd fields. We join the nodes k^2 and k'^2 by a segment if the nodes are the degrees of the neighboring harmonics.

Since $\deg v = 1$, the components of the Poisson bracket of v with a homogeneous field of degree k^2 can be non-zero only in the degrees neighboring to k^2 in the above described graph.

The beginning segment of the graph is, according to § 3.4, of the form

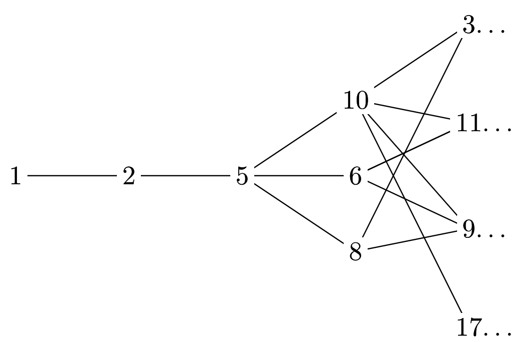

Fig. 1

3.7.1. Theorem. *Every even node is joined only with odd ones, and every odd node is joined only with even ones. No other segments, except those drawn, go out of the nodes* 1, 2, 5, 6, 8, 10.

In certain situations more convenient is the graph with a larger number of nodes that correspond not to the degrees k^2 but the orbits of the group PS of permutations of the coordinates of the wave vector and their sign changes.

The restriction A_i of the operator $\{v, \cdot\}$ onto the space F_i of homogeneous odd fields of degree i is represented as the finite sum of homogeneous summands

$$A_i = \bigoplus A_{i,j}, \text{ where } A_{i,j}\colon F_i \longrightarrow F_j \text{ is the homogeneous component of } A_i \tag{3.25}$$

and where j runs over all neighbors of i in the graph constructed.

If we replace each edge of our graph by a pair of arrows directed in opposite directions to each other, we get an (infinite) quiver. Above we have constructed a representation of this quiver: To the node i the space F_i corresponds and to each arrow there corresponds a linear operator $A_{i,j}\colon F_i \longrightarrow F_j$.

3.7.2. Theorem. *The dimensions of the first three spaces F_1, F_2 and F_5 are equal to* 1, 2 *and* 2, *respectively, with the bases H_1, H_2 and H_2', H_5 and H_5', see § 3.6.*

In these bases, the matrices of the operators $A_{i,j}$ are as follows:

$$(A_{1,2}) = \begin{pmatrix} -2 \\ 0 \end{pmatrix}, \quad (A_{2,1}) = \begin{pmatrix} 0, \frac{1}{2} \end{pmatrix}, \quad (A_{2,5}) = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, \quad (A_{5,2}) = \begin{pmatrix} -\frac{1}{2} & 0 \\ 1 & -1 \end{pmatrix}, \tag{3.26}$$

This is a reformulation of Theorem 3.6.1.

3.8. Galerkin 5-mode system

Ignoring in the Fourier series

$$H = (H)_1 + (H)_2 + (H)_5 + \dots \tag{3.27}$$

the terms of degree higher than 5 we get a Galerkin system of 5 linear equations for the components of the odd eigenvector

$$(H)_1 = a_1 H_1, \quad (H)_2 = a_2 H_2 + a_2' H_2', \quad (H)_5 = a_5 H_5 + a_5' H_5' \tag{3.28}$$

of the operator $A_R^{(5)} = p^{(5)}[R\{v, \cdot\} + \Delta]$, where $p^{(5)}$ is the projection onto $F_1 \oplus F_2 \oplus F_3$. Observe that, on the level of F_1 and F_2, the relations are the same as for the total (not Galerkin) system, i.e., the nodes 1 and 2 in the graph are only joined with the node 5.

3.8.1. Theorem. *The explicit form of the Galerkin system of the five odd modes is as follows:*

$$\begin{gathered}\left(\frac{R}{2}\right)a_2' = (\lambda+1)a_1, \quad -R\left(2a_1 + \frac{a_5}{2}\right) = (\lambda+2)a_2, \\ R(a_5 - a_5') = (\lambda+2)a_2', \quad Ra_2 = (\lambda+5)a_5, \quad \left(\frac{R}{2}\right)a_2' = (\lambda+5)a_5'.\end{gathered} \tag{3.29}$$

Proof of the formulas of § 3.7. Solving the system (3.29) we easily find the characteristic equation:

$$4(\lambda+1)(\lambda+2)^2(\lambda+5)^2 + 4(\lambda+1)(\lambda+2)(\lambda+5)R^2 + (5\lambda+21)R^4 = 0, \tag{3.30}$$

wherefrom we have

$$R^2 = \frac{-2(\lambda+2)(\lambda+5)[(\lambda+1) \pm 2\sqrt{-(\lambda+1)(\lambda+5)}]}{5\lambda+21}. \tag{3.31}$$

The graph if this function is depicted on the following Figure [5]:

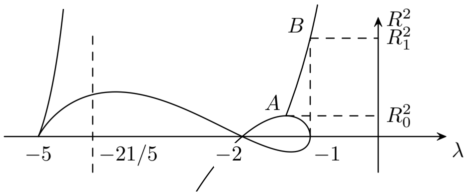

Fig. 2

If the viscosity is sufficiently high, we can ignore the higher harmonics. Therefore, for small values of the Reynolds number R we may use this graph to make predictions on the behavior of the increment of the odd mode of fastest growth (more precisely, of slowest decay). Comparison with numerical experiments for large Reynolds numbers shows a good agreement of the 5-mode Galerkin approximation under our study with the exact solution for $R^2 \ll 10$.

For $R = 0$, the spectrum is real ($\lambda = -1, -2, -2, -5, -5$). □

3.8.2. Theorem. *In the* 5*-mode Galerkin system, as* R *grows the two eigenvalue (*$\lambda = -1$ *and* $\lambda = -2$*) begin to move towards each other. At the critical value* $R_0 \approx 0.9324$ *the two real values coincide (*$\lambda \approx -1.356$*) and descend in the complex domain (point* A *on Fig. 2). Then the real part* γ *of the newborn pair of complex eigenvalues grows together with* R *(curve* AB *on Fig. 2).*

[5] Here the dotted line $x = -\dfrac{21}{5}$ is a vertical asymptote (to which the part of the graph under the λ-axis approaches rather slowly).

Around point $R_1 \approx 2.032$ *(point* B*) the value of* γ *reaches* -1 *and the odd mode outruns the even mode of* v *for which* $\gamma = -1$ *for all Reynolds numbers.*

Starting from this place the mode under study becomes the leading one (has the least damping). As R *grows further,* γ *also grows (asymptotically as* $\dfrac{R\sqrt{1+\sqrt{5}}}{2} \approx 0.9R$*) passing through* 0 *at* $R \approx 4.32$*.*

Although it is clear that for so big values of R it is inappropriate to ignore harmonics of degree greater than 5, the behavior of $\gamma(R)$ for $R < 5$ is satisfactory described by the 5-mode approximation; this approximation reveals the origin of the growing complex mode from the two colliding pair of real decaying modes of degrees 1 and 2 (in the exact system we have $R_0 \approx 0.96$, $\lambda_0 \approx -1.32$, $R_1 \approx 2.286$).

3.9. The diagram technique

Our calculation give, actually, more than Galerkin approximations: We also get the exact expressions for the Taylor (or Puiseux [6)]) series expansions of the eigenvalues of the total (no Galerkin) system (3.1) in powers of $R = \dfrac{1}{\varepsilon}$.

The answer is given in terms of the above constructed representation of the quiver: The terms of degree n in the Taylor series of the root born of $-k^2$ correspond to the loops of length n and the source at the node k^2.

For simplicity, we begin with a simple (of multiplicity 1) root $\alpha_0 = -1$.

3.9.1. Theorem ([3]). *The first (apart from* $\alpha_0 = -1$*) non-zero coefficient of* R^n *in the Taylor series expansion of the eigenvalue*

$$\lambda = \alpha_0 + \alpha_1 R + \alpha_2 R^2 + \ldots \tag{3.32}$$

of the Laplace operator $\Delta + R\{v, \cdot\}$ *is given by the sum*

$$\alpha_n = \sum \frac{A_{i_{n-1},i_0} \ldots A_{i_1,i_2} A_{i_0,i_1}}{(\alpha_0 - \lambda_{i_{n-1}}) \ldots (\alpha_0 - \lambda_{i_2})(\alpha_0 - \lambda_{i_1})}, \tag{3.33}$$

which runs over all loops $i_0 \longrightarrow i_1 \longrightarrow \ldots \longrightarrow i_{n-1} \longrightarrow i_0$ *of length* n *and with the source and target in the node* i_0*.*

In our case, $i_0 = 1$, the operators $A_{i,j}$ are introduced in § 3.7.

Proof of relation (3.33) is not difficult; we will only show how to use it.

1. Eq. (3.33) implies that α_n with any odd subscript vanish since the length of any loop on our quiver is even.
2. The only loop of length 2 with source at node 1 is $1 \longrightarrow 2 \longrightarrow 1$ Since, due to formulas of § 3.7 we have $A_{2,1} A_{1,2} = 0$, it follows that $\alpha_2 = 0$.

[6] http://en.wikipedia.org/wiki/Puiseux_expansion.

3. The only loop of length 4 with source at node 1 is $1 \longrightarrow 2 \longrightarrow 5 \longrightarrow 2 \longrightarrow 1$. We get
$$\alpha_4 = \frac{\left(0, \frac{1}{2}\right)\begin{pmatrix} -\frac{1}{2} & 0 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & -\frac{1}{2} \end{pmatrix}\begin{pmatrix} -2 \\ 0 \end{pmatrix}}{1 \cdot 4 \cdot 1} = -\frac{1}{4}. \tag{3.34}$$
Therefore, for the total infinite dimensional system, we have
$$\lambda = -1 - \frac{R^4}{4} + o(R^4),$$
the same as for the Galerkin approximation in § 3.8.
4. We also conclude that only a bounded part of the quiver affects α_n and can estimate which part.

In the general case where the root α_0 is a multiple one, the relation (3.33) determines not a number but a linear operator $\alpha_n \colon F_{i_0}$.

In the calculation of the perturbation of the multiple eigenvalue one can replace the infinite dimensional space of fields by a finite dimensional space F_{i_0} and the total operator by the matrix
$$\alpha_0 E + R\alpha_1 + R^2\alpha_2 + \dots \tag{3.35}$$
For example, exactly two loops of length 2 begin in node 2, namely $2 \longrightarrow 1 \longrightarrow 2$ and $2 \longrightarrow 5 \longrightarrow 2$. The corresponding summands are
$$\frac{\begin{pmatrix} -2 \\ 0 \end{pmatrix}\left(0, \frac{1}{2}\right)}{-1} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \frac{\begin{pmatrix} -\frac{1}{2} & 0 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}}{3} = \begin{pmatrix} -\frac{1}{6} & 0 \\ \frac{1}{3} & -\frac{1}{6} \end{pmatrix}. \tag{3.36}$$
Thus,
$$\alpha_0 E + R^2\alpha_2 = \begin{pmatrix} -2 - \frac{R^2}{6} & R^2 \\ \frac{R^2}{3} & -2 - \frac{R^2}{6} \end{pmatrix} \tag{3.37}$$
and, to find λ with accuracy $O(R^2)$, we have to know α_4 as well. For the Galerkin system, we have
$$\alpha_4 = 0 \text{ and } \lambda = -2 + R^2\Big(-\frac{1}{6} \mp \sqrt{\frac{1}{3}}\Big) + O(R^4) \tag{3.38}$$
in good agreement with formulas of § 3.8.

References

[1] Henon M. Sur la topologie de lignes de courant dans un cas particuler. C.R. Acad. Sci., Paris, v. 262, 1966, 312–314.

[2] Arnol'd V.I. Several remarks on anti-dynamo theorem. Moscow Univ. Mathematics Bull., ser. matem., meh., 1982, no. 6. 50–57.

[3] Arnol'd V.I. Remarks on perturbatioin theory for Mathieu-type problems. Russian Math. Surveys, t.38, 1983, no. 4, 189–203.

[4] Arnol'd V.I., Korkina E.I. The growth of magnetic field in three-dimensional stationary low of incompressible liquid. Moscow Univ. Mathematics Bull., ser. matem., meh., 1983, no. 3, 43–36.

Index