CS728

An Introduction to Geometric Complexity Theory

Milind Sohoni

July 2023

Module 1: Group Actions and orbits.

About the course.

1. What are objectives of the course. Prerequisites: Groups, Linear Algebra. Commutative Algebra. Basics of Ideals and Varieties. Artin -Algebra covers all. Lectures and conduct. Taking notes. Exams and quizzes. Reading work.

2. Algorithms and computational complexity. The function as a subroutine. The reduction in computer science. Examples - The flow problem, LP, Hamiltonian circuit. The classes P and NP.

3. The reduction by algebraic substitutions. The singular substitution and the problem of orbit closure.

4. Formula size and Valiant's result of the universality of the determinant, i.e., Combinatorial ⇒ Algebraic. What is GCT.

Simple examples of the orbit closure problem.

1. Examples of configurations and actions. 15-puzzle, the elevator, the milk distribution problem, n-ball problems. Some are group problems, most are not. Invertibility and universal applicability.

2. The rubik cube -actions and configurations and a typical question. The necklace problem under the rotation and dihedral group. The definition of the orbit.

Review of groups

1. The basic axioms. The basic examples including $\mathbb{Z}_n, S_n, D_n$ and GL_n, O_n. Homomorpshisms.

2. The group of substitutions $x \to ax + b$. Its subgroup $x \to x + b$.

3. Subgroups generated by elements - the finite and the infinite case. Groups as functions on sets.

Group actions and Orbits

1. Group acting on sets - definition $\rho : G \to Bij(S)$. Examples - D_n and S_n. Definition of the orbit. The equivalence $\sim$ and the partition of S into orbits.

2. Associated actions on $S \times S, 2^S$. The action of S_3 on $S \times S \times S$. Diagonal action on S_1 and S_2. Whyis it important.

3. Stabilizers and examples -D_n, S_n and GL_n.

4. What is the orbit and how does it connect with cosets. The symmetries of the cube and its accounting in different ways.

Orbits and Vector Spaces

1. The vector space $\mathbb{C} \cdot S$ and the action of G as permutations $\rho(g)$. Necklaces as vectors. Orbits for D_n and S_n.

2. Motivating invariants. How do I check if the two necklaces are the "same"? How many distinct necklaces exist? How do I quickly check if two necklaces are the "same"? Functions on $\mathbb{C} \cdot S$ and the need for invariants. $\mathbb{C} \cdot S$ and its dual $X = \{x_1, \ldots, x_n\}$. The combined action on the dual and the primal. Polynomials as functions on necklaces and the action of G on polynomials.

3. The Fourier functions $f_k(x) = \sum_{i=0}^{n-1} x_i \alpha^{ik}$ and their properties. The FFT and proof that the Fourier functions separate orbits for Z_n.

4. The action of G on $\mathbb{C}[S]$ the ring of polynomials. The symmetric group on subsets of a set. And on vectors. The symmetric polynomials. Proof that the symmetric polynomials separate orbits.

5. Polynomials in one variable with the substitution $x \to ax + b$. Why is this a subgroup of GL_2? How does this act on the vector space of polynomials? Writing matrices for this action. Invariants as elimination of variables.

6. The Galilean group on colored points in $\mathbb{R}^2$. Writing matrices for this action. The distance invariants and the determinant for orientation. Exact match and almost exact match. Nearby orbits and the importance of invariants. The question of 2D photographs of 3D objects.

Functions and Invariants

1. The general group action $\rho : G \to GL(V)$, where V is an n-dimensional vector space with coordinate functions $X = \{x_1, \ldots, x_n\}$. Examples from last class. New examples: The adjoint action of GL_2 on 2×2-matrices.

2. Getting the action on $x_1, \ldots, x_n$. The action on a space and its dual. Definition of the group action on functions and its associativity. The action so that the diagonal action of G on $\mathbb{C}[V] \times V$ such that $g(f, v)$ is invariant, in other words $g(f(v)) = f^g(g \cdot v)$. The associativity of $G \times \mathbb{C}[V] \to \mathbb{C}[V]$.

3. The notion of a general invariant $f^g = \phi(g) \cdot f$. Ring of invariants. What does it signify?

4. The big question: Can orbits be separated by invariants? The finite group case. The averaging operation. The separation of points sets by Lagarange interpolation.

Summary

- Group actions. Orbits and stabilizers. Quotients and Invariants.

- The Z_n fourier invariants and the usual invariants. Is it possible to determine the fourier invariants from the usual?

Module 2: Vector Spaces and Maps

Basic definitions

1. The field $\mathbb{C}, \mathbb{R}$ etc. The basic definitions - linear independence and subspaces. The existence of a basis. Dimension. Isomorphism and Homomorphism. Kernel, Image and quotient. The dimension theorem.

2. Choice of basis and the matrix representation. Change of basis and the invertible matrix. The matrix, its row space and column space. The nullspace of a matrix and the rank-nullity theorem. The row echelon form and the expression of a matrix $M = TR$, where R is the row echelon form. The structure of T and R. The equality of dimensions of the row space and column space.

3. Examples: $\mathbb{R}^n$, polynomials - various bases and their importance. Matrices and subspaces with special properties. Lie algebras.

4. Tangent spaces: Two definitions - small movements and derivations (but what are functions on solutions of equations?). Computation of tangent spaces from polynomial equations. The gradient nullity form.

   The dimension of tangent spaces and their significance. Non-singular and singular points as examples. The sphere, SL_n, the orthogonal group, the elliptic curve and the singular cubic.

   Maps between manifolds and the tangent map. The parametrization of the surface of the sphere. The map from rank 1 2×3-matrices to the 3 determinants.

5. The membership problem. Given a subspace $W \subseteq V$ in terms of a basis of V, to answer : Is $w \in W$?. Simplification of a basis of a subspace in terms of another. The LU decomposition. Nested subspaces and dimensions. Classification of subspaces.

The Linear Map

1. $V = M^{n \times r}$ as a left GL_n-module. The orbits and orbit closures. The REF as a section. SL_n and GL_n-orbits. Stabilizer and their dimensions.

2. The basic orbit structure of an SL_n-module. The stable and semi-stable points and the hull-cone. The ring of invariants and what they can separate. Example: binary forms. Writing matrices for binary forms and Sym_n-modules. Writing matrices for the adjoint action.

3. How things change with the GL_n-orbits. The $M^{n\times r}$ again. GL_n-orbit closure and modules as separators.

4. The determinant, basic operations and invariance and its universality. The product law. The proof of invertibility $\Leftrightarrow$ non-singularity. The rank condition for general matrices. The quadratic relations. The matrices for the $\wedge_n$-modules.

5. The orbit structure of $Hom(V, W)$. Analysis of a single linear transformation ϕ. Change of basis and the conjugation action. The Cayley Hamilton theorem and the Jordan representation. The rank closure conditions and filtration of Jordon blocks. Orbits, Hilbert'ss 1-PS criteria (without proof) and Closures. Invariants. Stabilizers and their dimension. The GIT of the conjugate action.

Module 3: Geometric Invariant Theory and Geometric Complexity Theory

The Gordan-Hilbert historical problem

1. The homogeneous action of $G \subseteq GL(X)$ on a general space V and on $\mathbb{C}[V]$. The basic questions - What is the space of orbits? Does the ring of invariants separate orbits? Is the ring of invariants finitely generated. The importance of finite generation.

2. Hilbert's solution for SL_n. The null-cone and the extent of non-separation. The definition of unstable points and Hilbert's 1-PS solution. Stable points and their extent. The GIT structure the of $Sym^k(\mathbb{C}^n)$ and $End(V)$ and the core invariants.

3. Later developments. The semi-stable and the open stable points. The set $null(z)$ of all points which close onto z. The Mumford-Kempf criteria, optimal 1-PS and stabilizers.

Rings and ideals

1. The Ring $\mathbb{C}[V]$. Algebraic sets, ideals and varieties. Correspondence between radical ideals and varieties. maximal ideals. The coordinate ring and $\mathbb{C}[V]/I$. Finite generation and Hilbert basis theorem.

2. The resultant in $R[x]$ and its cases. An example. The easier version of Hilbert's Nullstelensatz. Its consequences. The harder version. The orbit is an almost algebraic set.

3. Dimension and the Jacobian condition. Singularity. The dimension of orbits and the complementarity of stabilizers. Examples.

4. Group actions and the map $\rho^* : \mathbb{C}[V] \to \mathbb{C}[V] \otimes \mathbb{C}[G]$. Its consequences - homogeneity and finite dimensionality of modules.

5. Lie alegbras, their definitions and examples. Lie algebra actions. The computation of $\rho_1 : \mathcal{G} \times V \to V$. Stabilizer conditions. Examples.

Groups and reductivity

1. Algebraic groups as subgroups of $GL(X)$. Computation of the coordinate rings. Examples of GL_n, SL_n, O_n and computation of the ρ^*-map for Sym^d and $\wedge^d$.

2. Definition of reductivity of groups and irreducible modules. Basic categorical properties.

3. The action of reductive groups on $\mathbb{C}[V]$ and the Π^G Reynolds operator.

Geometric Invariant Theory and Geometric Complexity Theory

1. GIT: the fundamental theorems. Finite generation and separation of closed sets. The Nagata equivalence relation. Closed orbits and The statement of the Hilbert-Mumford 1-PS condition. Examples.

2. GCT I: The affine pull-back problem of g from f. The homogenization and the 1-PS formulation. The orbit closure and the witness formulation. The Peter-Weyl condition and the Obstruction Cojecture.

3. GCT II: The action of λ on V and $\mathcal{G}$. The basic equation:
$$\lambda(t)y = t^d y_d + t^e y_e + \ldots + t^D y_D$$
The weight spaces and leading terms. The first theorem: leading term Lie algebra $\hat{\mathcal{H}}$ of $\mathcal{H}$ and module $\hat{N}$ of N. The second theorem: $\lim_{t\to} t^{-d} y(t) = z$ and its implication $\widehat{\mathcal{G}_y} \to \mathcal{H}_{\overline{y_e}} \subseteq \mathcal{H}$ (where $\mathcal{H} = \mathcal{G}_z$).

4. Alignment and the dichotomy result. Consequences of alignment - rectangular decompositon. The absence of alignment and intermediate G-varieties.

Notes - Lecture 1

02-Aug-2023

Scribe: Vaibhav Krishan

The overall goal of the current offering of this course is to introduce the participants to Geometric Complexity Theory, a framework designed by Mulmuley and Sohini [MS01] where they put forth an approach for resolving some conjectures in computer science. The focus will be on applications to complexity theory, an area of computer science concerned with understanding which functions are "easy" and which are "hard" to compute, the precise notions will be defined later.

A basic familiarity with Linear Algebra, Basic Algebra is assumed. Familiarity with Commutative Algebra, Theory of Group Actions, and Group Representations will certainly be of help. Algebra by Artin [Art11] will be the standard reference for the basics. Along with this, a basic familiarity with asymptotic notation, i.e. $O(), \Omega(), \Theta(), o(), \omega()$, is also assumed.

1 Introduction

We start by describing a central theme in complexity theory i.e. understanding functions that can be computed easily versus functions such their value can be verified easily. Let us understand what it means to compute or verify a function.

A program or an algorithm for a function is a series of steps that, when given an input of the function, leads to the value of the function on that input. The number of steps in an algorithm is called the running time of the algorithm. An important notion is the size of the input, with respect to which the asymptotic growth of the running time of an algorithm is considered in complexity theory.

A program or an algorithm for verifying a function is an algorithm which, when given as input an input to the function along with a series of steps that are supposed to lead to the value of the function aka a "proof", can verify the value of the function. Note that the size of the input does not include the size of the "proof".

Example 1.1. *Consider the function f which, on begin input an array $A = \{a_1, \dots, a_n\}$ of n elements, outputs the elements in sorted order i.e. it outputs $\{a_{i_1}, \dots, a_{i_n}\}$ such that $a_{i_1} \leq \dots \leq a_{i_n}$.*

Various algorithms are known for computing this function. Bubble sort can compute this function in $O(n^2)$ steps while merge sort can compute this function in $O(n \log(n))$ steps. It is also known to be the best asymptotic running time of any algorithm computing this function.

Compare this with the following algorithm for verifying this function. Let the algorithm be given an input array and the supposedly sorted array. The algorithm will first check that the frequency of each element in both the arrays match. Then it will simply iterate over the latter array and check if it is sorted. If yes, it will output that array, and halt otherwise. This will take $O(n)$ steps.

1.1 P vs NP

Here, we state a conjecture about two sets of functions.

Note that, as any real computer can only work on inputs that have a finite size and output a finite size value, both of which can be represented as a Boolean valued string, we can assume without loss of generality that functions are Boolean valued over Boolean inputs. It is indeed common in complexity theory to study Boolean valued functions. Hence functions will be Boolean valued unless specified otherwise.

Definition 1.2. *Let there be a function $f : \{0,1\}^* \to \{0,1\}$ where $\{0,1\}^*$ denotes the set of all finite sets with $0, 1$ as entries.*

An algorithm $A : \{0,1\}^* \to \{0,1\}$ is a series of steps that computes *f if and only if $A(x) = f(x)$ for all $x \in \{0,1\}^*$.*

An algorithm $V : \{0,1\}^* \times \{0,1\}^* \to \{0,1\}$ verifies *f if and only if for each $x : \{0,1\}^*$, there exists a $y \in \{0,1\}^*$, such that $V(x,y) = 1$ if and only if $f(x) = 1$. The following logical formula is an equivalent restatement of the above:*

$$\forall x \in \{0,1\}^*, ((\exists y \in \{0,1\}^*, V(x,y) = 1) \iff f(x) = 1)$$

Here y is called the proof/certificate for f.

The asymptotic rate of growth of running time of an algorithm, denoted by $T(A,n)$, is called the complexity of the algorithm. We will often omit A when it is clear from the context. The complexity of a function, denoted by $T(f,n)$, is the slowest growing complexity of an algorithm computing it. The non-deterministic time complexity of a function, denoted by $NT(f,n)$, is similarly the slowest growing complexity of an algorithm verifying it. We will often omit f if it is clear from the context.

It should be obvious from the definition that $NT(f,n) \leq T(f,n)$ as any algorithm for verifying the function can simply ignore the proof and just run the algorithm to compute f.

A long standing line of inquiry is to find functions such that their non-deterministic time complexity is much lower than their time complexity. We define two sets and state a conjecture about their relationship.

Definition 1.3. *A function $f : \{0,1\}^* \to \{0,1\}$ belongs to the set* P *if $T(f,n) = n^{O(1)}$. It belongs to the set* NP *if $NT(f,n) = n^{O(1)}$.*

It should be immediately clear that $\mathrm{P} \subseteq \mathrm{NP}$. A long standing conjecture is that $P \subsetneq \mathrm{NP}$.

Conjecture 1.4. $\mathrm{P} \subsetneq \mathrm{NP}$.

An example of functions that are in NP but believed to not be in P are: Hamiltonian cycles in a graph i.e. is there a cycle in a graph that visits each vertex exactly once, subset sum problem i.e. given a collections of integers is there a subset that sums up to a certain given value etc.

Several approaches have been considered in the past for resolving this conjecture and have lead to beautiful results, see [Sha92] for example. But barriers have been found for many of these approaches, proving that those techniques cannot work for proving the aforementioned conjecture, look at [BGS75, RR94, AW09]. Hence, there was a need for approaches for which these barriers do not apply. Geometric complexity theory offers one such approach for which it has not been proved that current known barriers apply.

1.2 Reductions

We will study an important notion in complexity theory, which has a deep connection with geometric complexity theory, called reductions. Simply, a reduction is using a black-box computing a function, aka an oracle, to compute another function. The precise definition follows.

Definition 1.5 (Reduction)**.** *Let there be an algorithm A computing a function $f : \{0,1\}^* \to \{0,1\}$. A function g is said to be reducible to f, denoted by $g \prec f$, if there is an algorithm B that on input x is allowed to use A to compute f on some inputs $y_1, \ldots, y_k$, called* querying f*, and outputs $g(x)$.*

g is polynomial time reducible to f, denoted by $g \prec_p f$, if the algorithm B takes $n^{O(1)}$ time to compute $g(x)$ where $|x| = n$. Note that each query counts only as one step, the time taken by A is not taken into account. Also, for B to take polynomial amount of time, each y_i must have polynomial size and the number of queries must be polynomial as well.

Example 1.6. *A trivial example of a reduction is deciding if a number is prime using its prime factorisation.*

A slightly non-trivial example is as follows. A matrix M of size $n \times n$ is said to be sorted if $M_{i,j} \leq M_{i+1,j}$ and $M_{i,j} \leq Mi, j+1$ for all $1 \leq i, j \leq n$. Sorting a matrix is reducible to sorting an array as follows. Collect all the entries and sort them as an array. Use this to fill the matrix back in a straightforward way.

Some more examples are Hamiltonian cycle being reducible to subset sum and vice-versa. Many such functions are also reducible to integer linear programming.

Finding a reduction amounts to constructing an algorithm while proving that no reduction exists must somehow exhaust all possibilities. This makes it seemingly harder to prove non-existence of reductions as compared to finding them.

Remark 1.7. *An easy to compute function f generally makes it harder to find a reduction from other functions. The idea is that because the algorithm for f is not "doing much", it does not lend enough power as an oracle to compute other functions.*

We will see later how reductions play an important role in geometric complexity theory.

2 Algebraic Problems

Geometric complexity theory studies the connections of the combinatorial objects defined above with algebra and geometry. In a way, it tries to reduce combinatorial objects to algebraic functions. We look at a beautiful result by Valiant [Val79] that shows how one algebraic function is able to capture the power of a significant class of combinatorial objects. We first define a few algebraic functions.

Definition 2.1 (Permanent and Determinant)**.** *Let M be a matrix of size $n \times n$ with $x_{i,j}$ as formal variables being its entries. Its permanent and determinant, denoted by $perm(M), det(M)$, are defined as follows:*

$$perm(M) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} x_{i,\sigma(i)}$$

$$Det(M) = \sum_{\sigma \in S_n} sign(\sigma) \prod_{i=1}^{n} x_{i,\sigma(i)}$$

where S_n is the set of all permutations on n variables and sign(σ) is defined as follows. Let the set over which the permutation applies be $\{1, \ldots, n\}$. A pair i, j with $1 \leq i < j \leq n$ is called an inversion in σ if $\sigma(i) > \sigma(j)$. sign(σ) is -1 if there are odd number of inversions in σ and 1 otherwise.

The combinatorial objects under consideration are called formulas, defined below.

Definition 2.2. *A formula in the ring of polynomials $R = \mathbb{F}[X_1, \ldots, X_n]$ over a field $\mathbb{F}$ is defined recursively as follows.*

- *A field value $c \in \mathbb{F}$ is a formula.*
- *An indeterminate X_i is a formula.*
- *$f_1 + f_2$ and $f_1 \times f_2$ are formulas where $+, \times$ are rings operations in R.*

The size of a formula is the number of ring operations it uses.
Let there be a polynomial $P \in R$. A formula calculates P in an obvious fashion.

Valiant proved that all formulas, that are combinatorial objects, can be reduced to the determinant function, an algebraic function.

Theorem 2.3 ([Val79]). *Let there be a formula of size m computing a polynomial P over n variables. Then there is a matrix M of size $(m + 2) \times (m + 2)$, with linear functions of the form $l(x) = \sum_i a_i x_i + b$ for some $a_1, \ldots, a_n, b$ as its entries, such that $Det(M) = P$.*

This exemplifies the idea that "clever" combinatorial constructions can be captured efficiently by algebraic structures. This inspires exploring these connections deeper, which we will delve into next. We start with some preliminaries.

2.1 Transition Systems

Let V be a set of vertices and let G be a set of functions $g : V \to V$. A transition system T is a subset $T \subseteq V \times V$ where an element $g \in G$ acts on a vertex v and transitions to $g(v)$. Some examples are as follows.

Example 2.4.

1. *Consider all the configurations of a 15 tiles puzzle. A square is partitioned into 16 tiles of equal size, 15 with numbers on them and a tile with no number. Allowed movements are moving the empty tile left, right, up and down. The objective is to reach the configuration where all the numbered tiles are arrange in a sorted fashion and the empty tile in the bottom right corner.*
2. *Another example is the Rubik's cube, partitioned into cubes of equal sizes. Outside face of each smaller cube has one of six colors. Allowed movements are rotating a face. The goal is reach a configuration where each face has cubes of the same color on it.*

Elements $g \in G$ may not act on all elements of V. Consider example 1 in which the right and down movements do not act on the state where the empty tile is on the bottom right corner. While all movements in example 2 work on all states. A group transition system is a transition system if the action is applicable everywhere, invertible and associative. G is called a group action over V. Formally, V, T, G is a group transition system with an operation $\cdot : G \times G \to G$ if:

1. Every $g \in G$ acts on every $v \in V$.
2. There exists an element $e \in G$ such that $e \cdot g = g \cdot e$ for all $g \in G$.
3. $g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3$ for all $g_1, g_2, g_3 \in G$.
4. For all $g \in G$ there exists a $g^{-1} \in G$ such that $g \cdot g^{-1} = g^{-1} \cdot g$.

Consider the following group transition system. Let V be the collection of evenly spaced n beads of k colors. The set of actions allowed is to rotate the necklace around the axis perpendicular to its plane and flipping the necklace, i.e. rotating around an axis in the plane, such that the beads in the new and old configuration overlap.

The group action of these transitions is called the Dihedral group, denoted by either D_{2n} or D_n. We will fix D_n as the convention in these notes.

This group will be discussed in the next lecture.

Notes - Lecture 2

Scribe: Roshan Raj

3 Group Action and Group Homomorphism

In the last lecture, we saw the definition of a group. Now, we will see some examples.

3.1 Examples of Group

Example 3.1. $(\mathbb{Z}_n, +, 0)$ *Here n is some positive integer, and $\mathbb{Z}_n$ is the set of integers from 0 to $n-1$. For two elements $a, b \in \mathbb{Z}_n$, $a + b = (a + b) \mod n$. On RHS, the '+' refers to the regular addition operation for integers. 0 is the identity element, and for any element $a \in \mathbb{Z}_n \setminus 0$, $n - a$ is its inverse.*

Example 3.2. $(\mathbb{Z}_n^*, \times, 1)$ *Here n is some positive integer, and $\mathbb{Z}_n^*$ is the set of integers between 1 and $n-1$ which has gcd 1 with n. For two elements $a, b \in \mathbb{Z}_n$, $a \times b = (a \times b) \mod n$. The '×' on RHS refers to the regular integer multiplication operation. 1 is the identity element. For any element $a \in \mathbb{Z}_n^*$ since $gcd(a, n) = 1$, there exists integers α, β such that $\alpha a + \beta n = 1$. It is not difficult to see that $(\alpha \mod n)$ is an element of the group $\mathbb{Z}_n^*$ and is the inverse of a.*

Both of these groups are also commutative i.e. for any two elements a, b in the group, $a.b = b.a$ where '.' is the group operation. Following is an example of a group which is not commutative.

Example 3.3. $(GL_n(\mathbb{F}))$ *This group is known as the general linear group of degree n. Here, $\mathbb{F}$ is some 'field.' Right now, you can assume it to be the set of all real numbers or rational numbers or complex numbers. It consists of all $n \times n$ invertible matrices with the identity matrix as the identity element. The group operation is the operation of ordinary matrix multiplication. This forms a group because the product of two invertible matrices is again invertible, and the inverse of an invertible matrix is invertible.*

Now we return to the dihedral group that we saw in the first lecture.

Dihedral Group

The dihedral group denoted by D_n is the group of symmetries of a regular n-gon. It consists of $2n$ elements. Suppose, we label the vertices of the regular polygon by integers 0 to n-1 in the clockwise direction. Each element of the group corresponds to some configuration of the polygon. All the elements of D_n can be defined using just 2 elements i.e. σ, τ. You can think of identity element e as the initial configuration of the polygon. Now, if we rotate the n-gon clockwise around the axis perpendicular to the plane passing through its center by $2\pi/n$, we get a new configuration corresponding to element σ. In this new configuration, vertex 0 is at the place of vertex 1 in the initial configuration; vertex 1 is at the place of vertex 2, and so on. The configuration obtained by performing this operation $i \leq n$ times corresponds to the element σ^i. If we perform this operation n times on the n-gon, we return to the initial configuration. Hence, $\sigma^n = e$. The configuration obtained by flipping the polygon corresponds to the element τ. In this new configuration, vertex i is at the place of vertex $(n - i) \mod n$ in the initial configuration. If we flip the polygon twice, we return to the original configuration. Hence, $\tau^2 = e$. The remaining elements of the group are denoted by $\tau\sigma^i$ for $0 < i < n$. $\tau\sigma^i$ corresponds to the configuration obtained by first flipping the n-gon and then rotating it i times by angle $2\pi/n$ in the clockwise direction.

The elements of D_n are $\{\sigma^0, \sigma^1, \ldots, \sigma^{n-1}, \tau\sigma^0, \tau\sigma^1, \ldots \tau\sigma^{n-1}\}$. Let '·' denote the group operation. Then, the group operations between the elements of D_n are defined in the following way. You can verify that these operations follow from the above discussion.

$$\begin{aligned}
\sigma^i \cdot \sigma^j &= \sigma^{(i+j)\%n} \\
\tau\sigma^i \cdot \tau\sigma^j &= \sigma^{((n-1)i+j)\%n} \\
\sigma^i \cdot \tau\sigma^j &= \tau\sigma^{((n-1)i+j)\%n} \\
\tau\sigma^i \cdot \sigma^j &= \tau\sigma^{(i+j)\%n}
\end{aligned}$$

3.2 Subgroups

Given a group G under a binary operation $*$, a subset H of G is called a subgroup of G if H also forms a group under the operation $*$. More precisely, H is a subgroup of G if for any two elements $a, b \in H$, $a * b \in H$. This is denoted by $H \leq G$. Let e_G and e_H denoted the identity elements of G and H, respectively. Then, $e_G = e_H$.

For any element $s \in G$, let $s^1 = s$ and $s^i = s^{i-1} * s$ for $i > 1$. For any element s of G, $\langle s \rangle$ denotes the subgroup $\{s, s^2, \ldots, s^k\}$ where k is the smallest positive integer such that $s^k = e$ (the identity element). It is also called the subgroup generated by s. It is easy to verify that $\langle s \rangle$ satisfy all the axioms of a group. Following are some examples.

Example 3.4. *For the group $\mathbb{Z}_8$, the subgroup generated by 2, $< 2 >= \{0, 2, 4, 6\}$. Similarly, the subgroup generated by 3, $< 3 >= \mathbb{Z}_8$.*

Example 3.5. *For the Dihedral group D_n, the subgroup generated by σ, $\langle\sigma\rangle = \{\sigma^0, \sigma^1, \dots \sigma^{n-1}\}$.*

3.3 Group Homomorphism

Let G and H be two groups with '$*$' and '$\cdot$' as the group operations of G and H, respectively. Then, a mapping from elements of G to elements of H, denoted by $\phi : G \to H$, is called a group homomorphism if ϕ satisfies the following property:

$$\phi(a * b) = \phi(a) \cdot \phi(b) \quad \text{for any two elements } a, b \text{ of } G$$

Let $\phi : G \to H$ be a group homomorphism. Let e_G and e_H be the identity elements of G and H, respectively. Then, for any group element a of G, $\phi(a) = \phi(a * e_G) = \phi(a) \cdot \phi(e_G)$. This implies $\phi(e_G) = e_H$. Let's see one example.

Example 3.6. *Let ϕ be a mapping from D_n to $\mathbb{Z}_2$ that is defined as follows:*

$$\phi(\sigma^i) = 0 \quad \textit{and } \phi(\tau\sigma^j) = 1.$$

It is easy to verify that ϕ mentioned above is a group homomorphism. For a group homomorphism $\phi : G \to H$,

$$Kernel(\phi) = \{g \in G \mid \phi(g) = e_H\} \text{ and } Image(\phi) = \{h \in H \mid \exists g \in G, \phi(g) = h\}.$$

For any two elements $a, b \in Kernel(\phi)$, $\phi(a * b) = \phi(a).\phi(b) = e_H.e_H = e_H$. Hence, $Kernel(\phi) \leq G$.

3.4 Group Action

In the last lecture, we discussed transition systems. Let's see one more example. Let G be the set of all pairs (a, b) such that $a, b \in \mathbb{R}$ and $a \neq 0$. For $(a, b) \in G$, the transition $\psi : \mathbb{R} \to \mathbb{R}$ is defined as $\psi(x) = ax + b \; \forall x \in \mathbb{R}$. Suppose we compose transition (a, b) with (a', b'). When we first apply transition (a', b'), x is mapped to $a'x + b'$. Then on applying transition (a, b) on $a'x + b'$, we get $aa'x + ab' + b$. Finally, x is mapped to $aa'x + ab' + b$. Based on this, we can define G as a group with the group operation ('$*$') defined as follows. For $(a, b), (a', b') \in G$, $(a, b) * (a', b') = (aa', ab' + b)$. Note that, $(1, 0)$ is the identity element of G and $(1/a, -b/a)$ is inverse of (a, b).

Definition 3.7 (Group Action)**.** *If G is a group with identity element e and group operation '$*$', and S is a set, then a group action ϕ of G on S is a function $\rho : G \times S \to S$ that satisfies the following two axioms*

1. *$\phi(e, s) = s$, $\forall s \in S$*
2. *$\phi(g_1, \phi(g_2, s)) = \phi(g_1 * g_2, s)$, $\forall g_1, g_2 \in G$ and $s \in S$.*

When the group action is clear, we omit ϕ and use $g(x)$, or simply $g.s$, instead of $\phi(g, s)$. Following is an example of group action.

Example 3.8. *Let the group be D_n and the set S be $\{0, 1, \dots, n-1\}$. We discussed earlier that each group element of D_n can be identified as one of the $2n$ labeled configurations of a regular n-gon. For a fixed group element g, $g(x)$ is the position of the xth vertex of n-gon in the configuration corresponding to g. For example, for all $x \in S$, $e(x) = x$, $\sigma(x) = (x + 1)\%n$, $\tau(x) = n - x$ and so on.*

Notes - Lecture 3

Scribe: Kushagra Shandilya

In the last lecture, we defined group homomorphisms and group actions. We will now take a deep dive intro group actions.

4 Group Actions

A group action is a function of the form $\rho : G \times S \to S$ which follows certain properties, namely

1. $\phi(e, s) = s$, $\forall s \in S$
2. $\phi(g_1, \phi(g_2, x)) = \phi(g_1 * g_2, s)$, $\forall g_1, g_2 \in G$ and $s \in S$.

Let us fix an element g of G. Then g acts on S to give a permutation of S (Why is it a permutation?). Thus, equivalently, ϕ is a map from G to $Sym(S)$, the group of permutation of S. More formally, actions of a group G on a set S are the same as group homomorphisms from G to $Sym(S)$.

4.1 Orbit and Stabilizer

For each $s \in S$, its *orbit* is

$$Orb_s = \{s' \mid \exists g \in G, g.s = s'\}, \tag{1}$$

and its *stabilizer* is

$$Stab_s = \{g \mid g.s = s\}. \tag{2}$$

We also use the notation G_s to denote the stabilizer of an element s.

We can define a relation $\sim$ on $S \times S$. For any $s, s' \in S$, $s \sim s'$ if there exists some $g \in G$ such that $g.s = s'$. It can be easily verified $\sim$ is an equivalence relation on S. From this, we can conclude there exist some $T \subseteq S$ such that $S = \bigcup_{t \in T} Orb_t$.

Example 4.1. *Let $S = \{0, 1, \ldots, n-1\}$ and the group action be $\mathbb{Z}_n \times S \to S$ defined as $(i, j) \mapsto i + j \mod n$. Then, $Orb_0 = \{0, 1, \ldots, n-1\} = S$. When there is only one equivalency class, we say S has one orbit.*

Example 4.2. *$D_n \times S \to S$ defined as $(\sigma, i) \mapsto i + 1 \mod n$ and $(\tau, i) \mapsto n - i$. S has one orbit.*

Example 4.3. *$S_n \times S \to S$ defined as $(\sigma, i) \mapsto \sigma.i = \sigma(i)$. Here also S has one orbit.*

Lemma 4.4. *$\phi : G \to Sym(S)$ then G acts on $S \times S$ as $g.(s_1, s_2) = (g.s_1.g.s_2)$. Similarly, G can act on S^n for any $n > 0$.*

Example 4.5. *$S_n \times S \times S \to S$. Then, $Orb_{(1,1)} \neq Orb_{(1,2)}$. This action has two orbits which are (i, j), where $i \neq j$, and (i, j), where $i = j$.*

Similarly, $S_n \times S \times S \times S \to S$ have orbits of the type $(i, i, i), (i, j, k), (i, j, i), (i, i, j)$ and (j, i, i).

Example 4.6. *$D_n \times S \times S$ have orbit classes $Orb_{(1,1)}, Orb_{(1,2)}, Orb_{(1,3)}$ and $Orb_{(1,4)}$.*

Example 4.7. *Let S_n acts on set $S = \{1, 2, \ldots, n\}$. Then $Stab_1$ are the permutations which do not move 1. Thus, $|Stab_1| = (n-1)!$ and $|Orb_1| = n$. Note that $|Stab_1| \times |Orb_1| = n! = |S_n|$. We will see soon that this is a trend in general.*

Lemma 4.8. *The stabilizer of an element $s \in G, Stab_s$ is a subgroup of G.*

The above lemma is easy to verify and so we leave it as an exercise for the reader.

Lemma 4.9. *Let group G acts on a set S. Let $s \in S$. Then, $Stab_{gs} = gStab_s g^{-1}$.*

Proof. Let $h \in Stab_s$. Then, $hs = s$.

$$\begin{aligned} ghg^{-1}(g.s) &= gh.s \\ &= g(h.s) \\ &= g.s \end{aligned} \tag{3}$$

Thus, $gStab_s g^{-1} \subseteq Stab_{gs}$.

For the other side, let $j \in Stab_{gs}$. Then,

$$jg.s = g.s$$
$$g^{-1}jg.s = s$$

Thus, $g^{-1}jg \in Stab_s$. Let $g^{-1}jg = h$. Then, $j = ghg^{-1}$, which implies $j \in gStab_s g^{-1}$. Consequently, $Stab_{gs} \subseteq gStab_s g^{-1}$. □

4.2 Cosets and Lagrange's theorem

Let H be a subgroup of G. Then, we define a left coset of H by $gH = \{gh \mid h \in H\}$, where $g \in G$. Thus, for every g, there is a left coset of H.

Proposition 4.10. *Since $g = g.e$, every element $g \in G$ lies in some coset of H, specifically the left coset of the element itself.*

Proposition 4.11. *Two cosets are either the same or disjoint.*

Proof. Let g_1H, g_2H be two left cosets of H and $g_1H \cap g_2H \neq \phi$. Then, $g_1h_1 = g_2h_2$ for some $h_1, h_2 \in H$. Thus, $g_1 = g_2h_2h_1^{-1}$. Each element of g_1H is of the form g_1h for some $h \in H$. Thus, $g_1h = (g_2h_2h_1^{-1})h$. Consequently, $g_1H \subseteq g_2H$. Similarly, we can prove $g_2H \subseteq g_1H$. Therefore, $g_1H = g_2H$. □

Proposition 4.12. *Each coset has the same size which is the size of the subgroup H.*

Proof. Let gH be a left coset of $H = \{h_1, h_2, \dots\}$ and $gH \neq H$. If $h_i \neq h_j$, then $gh_i \neq gh_j$. Thus, each gh_i gets mapped to a different element, or equivalently, there is a one-one correspondence between H and gH. □

Theorem 4.13. *(Lagrange's theorem) When G is a finite group, $[G : H] = |G|/|H|$.*

Proof. Since G is finite, H has finitely many left cosets, let them be $g_1H, g_2H, \dots, g_tH$. We know that two cosets are either the same or disjoint. Thus, $G = g_1H \cup g_2H \cup \dots \cup g_tH$. Here, the union is disjoint.

Since, every coset has the same size which is the size of the subgroup H, therefore, $|G| = [G : H]|H|$. □

4.3 Orbit Stabilizer theorem

Theorem 4.14. *(Orbit Stabilizer theorem) Let G be a group that acts on S. Then, for any $s \in S$, we have*

$$|Orb_s| = [G : Stab_s], \tag{4}$$

where $[G : Stab_s]$ is the cardinality of left cosets of $Stab_s$, also called the index of $Stab_s$ in G. When G is a finite group, then

$$|G| = |Stab_s||Orb_s|. \tag{5}$$

Proof. We define a map $\psi : Orb_s \to G/Stab_s$ by $gs \mapsto gStab_s$. We claim ψ is well defined and is a bijective map.

- Suppose $g_1, g_2 \in G$ and $g_1.s = g_2.s$. Then, $g_2^{-1}g_1.s = s$. Thus, $g_2^{-1}g_1 \in Stab_s$. which is equivalent to $g_1Stab_s = g_2Stab_s$. This proves ψ is well defined. The above argument in the opposite direction proves the injectivity as well.
- ψ is onto since for any $gStab_s, \psi(gs) = gStab_s$.

Thus, ψ is bijective and $|Orb_s| = [G : Stab_s]$. When, G is a finite group, from Lagrange's theorem, we get $|G| = |Stab_s||Orb_s|$. □

Example 4.15. *$D_6 \times S \times S$ have orbit classes $Orb_{(1,1)}, Orb_{(1,2)}, Orb_{(1,3)}$ and $Orb_{(1,4)}$.*

$Orb_(1,1) = \{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\}$. Then, $|Stab_(1,1)| = |G|/|Orb_(1,1)| = 12/6 = 2$, which are the specifically the identity and the reflection elements.

Example 4.16. *Let G be the group of rotations of a cube around body diagonals and $|G| = 24$. If s is some corner of the cube then $|Orb_s| = 8$, which are all the corners, then $|Stab_s| = 3$, the rotations which keep the corner fixed. Similarly, if s is some face, then $|Orb_s| = 6$ and $Stab_s = 4$.*

4.4 Actions on vector spaces

Let S be a set. Then $\mathbb{C}.S = \{\sum \alpha_i s_i \mid \alpha_i \in \mathbb{C}\}$ is a vector space of dimension $|S|$. A group G acts on a vector $v \in \mathbb{C}.S$ as follows,

$$g.(\alpha_1 s_1 + \dots + \alpha_k s_k) = (\alpha_1 g.s_1 + \dots + \alpha_1 g.s_k). \tag{6}$$

Notes - Lecture 4

Scribe: Vempalli Venkata Sai Keerthana

5 Recall

Lets recall few definitions from previous classes

$$Orb_s = \{s' \mid \exists g \in G, g.s = s'\}, \tag{7}$$

$$Stab_s = \{g \mid g.s = s\}. \tag{8}$$

$$|G| = |Stab_s||Orb_s|. \tag{9}$$

6 Actions on vector spaces

Let S be a set. Then $\mathbb{C}.S = \{\sum c_i s_i \mid c_i \in \mathbb{C}\}$ is a vector space of dimension $|S|$. A group G acts on a vector $v \in \mathbb{C}.S$ as follows,

$$g.(c_1 s_1 + \cdots + c_k s_k) = (c_1 g.s_1 + \cdots + c_1 g.s_k). \tag{10}$$

Let $x_i(n)$ is the coefficient of s_i Lets consider a necklace of length 6. 1->2->3->4->5->6->1

Let $c_1(n) = 0; c_2(n) = 1; c_3(n) = 1; c_4(n) = 0; c_5(n) = 0; c_6(n) = 2$

Then $\sigma(n) = 0.s_2 + 1.s_3 + 1.s_4 + 0.s_5 + 0.s_6 + 2.s_1$

Therefore $x_1(n) = 0; x_1(\sigma n) = 2$

6.1 Action of groups on functions

Similar to the action σ we can say g(f(n))=$f(g^{-1}(n))$. Similarly $g_1 g_2(f(n)) = f(g_2^{-1} g_1^{-1}(n))$. This can be proved ,by considering $g_2 f$ as f'. Then $g_1(f')(n) = f'(g_1^{-1}(n)) = g_2(f)(g_1^{-1} n) = f(g_2^{-1} g_1^{-1} n)$.

Action on f acts as an inverse action on necklace

Let $X_1, X_2, ..., X_n$ be functions from $\mathbb{C}$ S to $\mathbb{C}$ and β be the nth root of unity(β^n=1).

Let us define $f_\beta(n) = \beta X_1(n) + \beta^2 X_2(n) + \cdots + \beta^n X_n(n)$

Lets us look at the function $f : \mathbb{C}\ S \to \mathbb{C}$ such that $f(n) = f(\sigma_n) \forall \sigma \in G$ and $f(n) \neq f(n')$ for some n' not in orbit of n. This function is expected to separate the orbits.

$\sigma(f_\beta) = \beta X_n + \beta^2 X_1 + \cdots + \beta^n X_{n-1}$

$\implies \sigma(f_\beta) = \beta(f_\beta)$

Similarly $\sigma(f_{\beta^2}) = \beta^2(f_{\beta^2})$

So $\sigma^i(f_{\beta^j}) = \beta^{ij}(f_{\beta^j})$

Suppose if we have the values of $f_\beta(n), f_{\beta^2}(n), \ldots, f_{\beta^n}(n)$ then we can determine the values of $c_1, c_2, \ldots, c_n$ exactly by

$$\begin{bmatrix} f_\beta(n) \\ \vdots \\ f_{\beta^n}(n) \end{bmatrix} = \begin{bmatrix} \beta & \beta^2 & \ldots & \beta^n \\ \beta^2 & \beta^4 & \ldots & \beta^{2n} \\ \vdots & \vdots & \ldots & \vdots \\ \beta^n & \beta^{2n} & \ldots & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

Given LHS we can compute RHS, Hence $c_1, c_2, \ldots, c_n$

6.1.1 Construction of invariants

How to construct invariants Given f_1, f_2

g(f_1)=ψ_1(g)f_1

g(f_2)=ψ_2(g)f_2

g($f_1 f_2$)=$g(f_1)g(f_2)$

$(f_1f_2)\bar{n} = f_1(\bar{n})f_2(\bar{n})$

$g(f_1f_2)(\bar(n) = (f_1f_2)(g^{-1}\bar{n}) = f_1(g^{-1}\bar{n})f_2(g^{-1}\bar{n}) = (gf_1)(\bar{n})(gf_2)(\bar{n})$

$g(f_1f_2) = \psi_1(g)\psi_2(g)(f_1f_2)$

$\sigma f_\beta = \beta f_\beta$

$\sigma f_{\beta^{n-1}} = \beta^{n-1} f_{\beta^{n-1}}$

$\sigma(f_{f_\beta \beta^{n-1}}) = f_\beta f_{\beta^{n-1}}$

$\beta X_n + \beta^2 X_1 + \cdots + \beta^n X_{n-1}$

$\beta^{n-1} X_n + \beta^{(n-1)2} X_1 + \cdots + \beta^n X_{n(n-1)}$

Question 1: $(iX_1 - X_2 - iX_3 + X_4)(-iX_1 - X_2 + iX_3 + X_4)$ calculate the invariants ....

Question 2: $f_i f_j f_k$ such that $i + j + k \equiv 0(\mod n)$
$y_0, y_1, \ldots, y_{n-1}$

$y_0^{d_0} y_1^{d_1} \ldots y_{n-1}^{d_{n-1}}$

$d_0.0 + d_1.1 + d_2.2 + \cdots + d_{n-1}.n - 1 = 0$

all such values
$S_n- > S_1, S_2, \ldots, S_n$

$\sigma(S_i) = S_{\sigma(i)}$

$\bar{c} = c_1 s_1 + c_2 s_2 + \cdots + c_n s_n$

n=5

$$\begin{bmatrix} 3 & 0 & 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{bmatrix}$$

Given $\bar{c}, \bar{d} \in c^n$ are they in the same orbit
Algorithm: Sort $\bar{c}, \bar{d}$ and check equality
Algebraic: $\alpha \in \mathbb{C}$ $\prod_{i=1}^{n}(c_i - \alpha) = \prod_{i=1}^{n}(d_i - \alpha)$

6.2 Gallelian Group

Group which consists of physics

$$M(\theta, a, b) \ = \left[\begin{array}{ccc} \cos\theta & -\sin\theta & a \\ -\sin\theta & \cos\theta & b \\ 0 & 0 & 1 \end{array} \right] \left[\begin{array}{c} x \\ y \\ 1 \end{array} \right]$$

where, $y' = x\cos\theta + y\sin\theta + a and x' = -x\sin\theta + y\cos\theta + b$

$$M(\theta, a, b) \ = \left[\begin{array}{c} x_1 \\ y_1 \\ 1 \end{array} \right] = \left[\begin{array}{c} x_1' \\ y_1' \\ 1 \end{array} \right]$$

Notes - Lecture 5

Scribe: Kunal Kundwani

Abstract

In this lecture, we would mostly cover linear algebra and look at some important vector spaces.

7 Recall

Lets first recall a few definitions, concepts and results from previous classes.

Definition 7.1 (Group Action)**.** *Given a group G and a set S, a function $\rho\colon G \longrightarrow$ Bij(S) (the group of bijections on S) is called a group action of G on S if it is also a homomorphism between the above mentioned domain and range groups.*

Definition 7.2 (Group action on functions)**.** *Given a group G, a set S and an action ρ of G on S, the corresponding group action on the set of functions from S to S is defined as, $(g.f)(s) = f(g^{-1}.s)$ for all functions f on S and $s \in S$.*

Definition 7.3 (Orbit)**.** *For a group action ρ of G on S, the Orbit O_s corresponding to $s \in S$ is defined as the set $\{v \in S \mid \exists g \in G \text{ s.t. } v = g.s\}$.*

Definition 7.4 (Stabilizer)**.** *For a group action ρ of G on S the stabilizer G_s corresponding to $s \in S$ is defined as the set $\{g \in G \mid g.s = s\}$.*

Observe that, the stabilizer of any element in S is a subgroup of G.

Theorem 7.5. *For a finite group G acting on a set S:*

$$|O_s|.|G_s| = |G|, \forall s \in S$$

Definition 7.6 (Orbit invariant functions)**.** *For a group G acting on a set S, a function f with S as its domain is said to be orbit invariant if it maps elements of the same orbit to the same value.*

8 Vector Spaces

Definition 8.1 (Vector Space)**.** *A vector space is a set V equipped with a field $\mathbb{F}$ and two binary operators $(+) : V \times V \longrightarrow V$ and $(\cdot) : \mathbb{F} \times V \longrightarrow V$ satisfying:*

- *V is a group under $(+)$*
- $a \cdot (b \cdot v) = (ab) \cdot v \ \forall a, b \in \mathbb{F}, v \in V$
- *$1 \cdot v = v \ \forall v \in V$ where 1 is the identity of $\mathbb{F}$*
- $a \cdot (u + v) = a \cdot u + a \cdot v, \ \forall u, v \in V, a \in \mathbb{F}$
- $(a + b) \cdot v = a \cdot v + b \cdot v \ \forall a, b \in \mathbb{F}, v \in V$

The field $\mathbb{F}$ is generally taken to be $\mathbb{C}$ or $\mathbb{R}$. Also, $a.v$ is generally also written as av. Elements of $\mathbb{F}$ are also referred to as scalars

Definition 8.2 (Subspace)**.** *A subset W of V is said to be a subspace of V if it itself is a vector space with the same operations $(+)$ and $(\cdot)$ and field $\mathbb{F}$. An alternate definition could be that a subset W of V is called a subspace of V if it is closed under addition and scalar multiplication.*

Definition 8.3 (Linear Independence)**.** *A set $S \subseteq V$ is said to be linearly independent if for all finite subsets T of S s.t. $T = \{v_1, v_2, \ldots, v_k\}$, there do not exist non-zero scalars $a_1, a_2, \ldots, a_k$ such that $\sum_{i=1}^{i=k} a_i v_i = 0$.*

Definition 8.4 (span)**.** *For a given subset S of V, its span is defined as the set of all $v \in V$ for which there exist (finitely many) vectors $v_1, v_2 \ldots v_k \in S$ and scalars $a_1, a_2, \ldots, a_k$ such that $v = \sum_{i=1}^{i=k} a_i v_i = v$, and usually denoted as span(V).*

It is easy to verify that span of any subset of S is actually a subspace of V.

Definition 8.5 (basis)**.** *A subset B of V is said to be a basis of a vector space V if B is linearly independent and span(B)= V.*

Theorem 8.6. *If B_1 and B_2 are two bases of a vector space V and at least one of them is finite in cardinality, then $|B_1| = |B_2|$.*

Proof. Assume, to the contrary, that $|B_1| \neq |B_2|$ and WLOG let $|B_1| < |B_2|$. Hence B_1 is finite. Let $B_1 = \{v_1, v_2, \ldots, v_m\}$. Hence, $v_1 = \sum_{i=1}^{i=n} a_i w_i$ for some n and $w_i \in B_2, a_i \neq 0\ \forall i \in [n]$. Observe that $B_2 \backslash \{w_1\} \cup \{v_1\}$ is also a basis. (left as an exercise for the reader) Hence update B_2 as above, i.e., by removing w_1 and adding v_1 to it. Now, we can continue this swapping procedure inductively as follows: write v_i for $i \leq m$, as a linear combination of vectors in (the updated) B_2. Atleast one of the vectors involved in this linear combination will not belong to B_1 since otherwise B_1 would become linearly dependent. Let w_i be one such vector, and swap w_i in B_2 with v_i.

Hence, after doing this once for every element in B_1, we have replaced atmost m vectors in B_2 by some vectors in B_1 while maintaining the cardinality and the span of B_2. Hence, since $|B_2| > |B_1|$, there exists an element w in B_2 which was never replaced and remained in B_2 throughout. However since B_1 is a basis, the set $\{v_1, v_2, \ldots, v_m, w\} \subseteq B_2$ is linearly dependent, contradicting the linear independence of B_2. Thus, our assumption of $|B_1| \neq |B_2|$ must be wrong, and we are done.

Note that the same proof actually works for the case when one of the sets is countable! □

Definition 8.7 (Dimension)**.** *If a vector space has a finite basis, then all its bases will have the same size. Then the size of any basis set of V is defined as its dimension, and usually denoted as $dim(V)$.*

Definition 8.8 (Linear Map)**.** *Given vector spaces V and W, a function $T : V \longrightarrow W$ is called a linear map if $\forall u, v \in V, a, b \in \mathbb{F}, T(au + bv) = aT(u) + bT(v)$.*

Following are some (easy, but useful) results:

Theorem 8.9. *Let V be a vector space and W be any subspace of it. Then,*

- *$dim(W) \leq dim(V)$.*
- *If B_W is a basis of W then there exists $S \subseteq V$ such that $S \cup B_W$ is a basis of V.*
- *$dim(W) = dim(V)$ iff $W = V$.*

Definition 8.10 (Kernel)**.** *Given a linear map $T : V \longrightarrow W$ between two vector spaces, its kernel is defined as the set $\{v \in V \mid Tv = 0\}$, represented by $ker(T)$.*

Definition 8.11 (Image)**.** *Given a linear map $T : V \longrightarrow W$ between two vector spaces, its image is defined as the set $\{w \in W \mid \exists v \in V \text{ s.t. } Tv = w\}$. Also represented as $Im(T)$.*

It is easy to verify that the kernel and the image of any linear map are indeed vector spaces.

Theorem 8.12. *Let $T : V \longrightarrow W$ be a linear map. Then,*

$$dim(ker(T)) + dim(Im(T)) = dim(V)$$

Proof. Let B be a basis of the kernel of T. Then, by the previous theorem, there exists $S \subseteq V$ such that $B \cup S$ is a basis of V. Observe that the set $X = \{Ts \mid s \in S\}$ is a basis for $Im(T)$. (More precisely,$|X| = |S|$ and X is a basis for $Im(T)$) (Easy exercise). Thus,

$$dim(V) = |B| + |S| = |B| + |X| = dim(ker(T)) + dim(Im(T))$$

□

Here are a few examples of some well known bases of some well known vector spaces. From here on, V will denote the vector space in concern and B its basis.

Example 8.13.

$$V = \mathbb{R}^n \text{ with the standard basis } B = \{e_1, e_2, \ldots, e_n\}$$

where e_i is the vector whose entries are all 0 except for the ith one, which is 1. ($\forall i \in [n]$)

Example 8.14 (Taylor Basis)**.**

$$V = \{\text{polynomials of degree} \leq n\}$$

$$B = \{1, x, x^2, \ldots, x^n\}$$

This basis is called the Taylor basis since these terms appear in the Taylor expansion of any function around 0.

Example 8.15 (Lagrangian Basis)**.**

$$V = \{\textit{polynomials of degree} \leq n\}$$

Let $A = \{x_0, x_1, \dots, x_n\}$*, then the Lagrangian basis of* V *for the set* A *is defined as:*

$$L_A = \left\{ L_i := \frac{\prod\limits_{\substack{j=1 \\ j\neq i}}^{j=n} (x - x_j)}{\prod\limits_{\substack{j=1 \\ j\neq i}}^{j=n} (x_i - x_j)} \Bigg| i \in \{0, 1 \dots, n\} \right\}$$

These terms appear in the Lagrangian polynomial interpolation formula and hence the name.

Example 8.16 (Bernstein Basis)**.**

$$V = \{\textit{polynomials of degree} \leq n\}$$

The Bernstein basis is defined as:

$$B = \{B_i^n \mid i \in 0, 1, \dots, n\}$$

where,

$$B_i^n := \binom{n}{i} x^i (1-x)^{n-i}$$

The name and significance of this basis comes from the following theorem.

Theorem 8.17 (Bernstein-Weierstrass Theorem)**.** *Given any function* $f \in C[0,1]$ *and any real* $\epsilon > 0$*, then* $\lim_{n\to\infty} B^n(f) = f$ *(point-wise) where*

$$\forall x \in [0,1], B^n(f)(x) := \sum_{i=0}^{i=n} f\left(\frac{i}{n}\right) B_i^n(x)$$

Example 8.18.

$$V = \mathbb{R}^{m\times n}$$

$$B = \{e_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\} \textit{ where,}$$

e_{ij} *is the* $m \times n$ *matrix with all entries* 0 *except for the* (i,j)*th one, which is* 1*.*

Definition 8.19 (Lie Bracket)**.** *Given* $A, B \in \mathbb{R}^{n\times n}$*, their lie product is defined as,*

$$[A, B] = AB - BA$$

Observe that the set of $n \times n$ (anti)symmetric matrices is closed under Lie product, but not under normal matrix product!

9 Tangent Spaces

Let's start with an example for motivation.

Example 9.1. *Given a sphere* $x^2 + y^2 + z^2 = 1$*, and a point* $(1,0,0)$ *on it, we want to know all the directions in which i can "move slightly" while still being "almost" on the sphere.*
Hence, We want to find the direction vectors $v = (x_0, y_0, z_0)$ *for which* $(1,0,0) + \epsilon v$ *is "almost on the sphere" for small* $\epsilon > 0$*, i.e.*

$$(1 + \epsilon x_0)^2 + \epsilon^2 y_0^2 + \epsilon^2 z_0^2 \approx 1$$

$$\implies 2x_0\epsilon + (x_0^2 + y_0^2 + z_0^2)\epsilon^2 \approx 0$$

Thus $x_0 = 0$ *suffices since it makes the* ϵ *term vanish. Hence,* $y - z$ *plane is our desired set of directions!*

Lets try formalising all this,

Let us say, in $\mathbb{R}^n$(or some other suitable space), we are given a set of scalar coordinate wise differentiable functions $F = \{f_1, f_2, \dots, f_r\}$. Then,

Definition 9.2 (Variety)**.** *Variety of this set* F *is defined as the set of vectors* p *for which* $f_i(p) = 0, \forall i \in [r]$*. Let's denote it by Var(F).*

Then our problem essentially is to find all directions v for which

$$\forall p \in \mathrm{Var}(F), i \in [r], \lim_{h \to 0} \frac{f_i(p + hv)}{h} = 0$$

The expression in the above limit is nothing but the derivative of f_i in the direction v at the point p ($|v|$ times the derivative, to be precise).

Thus, we arrive at a rather formal definition of Tangent spaces as follows:

Definition 9.3 (Tangent Space)**.** *The tangent space of F is defined as the set of all vectors $v \in \mathbb{R}^n$ for which,*

$$v.((\nabla f_i)(p)) = 0 \ , \forall p \in \mathit{Var}(F), i \in [r]$$

Notes - Lecture 6

Scribe: Vaibhav Krishan

In this lecture, we will study some more properties of linear transformations and understand which invariant holds while applying a certain type of linear transformation.

10 Linear Transformations

Let V be a vector space, over $\mathbb{R}$ or $\mathbb{C}$ for now. We start by defining a basis for V.

10.1 Basis

Vectors $v_1, \ldots, v_k \in V$ are called linearly dependent if there exists $\lambda_1, \ldots, \lambda_k$, not all 0, such that $\sum_{i=1}^{k} \lambda_i v_i = 0$. $v_1, \ldots, v_k$ are called linearly independent if and only if they are not linearly dependent.

Definition 10.1 (Basis)**.** *A basis for V is a maximal set of independent vectors in V.*

Let there be a basis $b_1, \ldots, b_n$ for V. Then any vector $v \in V$ can be written as a linear combination of $b_1, \ldots, b_n$, as otherwise $v, b_1, \ldots, b_n$ would be linearly independent which contradicts the maximality of the given basis. Moreover, the linear combination is unique, as otherwise there is a linear combination of $b_1, \ldots, b_n$ that equals 0. Therefore, given a set of vectors $c_1, \ldots, c_m$, one can find a unique matrix A, more abstractly called a linear transformation, of dimensions $m \times n$ such that

$$\begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} = A \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

We would like to understand the conditions for $c_1, \ldots, c_m$ to be a basis as well. The following result proves the conditions that are equivalent to it being another basis.

Lemma 10.2. *$c_1, \ldots, c_m$ is a basis for V if and only if:*

1. *$m = n$.*
2. *A is an invertible matrix.*

Proof. $m = n$ can be proved using Steinitz exchange lemma, see Steinitz exchange lemma.

To see that A is invertible, note that as both $b_1, \ldots, b_n$ and $c_1, \ldots, c_n$ are a basis for V, we can also write

$$\begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = A' \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

Therefore

$$\begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = A'A \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

As $b_1, \ldots, b_n$ are linearly independent, this is only possible if $A'A = I$. Hence $A' = A^{-1}$, therefore A is invertible. □

This also proves that the dimension of a vector space V can be well defined as the size of a basis of the vector space. This will be denoted by $dim(V)$.

10.2 Vector Spaces from Linear Transformations

Now we will study some vector spaces associated with a linear transformation. Consider a matrix $A \in \mathbb{C}^{m \times n}$ of the following form:

$$\begin{bmatrix} c_1 \ldots c_n \end{bmatrix} = A = \begin{bmatrix} r_1 \\ \vdots \\ r_m \end{bmatrix}$$

Then following four spaces will be of importance to us:

1. The column space of A, denoted by col-sp(A), is the subspace of $\mathbb{C}^m$ spanned by $c_1, \ldots, c_n$. The dimension of the column space is called the column rank of A, denoted by col-rank(A).
2. The row space of A, denoted by row-sp(A), is the subspace of $\mathbb{C}^n$ spanned by $r_1, \ldots, r_m$. The dimension of the row space is called the row rank of A, denoted by row-rank(A).
3. The column nullspace of A is col-null $= \{v \in \mathbb{C}^n : Av = 0\}$.
4. The row nullspace of A is row-null $= \{v \in \mathbb{C}^m : v^T A = 0\}$.

The following relationship holds between row space and row nullspace, and similarly between column space and column nullspace.

Theorem 10.3 (Rank-Nullity Theorem for Linear Transformations). *$dim(row\text{-}null(A)) + dim(row\text{-}sp(A)) = m$. $dim(col\text{-}null(A)) + dim(col\text{-}sp(A)) = m$.*

Proof. Consider the following linear map: $\phi : \mathbb{C}^m \to \mathbb{C}^n$ where $\phi(e_i) = r_i$. Then $Im(\phi) =$ row-sp(A) and $Ker(\phi) =$ row-null(A). Using the rank nullity theorem for linear maps proves the desired result. A similar argument works for column space and column nullspace. □

Example 10.4. *Consider the matrix*

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$$

It can be verified that the row space of A has dimension 2 and the row nullspace of A is spanned by $[1, -2, 1]$.

10.3 Row Echelon Form

Now we will look at a special matrix, called the row echelon form of a matrix, which will carry some important invariants.

Definition 10.5. *For every matrix A, with $dim(row\text{-}sp(A)) = k$, a matrix E is called the row echelon form of A, denoted by $ref(A)$, if the following holds for some $i_1 < \ldots < i_k$:*

1. *$row\text{-}sp(A) = row\text{-}sp(E)$,*
2. *$E[\{1, \ldots, k\}, \{i_1, \ldots, i_k\}] = I_{k \times k}$,*
3. *$E_{i,j} = 0$ for all $i > k$,*
4. *$E_{i,i_j} = 0$ for all $i > i_j$ for any j.*

Example 10.6. *Let*

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$$

Then its row echelon form is

$$E = \begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Note this form is obtained only by left multiplication of linear transformations. Reducing any remaining non-zero entry in E will require right multiplication which, as we will see in a later example, may not preserve row space and therefore, is not allowed.

Theorem 10.7. *For each matrix A, there exists a row echelon form $ref(E)$ and it is unique.*

The proof is left as an exercise.

Remark. *The uniqueness of row echelon form automatically implies that two matrices with the same row space have the same row echelon form. Hence there is a one-to-one correspondence between row echelon forms and row spaces.*

An application of row echelon form can be seen as follows. Let $w \in \mathbb{C}^n$ be a vector and we want to know whether $w \in$ row-sp(A) or not. This can be decided using ref(A) as follows. Simply check that $\sum_{i=1}^{k} w_{i_1} r_i = w$ where r_i denote the rows of ref(A) and $i_1, \ldots, i_k$ are as per definition 10.5. This is equivalent to $w \in$ row-sp(ref(A)) $\iff$ $w \in$ row-sp(A).

10.4 Invariants over Linear Transformations

Now we will see which invariants hold under certain types of linear transformations.

Lemma 10.8. *Let $T \in \mathbb{C}^{m\times m}$ be invertible. Then row-sp(TA) = row-sp(A).*

Proof. All rows of row-sp(TA) can clearly by written as a linear combination of rows of A. Therefore row-sp$(TA) \subseteq$ row-sp(A). It is easy to see that the converse holds as well as T is invertible. Therefore the two are equal, completing the proof. □

Let $\mathcal{V} = C^{m\times n}$ and let $G = \mathsf{GL}_m$ i.e. invertible matrices of dimension $m \times m$. Consider the group action $g \cdot v$ as left multiplying v by g. Then every orbit has the same row space. Therefore there is a one-to-one correspondence between orbits under this action and row echelon forms. Stated differently, each orbit intersects with the "REF condition" exactly once.

Now we will prove, using the row echelon form, that the column rank and row rank of a matrix are the same.

Theorem 10.9. *col-rank(A) = col-rank$(\mathsf{ref}(A))$ = row-rank$(\mathsf{ref}(A))$ = row-rank(A).*

Proof. Let $E = \mathsf{ref}(A)$. E and A have the same row rank as their row spaces themselves are the same. It is easy to show that the row rank and column rank of E are the same using its definition. Hence, all that needs to be proved is that column rank of A and E are the same.

We prove this using Theorem 10.3 as follows. Note that $E = TA$ for some invertible T. Therefore for any $v \in V$, $Av = 0 \iff TAv = 0 \iff Ev = 0$. Hence the column nullspace of A and E are the same. Using Theorem 10.3, it follows that their column rank are the same as well. □

Remark. *Note the following:*

- *Left multiplication by an invertible linear transformation preserves the column nullspace as well as the row space. This exemplifies a duality between the column nullspace and the row space.*
- *While the column rank of A and E are the same, their column spaces need not be the same, as will be seen in the next example.*

Example 10.10. *Consider the following transformations on the given matrix to reach its row echelon form:*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

It can be seen that the column space changes dramatically. While before the first step there could be non-zero entries in the third place of a column vector, they are always zero for the columns in the last two steps. Nonetheless, the column rank remains the same.

10.5 Tangent Spaces

Finally, let's return to the study of tangent spaces. Let $f_1, \ldots, f_r : \mathbb{C}^n \to \mathbb{C}$ be complex functions. Their variety is defined as $V(f_1, \ldots, f_r) = \{z \in \mathbb{C}^n \mid \forall i \in [r], f_i(z) = 0\}$. The tangent space at a point $p \in V(f_1, \ldots, f_r)$ is defined as $T_p(V(f_1, \ldots, f_r)) = \{v \mid \forall i \in [r], f_i(p + \varepsilon v) = 0, \varepsilon^2 = 0\}$.

The tangent space can be thought of as the "possibilities" that can arise by small movements from a certain point. It has numerous practical applications in civil engineering, mechanical engineering etc. as studying tangent spaces corresponds to studying the dynamics of a system around an equilibrium. Another example of importance of tangent spaces is in gradient descent.

The tangent space can be seen to be a vector space as following. For each f_i, consider its gradient vector $\boldsymbol{\nabla} f_i(p)$. Then $f_i(p + \varepsilon v) = 0 \iff \boldsymbol{\nabla} f_i(p) \cdot v = 0$. Therefore $T_p(V(f_1, \ldots, f_r)) = \{v \mid \forall i \in [r], \boldsymbol{\nabla} f_i(p) \cdot v = 0\}$. This can be restated as $\begin{bmatrix} \boldsymbol{\nabla} f_1(p) \\ \vdots \\ \boldsymbol{\nabla} f_r(p) \end{bmatrix} v = 0$. This shows that v is in the nullspace of the gradient matrix, therefore $T_p(v)$ is a vector space.

The dimension of the tangent space specifies the local behaviour around a point in the given variety.

Example 10.11. *Let $f_1(x, y, z) = x^2 + y^2 - z^2, f_2(x, y, z) = y - z$ be two functions from $\mathbb{C}^3$ to $\mathbb{C}$.*

Consider the point $p = (0, 1, 1)$. Then $\boldsymbol{\nabla} f_1(p) = (0, 2, -2), \boldsymbol{\nabla} f_2(p) = (0, 1, -1)$. As the first column is 0 and the last two columns are linearly dependent, the column rank is 1, hence the column nullspace has dimension 2. Therefore, the tangent space at p has dimension 2. The tangent space can be characterised as (c_1, c_2, c_2) for any $c_1, c_2 \in \mathbb{C}$.

Now consider the point $q = (1, 0, 0)$. Then $\boldsymbol{\nabla} f_1(q) = (2, 0, 0), \boldsymbol{\nabla} f_2(q) = (0, 1, -1)$. There are two linearly independent columns, and there are two rows, hence the column rank is 2 and the column nullspace has dimension 1. Therefore, the tangent space at q has dimension 1. The tangent space can be characterised as $(0, c, c)$ for any $c \in \mathbb{C}$.

10.6 Review

We studied what is a basis and the dimension of a vector space. We studied linear transformations, certain vector spaces associated to them, their row echelon form, and invariants that hold under certain linear transformations. We closed with the definition of a tangent space and why it is a vector space.

Notes - Lecture

Scribe: Venkata Sai Keerthana

$\rho : G \to GL(v)$

$G \curvearrowright v$

$\rho' : G \to GL(v')$

$\rho \otimes \rho' : G \to GL(v \otimes v')$

consider v,v' vectorspaces, $v \otimes v' = \{\sum_{i=1}^{K} v_i \otimes v_i' | K > 0, v_i \in V_i, v_i^{-1} \in V_i'\}$

$v \otimes (v_1' + v_2') = v \otimes v_1' + v \otimes v_2'$

$(v_1' + v_2') \otimes v' = v_1 \otimes v' + v_2 \otimes v'$

v has a basis $b_1, b_2, ..., b_m$

v' has a basis $b_1', b_2', ..., b_m'$

$v \otimes v'\ has\ a\ basis\ \{b_i \otimes b_j | b_i \in B, b_i' \in B'\}$

$\rho(g)\ is\ m \times mmatrix$

$\rho'(g)\ is\ m' \times m'matrix$

$\underbrace{\rho \otimes \rho'(g)\ is\ mm' \times mm'matrix}_{tensor\ product\ matrix}$

$\rho \otimes \rho' \curvearrowright v \otimes v'$

Example: GL_3

$$\begin{bmatrix} \rho(x_1) \\ \rho(x_2) \\ \rho(x_3) \end{bmatrix} \begin{bmatrix} \rho'(y_1) \\ \rho'(y_2) \\ \rho'(y_3) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

where $v \to (x_1, x_2, x_3)\ and\ v' \to (y_1, y_2, y_3)$

$X_1 \otimes Y_1 \to (a_{11}x_1 + a_{12}x_2 + a_{13}x_3) \otimes (a_{11}y_1 + a_{12}y_2 + a_{13}y_3)$

$X_1 \otimes Y_1 \to a_{11}^2 x1 \otimes y_1 + a_{11}a_{12}x1 \otimes y_2 + ... + a_{13}a_{13}x_3 \otimes y_3$

$$\begin{bmatrix} \rho(x_1) \\ \rho(x_2) \\ \rho(x_3) \end{bmatrix} \begin{bmatrix} \rho'(y_1) \\ \rho'(y_2) \\ \rho'(y_3) \end{bmatrix} = \begin{bmatrix} A \end{bmatrix} \begin{bmatrix} x_1 \otimes y_1 \\ \vdots \\ x_3 \otimes y_3 \end{bmatrix}$$

$Z_{12} = X_1 \otimes Y_2 - X_2 \otimes Y_1$

$Z_{13} = X_1 \otimes Y_3 - X_3 \otimes Y_1$

$Z_{23} = X_2 \otimes Y_3 - X_3 \otimes Y_2$

$\Rightarrow (a_{11}x_1 + a_{12}x_2 + a_{13}x_3) \otimes (a_{21}y_1 + a_{22}y_2 + a_{23}y_3) - (...)(...)$

$\to X_1 \otimes Y_2(a_{11}a_{22} - a_{21}a_{12}) + X_2 \otimes Y_1(a_{12}a_{21} - a_{22}a_{11})$

$\Rightarrow (a_{11}a_{22} - a_{21}a_{12})(z_{12})....$

$Z = \mathbb{C}\ z_{12} + \mathbb{C}\ z_{13} + \mathbb{C}\ z_{23}$

$\rho \otimes \rho'(G)(z) \subseteq z$

$S = \mathbb{C}\ \{X_1 \otimes Y_1, ..., X_1 \otimes Y_2 + X_2 \otimes Y_1, ...\}$

$v \otimes v' = z \otimes S$

where z is a 3-dimensional G-module $\neq\ \mathbb{C}3$

S is a 3-dimensional G-module.

Notes - Lecture 8

Scribe: Shantanu Nene

In this lecture, we will study determinants and other multilinear alternating maps.

11 Multilinear Alternating Maps

Consider the space $\mathbb{C}^{n\times r} = \underbrace{\mathbb{C}^r \times \mathbb{C}^r \times \cdots \times \mathbb{C}^r}_{n\text{times}}$. We think of the input as n row vectors of r entries each.

Definition 11.1. *A function $f : \mathbb{C}^{n\times r} \to \mathbb{C}$ is called multilinear and alternating if it is linear in each row, and* $f\left(\begin{bmatrix} \vdots \\ R_i \\ R_{i+1} \\ \vdots \end{bmatrix}\right) = -\begin{bmatrix} \vdots \\ R_{i+1} \\ R_i \\ \vdots \end{bmatrix}$ *when swapping any two adjacent rows.*

From now on, we assume f is a multilinear alternating map. It is evident from the definition that swapping any two rows in the input changes the sign of the output.

Lemma 11.2. *Let $m \in \mathbb{C}^{n\times r}$ such that two of its rows are equal. Then $f(m) = 0$.*

Proof. When we swap the two same rows, the output must change its sign, but m remains the same. Hence $f(m) = 0$ □

Lemma 11.3. *f transforms in the following way under elementary row transformations:*

1. $f\left(\begin{bmatrix} R_1 \\ \vdots \\ R_n \end{bmatrix}\right) = \frac{1}{a}\begin{bmatrix} aR_1 \\ \vdots \\ R_n \end{bmatrix}$ *for any $a \neq 0$.*

2. $f\left(\begin{bmatrix} \vdots \\ R_i \\ \vdots \\ R_j \\ \vdots \end{bmatrix}\right) = -\begin{bmatrix} \vdots \\ R_j \\ \vdots \\ R_i \\ \vdots \end{bmatrix}$ *for any two rows R_i, R_j.*

3. $f\left(\begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{bmatrix}\right) = \begin{bmatrix} R_1 \\ R_1 + aR_2 \\ \vdots \\ R_n \end{bmatrix}$ *for any $a \in \mathbb{C}$.*

Proof. The first two properties follow from definition of multilinear alternating, and the third follows from linearity and the previous lemma. □

We can get any matrix M into its row echelon form ref(M) using elementary row transformations, and using the above lemma we get $f(M) = \alpha f(\text{ref}(M))$ for some $\alpha \neq 0$. We first study 3 cases:

Case I: $n > r$.
In this case, the last row of ref(M) is 0, so $f(M) = 0$ always. Therefore f is identically 0.
Case II: $n = r$ and rank$(M) < n$.
In this case, $f(M) = 0$ because again the last row of ref(M) is 0.
Case III: $n = r$ and rank$(M) = n$.
In this case, ref(M) must be the identity matrix I_n, so $f(M) = \alpha f(I_n)$.

11.1 The Determinant

Definition 11.4. *For a matrix $M_{n\times n} = (m_{i,j})$, its determinant is defined as*

$$\det M = \sum_{\sigma\in S_n} (-1)^{sgn(\sigma)} \prod_{i=1}^{n} m_{i,\sigma(i)}.$$

Proposition 11.5. det *is a multilinear alternating map.*

The proof is left as an exercise.

Theorem 11.6. *Suppose $f : \mathbb{C}^{n\times n} \to \mathbb{C}$ is a multilinear alternating map with $f(I_n) = d$. Then $f(M) = d \cdot \det M$ for all $M \in \mathbb{C}^{n\times n}$.*

Proof. Consider the map $f' = f - d \cdot \det$. Then f' is also multilinear alternating, and $f'(I_n) = 0$, so from the above discussion f' is identically 0. □

Proposition 11.7. *For any $n \times n$ matrices A, B, $\det(AB) = \det(A)\det(B)$.*

Proof. Define $D_B : \mathbb{C}^{n\times n} \to \mathbb{C}$ by

$$D_B \begin{bmatrix} R_1 \\ \vdots \\ R_n \end{bmatrix} = \det \begin{bmatrix} R_1 B \\ \vdots \\ R_n B \end{bmatrix}.$$

It is easy to check that D_B is mulitlinear alternating. Further, $D_B(I_n) = \det B$. Hence $D_B(M) = \det(M)\det(B)$. Putting $M = A$, we see that the matrix becomes AB, so $\det(AB) = D_B(A) = \det(A)\det(B)$, as required. □

12 Plücker Vector

We now study orbits of $SL_n(\mathbb{C})$ acting via left multiplication on $\mathbb{C}^{n\times r}$ for $r \geq n$.

Definition 12.1. *Let $m \in \mathbb{C}^{n\times r}$. The Plucker vector $p(m)$ of m is a vector of size $\binom{r}{n}$, indexed by n-tuples $\alpha = (i_1, i_2, \ldots i_n)$ such that $1 \leq i_1 < i_2 < \cdots < i_n \leq r$, whose α-entry $p(m)(\alpha)$ is the determinant of the submatrix of m determined by columns $(i_1, i_2, \ldots i_n)$.*

Example 12.2. *Take*

$$m = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}$$

Then $p(m)$ is given by:

	(1,2)	*(1,3)*	*(1,4)*	*(2,3)*	*(2,4)*	*(3,4)*
p(m)	*-4*	*-8*	*-12*	*-4*	*-8*	*-4*

Proposition 12.3. *If $m \in \mathbb{C}^{n\times r}$ has rank r, then $p(m) \neq 0$.*

Proof. Follows from definition of rank. □

Theorem 12.4. *Two matrices $m, m' \in \mathbb{C}^{n\times r}$ with rank n are in the same $SL_n(\mathbb{C})$ orbit iff $p(m) = p(m')$.*

Proof. First assume that $m' = rm$ for some $r \in SL_n(\mathbb{C})$. Then,

$$p(m')(\alpha) = \det(r) \cdot p(m)(\alpha) = p(m)(\alpha)$$

for all indices α, which proves $p(m') = p(m)$.

Now assume m, m' are in different orbits. We apply elementary row operations in such a way that whenever we scale any row by $a \neq 0$, we scale the last row by $\frac{1}{a}$. Thus we remain in the same SL_n orbit. Since p is invariant under SL_n action, we can assume m, m' are in *almost* row echelon form, where the last row may have a non-one pivot entry. Now we use induction on r to prove that some entry of $p(m)$ is different from $p(m')$. Note that if we take the n pivotal columns, we see that the non-zero entry in the last row must be the same for m, m'; otherwise we have found a differing entry.

Base case $r = n$: Both m, m' are diagonal, so their last row entries must be different. Therefore their determinants must be different, proving $p(m) \neq p(m')$.

Now assume $r > n$. If there are two columns with non-pivotal entries (i.e., when $r \geq n + 2$), then deleting one of the columns would still leave m, m' distinct, so by induction we can find a differing Plücker entry. Else $r = n + 1$, and m, m' are identical except for the non-pivotal column c. In that case, for every combination of columns that includes c, the determinant is a constant times an entry in c, namely the entry corresponding to the column that was not chosen. Since c differs for m and m', one of these determinants must be different, as required. □

Let $V = \mathbb{C}^{n\times r}$, and let V^0 be the set of matrices having rank n. Note that V^0 is open and dense in V. Thus, after throwing away a "thin" set $V \setminus V^0$, we can distinguish orbits in V using algebraic invariant p.

For distinguishing GL_n orbits, we can identify vectors in $\mathrm{Im}(p)$ under scaling (since any matrix in GL_n is just a scaled version of a matrix in SL_n). Thus $\mathrm{Im}(p)$ can be thought of as a subvariety of $\mathbb{P}^{\binom{r}{n}-1}$.

12.1 Algebraic Relations

The entries of a Plücker vector satisfy many algebraic relations of the following form:

Proposition 12.5. *If* $m = [i_1 \mid i_2 \mid \cdots \mid i_n \mid j_1 \mid \cdots \mid j_n \mid j_{n+1}] \in \mathbb{C}^{n \times 2n+1}$, *then*

$$\sum_{r=1}^{n+1} (-1)^r \det([i_1, i_2, \ldots i_{n-1}, j_r]) \det\Big([j_1, \ldots, \hat{j_r}, \ldots, j_{n+1}]\Big) = 0$$

Here $\hat{j_r}$ *means that the column* j_r *is dropped.*

Proof. Take the LHS of above as a function $f(j_1, j_2, \ldots j_{n+1})$. We can easily check that it is multilinear and alternating (as a function on columns), on $\mathbb{C}^{n\times(n+1)}$. Therefore it must be identically 0. □

Notes - Lecture 10

Scribe: Shantanu Nene

In this lecture, we will study Hilbert's one parameter subgroup theorem and a few of its consequences.

13 Recall

We study orbits of $n \times n$ matrices under action of SL_n via conjugation. These orbits can be distinguished by representatives, namely Jordan canonical forms. For any matrix A, we can write $V = \mathbb{C}^n$ as a sum of generalized eigenspaces $\bigoplus V_\lambda$ where

$$V_\lambda = \{v \mid (A - \lambda I)^k v = 0 \text{ for some } k \in \mathbb{N}\}$$

Corresponding to this basis, A takes the form of a block diagonal matrix with blocks J_i, such that each J_i has some (fixed) eigenvalue on its diagonal, 1s on its subdiagonal, and 0s everywhere else.

The number and size of Jordan blocks corresponding to eigenvalue λ is uniquely determined by the nullities of $(A - \lambda I)^k$ for $k = 1, 2, \ldots$. More specifically, if nullity of $(A - \lambda I)^k$ is e_k, then for each $m \geq 1$, the number of Jordan blocks corresponding to λ having size at least m is $e_m - e_{m-1}$.

14 One Parameter Subgroups

Definition 14.1. *A one parameter subgroup is a group homomorphism $\lambda : \mathbb{C}^* \mapsto SL_n(\mathbb{C})$ (where $SL_n(\mathbb{C})$ acts on $\mathbb{C}^n$ via conjugation), such that, under some suitable change of basis, $\lambda(t)$ is a diagonal matrix with the $(i,i)^{th}$ entry being t^{d_i} for some integer d_i.*

Note that since determinant of $\lambda(t)$ is 1, $d_1 + d_2 + \cdots d_n = 0$.

Example 14.2. *Take $\lambda(t) = \begin{bmatrix} t^{-1} & 0 \\ 0 & t \end{bmatrix}$ and $v = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$. Then $\lambda(t)v(t)\lambda(t)^{-1} = \begin{bmatrix} 1 & 0 \\ 2t^2 & 3 \end{bmatrix}$. Taking limit as $t \to 0$, we get the matrix $\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$, which lies in the same orbit as v.*

However, if we take $u = \begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}$, then $\lambda(t)u\lambda(t)^{-1} = \begin{bmatrix} 2 & 0 \\ t^2 & 2 \end{bmatrix}$. Now if we take $t \to 0$, we get the diagonal matrix $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, which lies in a different orbit than u because u is already in its Jordan form.

The following is a theorem of Hilbert, characterizing orbit closures:

Theorem 14.3. *Let $u, v \in \mathbb{C}^{n \times n}$. Consider the action of $SL_n(\mathbb{C})$ via conjugation. Then $u \in \overline{O(v)}$ iff there exists a $w \in O(v)$ and a one parameter subgroup λ such that*

$$u = \lim_{t \to 0} \lambda(t) w \lambda(t)^{-1}.$$

14.1 Closed Orbits in $\mathbb{C}^n$

Theorem 14.4. *The orbits of diagonalizable matrices are the only closed orbits in $\mathbb{C}^n$.*

Proof. If w is diagonal, then for any 1-p.s. λ, $\lambda(t)w\lambda(t)^{-1} = w$, which is in orbit of w. Therefore $O(w)$ is closed. Conversely, assume w is not diagonalizable. Then w has a Jordan canonical form, and WLOG assume $w_{1,2} = 1$ (entry at $(1,2)$. Then if $\lambda(t) = \text{diag}(t, 1, \ldots, 1)$, $u = \lambda(t)w\lambda(t)^{-1}$ has the same entries as w, except $u_{1,2} = t$. Then taking $t \to 0$, u tends to a matrix in Jordan form which is different from Jordan form of w, so is in a different orbit. Therefore $O(w)$ is not closed. □

15 Gordon's Problem

Let G act on V rationally. That is, $\rho : G \to GL(V)$ is a representation such that, if $g = (g_{i,j})_{i,j}$, then $\rho(g) = (\phi_{k,l}(g_{i,j}))_{k,l}$ where each $\phi_{k,l}$ is a rational function. The following are Gordon's problems:

1. What is the space of orbits?
2. Do sufficient number of invariants (to separate orbits) exist?
3. Is the ring of invariants finitely generated?

These questions form the basis for Hilbert's invariant theory.

16 Basic Algebraic Geometry

16.1 Elimination Theory

Proposition 16.1. *Let R be an integral domain and let $f = a_dX^d + \ldots + a_0$ and $g = b_eX^e + \ldots + b_0$ be two polynomials in $R[X]$, with $a_d, b_e \neq 0$. Let $ResMat(f, g)$ be the $(d + e) \times (d + e)$-matrix as shown in class and $Res_X(f, g)$ be its determinant. Then:*

1. *$Res_X(f, g) = 0$ iff f and g have a common factor in the ring $R[X]$.*
2. *If $a = Res_X(f, g) \neq 0$, then $a \in (f, g)$, the ideal formed by f, g in $R[X]$.*

Remark 16.2. *If f, g have a common factor, say $h(X)$, then $a_d^r b_e^s h(X) \in (f, g)$, for some integers $r, s \geq 0$.*

Proposition 16.3. *Let $R = \mathbb{C}[Y_1, \ldots, Y_k]$ and $f, g \in R[X]$ be two polynomials as above with $a_D(Y_1, \ldots, Y_k)$ and $b_e(Y_1, \ldots, Y_k)$ as its leading coefficients. Suppose that they do not have a common factor. Let $h(Y_1, \ldots, Y_k) = Res_X(f, g)$. If $y \in \mathbb{C}^k$ is such that $h(y) = 0$ but $a_d(y) \neq 0$, then there is an $x \in \mathbb{C}$ such that $f(y, x) = 0$ and $g(y, x) = 0$. In other words, $(y, x) \in Var(f, g)$. Conversely, (y, x) is a common root of f, g then x is also a root of h.*

Proposition 16.4. *Let $f_0, \ldots, f_r \in \mathbb{C}[X_0, \ldots, X_n]$. Let $U_1, \ldots, U_r$ be a set of indeterminates and $g = U_1f_1 + \ldots U_rf_r$ and $a_d(X_1, \ldots, X_n)$ be its leading coefficient of f_0. Let $h = Res_{X_0}(f_0, g)$ with $h \neq 0$. Suppose that $h = \sum_\alpha U^\alpha h_\alpha(X_1, \ldots, X_r)$ is the expression of h in terms of the monomials U^α in the U's. Let $I = (h_\alpha) \subseteq \mathbb{C}[X_1, \ldots, X_n]$, the ideal formed by these coefficient polynomials. Then:*

1. *Let J be the ideal generated by $f_0, \ldots, f_r$ in $\mathbb{C}[X_0, \ldots, X_n]$, then $I \subseteq J$. Indeed, each element h_α is an element of the ideal J.*
2. *Suppose that $\overline{x} = (x_1, \ldots, x_k) \in \mathbb{C}^k$ is in the variety of I such that $a_d(\overline{x}) \neq 0$. Then there is an $x_0 \in C$ such that $(x_0, \overline{x}) \in \mathbb{C}^{k+1}$ is in the variety of J.*
3. *Conversely, if $(x_0, \overline{x})$ is a common root of $f_0, \ldots, f_r$, then $\overline{x} \in Var(I)$, the variety of I.*

Remark 16.5. *If $Rex_{X_0}(f_0, g) = 0$ then $f_0, \ldots, f_r$ have a common factor.*

Proposition 16.6. *Let $f_1, \ldots, f_r \in \mathbb{C}[X_1, \ldots, X_n]$. Let $V(I) \subseteq \mathbb{C}^n$ be its variety. Let $\pi_k : \mathbb{C}^n \to \mathbb{C}^k$ be the projection $\mathbb{C}^n$ onto its first k coordinates. Let $W = \pi_k(V(I))$. Then the closure of W within $\mathbb{C}^k$ is an algebraic variety.*

Definition 16.7. *Let $I \subseteq \mathbb{C}[X_1, \ldots, X_n]$ and $x \in V(I)$. We define the normal space $N_xV(I)$ as the vector space generated by the vector $\{\boldsymbol{\nabla}(f)(x)$ where $f \in I\}$.*

Proposition 16.8. *Let $I \subseteq \mathbb{C}[X_1, \ldots, X_n]$ and let $d(x) = dim(N_xV(I))$. Then there is a number d such that (i) $d(x) \leq d$ for all $x \in V(I)$, and (ii) there is closed subset $W \subseteq V(I)$ such that $d(x) = d$ for all $x \in V - W$. Then $n - d$ is called the dimension of $V(I)$. If I is a prime ideal then $dim(I) = n - d$ is called the dimension of I. Points $x \in V$ where $d(x) = d$ are called smooth points.*

Remark 16.9. *If I is prime with $dim(I) = n - k$, then there are actually $f_1, \ldots, f_k \in I$ and a closed set W such that for all $x \in V - W$, $d(x) = dim(\mathbb{C} \cdot \{\boldsymbol{\nabla}(f_1)(x), \ldots, \boldsymbol{\nabla}(f_k)(x)\})$. Moreover $V(f_1, \ldots, f_k) - W = V - W$.*

Corollary 16.10. *Let $V \subseteq V'$ be irreducible varieties with ideals I and I' and let x be a smooth point in V and V'. Then $d_V(x) \geq d_{V'}(x)$.*

Definition 16.11. *Let $V(I)$ be a variety in $\mathbb{C}^n$ of dimension k and let x be a smooth point. Let $N_xV(I)$ be given by the linear span of $\boldsymbol{\nabla}(f_1)(x), \ldots, \boldsymbol{\nabla}(f_{n-k})(x)$ treated as row vectors. Then the tangent space $T_xV(I)$ is the space of column vectors $v \in \mathbb{C}^n$ such that $wv = 0$ for all $w \in N_xV(I)$. The dimension of $T_xV(I)$ is then k.*

Proposition 16.12. *Let V be an irreducible variety of dimension d and $x \in V$ be a smooth point.*

16.2 Algebraic groups and orbits

Our primary concern will be algebraic subgroups $G \subseteq GL_n$. We will take $GL_n(\mathbb{C})$ as the mother of all groups. We take the matrix space $M = \mathbb{C}^{n\times n}$ and coordinate functions $(M_{ij})_{i,j=1}^n \in M^*$.

Let us first see how GL_n itself is an algebraic variety. Now, the group GL_n is the complement of the hypersurface $det((M_{ij})) = 0$, where $det((M_{ij}))$ is the determinant function expressed as $\sum_{\sigma\in\mathbb{S}_n}(-1)^{sign\sigma}\prod_{i=1}^n M_{i,\sigma i}$.

By standard algebraic geometry, the ring of regular functions $\mathbb{C}[GL_n]$ is obtained by adjoining an indeterminate Y to $\mathbb{C}[(M_{ij})]$ and going modulo the function $Y - det(M)$. Thus $\mathbb{C}[GL_n] \cong \mathbb{C}[(M_{ij}), Y]/(Y - det(M))$. Other groups $G \subseteq GL_n$ will be given by additional algebraic equations and hence will be closed subvarieties of GL_n. Thus $\mathbb{C}[G] \cong \mathbb{C}[GL_n]/I_G$, where $I_G \subseteq \mathbb{C}[GL_n]$ is a suitable ideal.

Definition 16.13. *An algebraic group $G \subseteq GL_n = GL_n(\mathbb{C})$ is a closed subvariety of GL_n, as above. Moreover, the multiplication map $G \times G \to G$ is an algebraic map.*

Remark 16.14. *Let G be a group of dimension r, then $dim(T_gG) = r$ for all $g \in G$.*

Let us now develop the algebraic geometry of a rational group action $\rho : G \times V \to V$. For V, we have $(V_i)_{i=1}^n \in V^*$, a basis of linear functions on V. Thus $\mathbb{C}[V] \cong \mathbb{C}[(V_i)]$, the polynomial algebra generated by the symbols (V_i).

Next, let us look at the LHS of ρ, viz., $G \times V$. By standard algebraic geometry, $\mathbb{C}[G \times V] \cong \mathbb{C}[G] \otimes_{\mathbb{C}} \mathbb{C}[V]$, the tensor product algebra.

Then, the map $\rho : G \times V \to V$ gives us a map $\rho^* : \mathbb{C}[V] \to \mathbb{C}[G] \otimes \mathbb{C}[V]$. We thus have for any $q \in \mathbb{C}[V]$:

$$\rho^*(q) = \sum_{j=1}^{K} p_j(M) q_j(V)$$

where $p_{ij} \in \mathbb{C}[G]$ and $q_j \in \mathbb{C}[V]$. Thus we have the important proposition:

Proposition 16.15. *Recall the action of G on $\mathbb{C}[V]$, where if $q \in \mathbb{C}[V]$ and $g \in G$, then we define $g \cdot q$ or simply $f^G \in \mathbb{C}[V]$ as the function:*

$$q^g(v) = q(g^{-1}v)$$

Then there are functions $q_1, \ldots, q_k \in \mathbb{C}[V]$ and $p_1, \ldots, p_k \in \mathbb{C}[G]$ such that for all g, we have:

$$q^g = \sum_{i=1}^{k} p_i(g) q_i$$

Remark 16.16. *We will see later that for* reductive *groups the vectors space M formed by the set $\{q_1, \ldots, q_k\}$ may itself be closed under the action of G, or in other words, a G-module.*

Given that ρ is linear on V, gives us the special case:

$$\rho^*(V_i) = \sum_{j=1}^{N} p_{ij}(M) V_j$$

We may write this in the matrix form:

$$\begin{bmatrix} \rho^*(V_1) \\ \vdots \\ \rho^*(V_n) \end{bmatrix} = \begin{bmatrix} p_{11}(M) & \ldots & p_{1N}(M) \\ \vdots & & \vdots \\ p_{N1}(M) & \ldots & p_{NN}(M) \end{bmatrix} \begin{bmatrix} V_1 \\ \vdots \\ V_n \end{bmatrix}$$

It is important to understand the above equation. If $v \in V$ and $g \in G$ such that $w = g \cdot v$, then we have the concrete matrix product:

$$\begin{bmatrix} V_1(w) \\ \vdots \\ V_N(w) \end{bmatrix} = \begin{bmatrix} p_{11}(g) & \ldots & p_{1N}(g) \\ \vdots & & \vdots \\ p_{N1}(g) & \ldots & p_{NN}(g) \end{bmatrix} \begin{bmatrix} V_1(v) \\ \vdots \\ V_N(v) \end{bmatrix}$$

Example 16.17. *$Sym^2(\mathbb{C}^2)$.*

Thus, the orbit of v is better seen as the **graph** of the group action on v, i.e., the set $\{(g, w) | w = g.v \text{ } and \text{ } g \in G\} \subseteq G \times V$ and the projection from $\pi : G \times V \to V$ with $\pi(g, x) = x$. Thus, if we use a fresh set of coordinates for V, viz $W_1, \ldots, W_N$, then the graph Y of v is given by the equations:

$$\begin{bmatrix} W_1 \\ \vdots \\ W_N \end{bmatrix} = \begin{bmatrix} p_{11}(M) & \ldots & p_{1N}(M) \\ \vdots & & \vdots \\ p_{N1}(M) & \ldots & p_{NN}(M) \end{bmatrix} \begin{bmatrix} V_1(v) \\ \vdots \\ V_N(v) \end{bmatrix}$$

These are N equations plus the equations needed to define I_G in $n^2 + N$ variables (M_{ij}) and (W_k). Then $O(v) = \pi(Y) \subseteq V$. Therefore, we have:

Corollary 16.18. *Let G be an algebraic group and V be a G-module via a rational map $\rho : G \to GL(V)$. Let $v \in V$. Then the closure of orbit $O(v)$ within V is an algebraic variety. The stabilizer G_v of v is an algebraic group.*

Proof: The first assertion follows from Prop. 16.6. For the second, we consider the algebraic subset of G defined by the equations:

$$\begin{bmatrix} V_1(v) \\ \vdots \\ V_N(v) \end{bmatrix} = \begin{bmatrix} p_{11}(g) & \ldots & p_{1N}(g) \\ \vdots & & \vdots \\ p_{N1}(g) & \ldots & p_{NN}(g) \end{bmatrix} \begin{bmatrix} V_1(v) \\ \vdots \\ V_N(v) \end{bmatrix}$$

16.3 Lie Algebras

Definition 16.19. *Let $G \subseteq GL_n$ be an algebraic group. The Lie algebra of G, denoted by $\mathcal{G}$ of $Lie(G)$ is the tangent space at the identity element I. In other words, $\mathcal{G} = T_I G$, the space of all $X \in M_n$ such that $I + \epsilon X \in G$, when $\epsilon^2 = 0$. $\mathcal{G}$ is closed under the lie bracket: if $X_1, X_2 \in \mathcal{G}$, then so is $[X_1, X_2] = X_1X_2 - X_2X_1$.*

Example 16.20. *SL_n and O_n.*

Proposition 16.21. *G acts on $\mathcal{G}$ via the* **adjoint** *representation. If $g \in G$ and $\mathfrak{g} \in \mathcal{G}$, then $adj(g)(\mathfrak{g}) = g\mathfrak{g}g^{-1}$, where the multiplication happens in GL_n.*

Lemma 16.22. *If $G \subseteq GL(X)$ is a group, then $dim(G) = dim(Lie(G))$. Moreover, if H is a subgroup of G, then $Lie(H) \subseteq Lie(G)$ and $dim(Lie(H)) = dim(H)$.*

Proposition 16.23. *For any group $G \subseteq GL(X)$, with Lie alegbra $\mathcal{G}$, there is an open neighborhood $O \subseteq \mathcal{G}$ around 0, and map $exp : O \to G$ which is a diffeomorphism on its image. Moreover, for any $X \in \mathcal{G}$, the image of the line tX, denoted by $\gamma_X(t) \subseteq G$, is a curve such that $\gamma'(0) = X$.*

Proposition 16.24. *Consider the map $exp : M_n \to GL_n$, given by $exp(X) = e^X$. Then exp has the following properties:*

1. *$exp(X)exp(-X) = I$, the identity. For any $A \in GL_n$, we have $Aexp(X)A^{-1} = exp(AXA^{-1})$.*
2. *If the eigenvalues of X are λ_1, λ_r with multiplicities $d_1, \ldots, d_r$, then $e^{\lambda_1}, \ldots, e^{\lambda_r}$ are the eigenvalues of e^X with the same multiplicities.*
3. *If X, Y commute, then $exp(X + Y) = e^X e^Y$.*

Proposition 16.25. *G acts on $\mathcal{G}$ via the* **adjoint** *representation. If $g \in G$ and $\mathfrak{g} \in \mathcal{G}$, then $adj(g)(\mathfrak{g}) = g\mathfrak{g}g^{-1}$, where the multiplication happens in GL_n.*

Example 16.26. *Exponential map for SL_n and O_n.*

Now, given the map $\rho : G \times V \to V$, we may pick points g, v and $w = \rho(g)(v)$ or simply gv and have the tangential map:

$$(\rho_*)_{g,v} : T_g G \times T_v V \to T_w V$$

More specifically, at $g = I$, the identity matrix, we have $T_I(G) = \mathcal{G}, v = w$ and $T_v \cong V$. Therefore, we have a map, which we denote by ρ_1:

$$\rho_1 : \mathcal{G} \times V \to V$$

This we call the *Lie algebra representation V*.

Definition 16.27. *For a Lie algebra $\mathcal{G}$, we say that V is a $\mathcal{G}$-representation of a $\mathcal{G}$-module, if we have a map $\rho_1 : \mathcal{G} \to End(V)$ such that for any $X_1, X_2 \in \mathcal{G}$, we have $\rho_1([X_1, X_2]) = [\rho_1(X_1), \rho_1(X_2)]$.*

Note that the Lie bracket on the right happens in $End(V)$ space.

Proposition 16.28. *Let $\rho : G \times V \to V$ be a representation. Let $\rho_1 : \mathcal{G} \to End(V)$ its tangent map. Then ρ_1 is a Lie algebra representation. Moreover, for any $g \in G$ and $\mathfrak{g} \in \mathcal{G}$:*

$$\rho(g)\rho_1(\mathfrak{g})\rho(g)^{-1} = \rho_1(adj(g) \cdot \mathfrak{g}) = \rho_1(g\mathfrak{g}g^{-1})$$

Finally, for a point v with stabilizer G_v, $Lie(G_v) = \{\mathfrak{g} \in \mathcal{G} | \rho_1(\mathfrak{g})(v) = 0\}$.

Note that $\rho_1 : \mathcal{G} \to End(V)$ is a linear map! Thus the Lie algebra of the stabilizer of a point v is a linear algebra computation, once representation ρ_1 is computed. How is $\rho_1(X)$ to be computed? The simplest is to use the matrix for ρ and differentiate the matrix $\rho(g)$ in the direction X. We will illustrate with two examples.

Example 16.29. *Let us look at the form $f = Ax^2 + Bxy + Cy^2$ under the infinitesimal Lie element:*

$$I + \epsilon \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 + \epsilon a & b \\ c & 1 + \epsilon d \end{bmatrix}$$

Its action on f is given by:

$$\begin{array}{rl} & A(x + \epsilon ax + \epsilon by)^2 + B(x + \epsilon ax + \epsilon by)(y + \epsilon cx + \epsilon dy) + C(y + \epsilon cx + \epsilon dy)^2 \\ = & (Ax^2 + Bxy + Cy^2) + \epsilon\{A(2ax^2 + 2bxy) + B(cx^2 + dxy + axy + by^2) + C(2cxy + 2dy^2) \\ = & \begin{bmatrix} x^2 & xy & y^2 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} + \epsilon \begin{bmatrix} x^2 & xy & y^2 \end{bmatrix} \begin{bmatrix} 2a & c & 0 \\ 2b & a + d & 2c \\ 0 & b & 2d \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} \end{array}$$

Compute the stabilizers of xy, x^2 and x^2+y^2. What is the tangent space at the point f of the orbit $O(f)$? This is done by taking a basis of $Lie(GL_2)$ and hitting it on f. This is obtained by putting $a=1, b=1, c=1$ and $d=1$ in turn. We get the 4 vectors:

$$\begin{bmatrix} 2A \\ B \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 2A \\ B \end{bmatrix} \begin{bmatrix} B \\ 2C \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ B \\ 2C \end{bmatrix}$$

This can be succintly written as the column space of the matrix below:

$$\begin{bmatrix} 2A & 0 & B & 0 \\ B & 2A & 2C & B \\ 0 & B & 0 & 2C \end{bmatrix}$$

So what are the forms $f \in Sym^2(\mathbb{C}^2)$ whose orbits have dimension ≤ 2? That is given by the condition that all 3×3-minors of the matrix should have determinant 0.

Example 16.30. *$Sym^3(\mathbb{C}^3)$ and adjoint action. Stabilizers.*

Corollary 16.31. *We have $dim(G) = dim(G_v) + dim(O(v))$. Moreover, the space $\{v | dim(G_v) \geq k\}$ is an algebraic subvariety of V.*

17 Groups, Reductivity and Invariant Theory

This section looks at finite-dimensional G-modules and the categorical structures which it allows. They form the basic building block for the First Theorem of Geometric Invariant Theory, that for reductive groups acting on a finiste dimensional vector space V, (i) disjoing closed G-invariant algebraic subsets can be separated by invariants, and (ii) the ring of invariants is a finitely generated algebra. The basic result which connects the two is:

Proposition 17.1. *Let $\rho : G \to GL(V)$ be a rational representation. Let $f \in \mathbb{C}[V]$ be any function. Then there is a finite dimensional G-invariant subspace $F \subseteq \mathbb{C}[V]$ such that $f \in F$ And so are all its translates f^g, for every $g \in G$.*

Proof: For any f as above, we have:

$$\rho^*(f) = \sum_i p_i q_i$$

where $p_i \in \mathbb{C}[G]$ and $q_i \in \mathbb{C}[V]$. Moreover, we have, for any $g \in G$ and $v \in V$:

$$f(gv) = \sum_i p_i(g) q_i(v)$$

Let Q be the finite dimensional space formed by (q_i). Then the above expression tells us that $f^g \in Q$. Let $F \subseteq Q$ be the space generated by all f^g as g ranges over G. Then F is the required space. □

Remark 17.2. *When the group action is linear, as in our case, we have $\mathbb{C}[V] = \oplus_d \mathbb{C}[V]_d$ and in fact, each $\mathbb{C}[V]_d$ is itself a G-module.*

Definition 17.3. *We say that an algebraic group G is reductive if for every finite dimensional representation V and aG-invariant subspace $W \subseteq V$, we have a G-invariant complement, i.e., a $W' \subseteq V$ which is itself G-invariant such that $W \oplus W' = V$.*

Lemma 17.4. *If $\rho : G \to V$ and $\rho' : G \to V'$ are two representations, then so is $Hom(V, W)$, the space of homomorphisms from V to V'. For any $\phi : V \to V'$, the action of $g \in G$ on ϕ is given by $\phi^g = \rho'(g) \circ \phi \circ \rho(g^{-1})$. Note that $\phi^g(v) = g \cdot \phi(g^{-1} \cdot v)$. An element ϕ is called a G-morphism if $\phi^g = \phi$, or equivalently for all $v \in V$, $\phi(gv) = g\phi(v)$.*

Lemma 17.5. *Let V, W be G-modules and $\phi : V \to W$ be a G-morphism. Then $ker(\phi)$ and $Im(\phi)$ are G-modules.*

Proposition 17.6. *Finite groups are reductive.*

Proof: Let V and $w \subseteq V$ be G-modules. Let $\phi : V \to W$ be a surjection such that $\phi(w) = w$ for all $\in W$. Let $\overline{\phi} = \frac{1}{|G|} \sum_{g \in G} \phi^g$. Then $\overline{\phi}$ is G-invariant. The kernel of $\overline{\phi}$ is the required complement. □

Proposition 17.7. *The group $G = (\mathbb{C}^*)$, also called the k-dimensional torus, is reductive. If $\chi = (n_1, \ldots, n_k) \in \mathbb{Z}^k$ and $\bar{t} = (t_1, \ldots, t_k) \in G$ then $\bar{t}^\chi = \prod_i t_i^{n_i}$. If V is any finite dimensional module, then $V = \oplus_\chi V_\chi$ where, for any $v \in V_\chi$, we have $\bar{t}v = \bar{t}^\chi v$.*

Corollary 17.8. *The group $(\mathbb{C}^*)^k$ is reductive.*

We assume henceforth that G is reductive. For general reductive groups, we build the averaging operator as follows.

Lemma 17.9. *Let W be a G-module. Let W' be the vector space generated by all elements $\{w - gw | w \in W, g \in G\}$, then W' is G-invariant. Moreover, there is a unique G-complement $W^G = \{w \in W | gw = w\}$.*

Proof: Suppose that $dim(W) = n$ and $dim(W') = r$. It is easy to check that W' is G-closed and thus there is a G-complement W''. For a $w'' \in W''$, we see that $gw'' = w''$ for all g. This is simply because $w'' - gw'' \in W'' \cap W' = 0$. Thus $W'' \subseteq W^G$ and $dim(W^G) \geq n - r$. Conversely, if $\pi : W \to W^G$ is a projection then for any w, g, $\pi(w - gw) = \pi(w) - \pi(gw) = 0$. Thus $W' \subseteq (W^G)'$, and $r \leq n - dim(W^G)$.

Proposition 17.10. *Let W be a G-module and W^G be all vectors $w \in W$ such that $gw = w$ for all $g \in G$. Then there is a canonical projection $\pi^G : W \to W^G$.*

Lemma 17.11. *Let V, W be G-modules such that $\phi : V \to W$ is G-equivariant and surjective. Then $\phi : V^G \to W^G$ is also surjective. Moreover $(V \oplus W)^G = V^G \oplus W^G$.*

Corollary 17.12. *For any G-module W and $w \in W$, let X be the G-module generated by w. Then we have $W = X \oplus Y$ and $W^G = X^G \oplus Y^G$. Moreover, $dim(X^G)$ is either 0 or 1. Thus $\pi^G(w)$ is decided by X and X^G and not by W.*

Note that $w - \pi(w) \in X'$ and so $X = X' + \pi(w)$.

Proposition 17.13. *Consider the algebra $\mathbb{C}[V]^G \subseteq \mathbb{C}[V]$. Then there is a canonical projection $\pi : \mathbb{C}[V] \to \mathbb{C}[V]^G$. Moreover, it preserves degree. Further, if $f \in \mathbb{C}[V]^G_d$ and $f' \in \mathbb{C}[V]_e$, the $\pi(ff') = f\pi(f')$.*

Warning: Note that π is NOT an algebra homomorphism.
Proof: The only part which needs proof is the second part that $\pi(ff') = f\pi(f')$ when $f \in \mathbb{C}[V]^G$. We may easily reduce to the case when f and f' are homogeneous. Let $M \subseteq \mathbb{C}[V]$ be the G-module generated by f'. Suppose now that $deg(f) = d$ and $deg(f') = e$. Consider the module $M' = \mathbb{C}f \otimes M \subseteq \mathbb{C}[V]_{d+e}$. It is clear that $M'^G = f \otimes M^G$. □

Proposition 17.14. *The ring $\mathbb{C}[V]^G$ is finitely generated.*

Proposition 17.15. *Let $O_1, O_2 \subseteq V$ be closed varieties such that $V_1 \cap V_2 = \phi$. Then there are two invariants $f_1, f_2 \in \mathbb{C}[V]^G$ such that $f_i \in I(V_i)$ and $f_1 + f_2 = 1$.*

We now come to some results on the structure of orbits.

Lemma 17.16.
1. *Let $v \in V$ and $O(v)$ be the G-orbit of V and $\overline{O(v)}$ be its closure. Let $w \in \overline{O(v)}$, then there is no invariant which separates v from w.*
2. *If $v' \in V$ is such that $\overline{O(v')} \cap \overline{O(v)} = \phi$ then there is an invariant $f \in \mathbb{C}[V]^G$ which separates them.*

Proposition 17.17.
1. *Let $v \in V$ and $O(v)$ be the G-orbit of V and $\overline{O(v)}$ be its closure. Then there is a unique orbit $O(w)$ of minimum dimension, and it is closed.*
2. *Let us define $\sim$ on V as $v \sim v'$ is $\overline{O(v)} \cap \overline{O(v')} \neq \phi$. Then $\sim$ is an equivalence relation. Moreover, $[v]_\sim$ is closed. If $v' \notin [v]_\sim$ then there is an invariant which separates them.*

Definition 17.18. *A $v \in V$ is called stable if its G-orbit is closed.*

Warning: The presence of closed orbits may be very different for a $GL(X)$-module V when restricted to $SL(X) \subset GL(X)$.

Theorem 17.19. *Suppose G is reductive and $\lambda(t) \subseteq G$ is a 1-PS. Let $\lambda(t)v = \sum_i t^i v_i$. Then if $v_i = 0$ for all < 0 then v is not stable. Conversely, if v is not stable, then there is a 1-PS $\lambda(t)$ and a $w \in V$ such that $\lim_{t \to 0} \lambda(t)v = w$. Note that $w = 0$ is also permitted.*

The converse part of the theorem is the Hilbert-Mumford theorem. Note that proving something as unstable is shown by a suitable $\lambda(t)$. However, very few techniques exist to prove stability.

18 Solvable Lie Algebras

Definition 18.1. *Let $\mathfrak{g}$ be a Lie algebra. Define a decreasing sequence of ideals $D_i(\mathfrak{g})$, called the derived series of $\mathfrak{g}$ by: $D_0(\mathfrak{g}) = \mathfrak{g}$ and $D_{i+1}(\mathfrak{g}) = [D_i(\mathfrak{g}), D_i(\mathfrak{g})]$ for all $i \geq 0$. Then, $\mathfrak{g}$ is called solvable if $D_n(\mathfrak{g}) = 0$ for some n.*

Theorem 18.2 (Cartan's Criterion for Solvability)**.** *Let V be a finite dimensional vector space and $\mathfrak{g}$ be a Lie subalgebra of $GL(V)$. Suppose for all $x \in [\mathfrak{g}, \mathfrak{g}]$ and $y \in \mathfrak{g}$, $Tr(xy) = 0$. Then $\mathfrak{g}$ is solvable.*

19 Semisimple Lie Algebras

Definition 19.1. *A Lie algebra $\mathfrak{g}$ is called simple if it is non-abelian and has no proper non-zero ideals.*

Definition 19.2. *A Lie algebra $\mathfrak{g}$ is called semisimple if it has no non-zero abelian ideal.*

Definition 19.3. *For a Lie algebra $\mathfrak{g}$, the radical $Rad(\mathfrak{g})$ is defined to be its greatest solvable ideal.*

Lemma 19.4. *A non-zero Lie algebra $\mathfrak{g}$ is semisimple iff $Rad(\mathfrak{g}) = 0$.*

Proof. By definition, any abelian ideal of $\mathfrak{g}$ is contained in $\mathrm{Rad}(\mathfrak{g})$, so the condition is necessary. Conversely, if $\mathrm{Rad}(\mathfrak{g})$ is non-zero, then the last non-zero ideal in its derived series is the required abelian ideal, making $\mathfrak{g}$ not semisimple. □

20 Structure of Semisimple Lie Algebras

Theorem 20.1. *Let $\mathfrak{g}$ be a non-zero semisimple Lie algebra. Then*

1. *There exists $t \in \mathbb{N}$ and ideals $s_1, \ldots, s_t \in \mathfrak{g}$ that are simple as Lie algebras, such that*

$$\mathfrak{g} = \bigoplus_{i=1}^{t} s_i$$

2. *If s is any ideal of $\mathfrak{g}$ that is a simple Lie algebra, then $s = s_i$ for some i.*

Corollary 20.2. *If $\mathfrak{g}$ is semisimple Lie algebra, then $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$.*

21 Complete reducibility of Finite Dimensional Representations

21.1 The Casimir Element

Lemma 21.1. *Let $\mathfrak{g}$ be a finite dimensional semisimple Lie algebra, and let (V, f) be a faithful finite dimensional representation of $\mathfrak{g}$. Then the map $\beta_f(x, y) = Tr(f(x)f(y))$ is an invariant, non-degenerate and symmetric bilinear form on $\mathfrak{g}$.*

Lemma 21.2. *Let $\mathfrak{g}$ be a finite dimensional semisimple Lie algebra, β be an invariant, non-degenerate and symmetric bilinear form on $\mathfrak{g}$, and let (V, f) be a representation of $\mathfrak{g}$. Let $(x_1, \ldots, x_n)$ be a basis of $\mathfrak{g}$ and let $(y_1, \ldots, y_n)$ be its dual basis wrt β. Then, the element*

$$c = \sum_{i=1}^{n} f(x_i)f(y_i)$$

is a $\mathfrak{g}$-invariant endomorphism of (V, f).

Proof. Direct calculation □

Such a c is called the Casimir element. Observe that $\mathrm{Tr}(c) = \dim(\mathfrak{g}) \neq 0$.

21.2 Weyl's Theorem

Lemma 21.3. *Let $\mathfrak{g}$ be a finite dimensional semisimple Lie algebra, and let (V, f) be a finite dimensional representation of $\mathfrak{g}$. If W is a codimension 1 subrepresentation of (V, f), then there exists a subrepresentation X of (V, f) such that $V = W \oplus X$.*

Proof. We proceed by induction on $\dim(V)$. If $\dim(V) = 1$, the claim is trivial. So now assume $\dim(V) > 1$, and let W be a subrepresentation of V with codimension 1. Note that, by going modulo te kernel of f, we may assume (V, f) is a faithful representation.

First assume W is simple. Consider the Casimir element c; it is a $\mathfrak{g}$-invariant endomorphism of V. Since $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$, $\mathfrak{g}$ acts trivially on one-dimensional representations; in particular on V/W. Therefore $f(V) \subset W$, and so $c(V) \subset W$. Further, by Schur's lemma, c is a scalar on W, say multiplication by $\lambda \in \mathbb{C}$. Since $\mathrm{Tr}(c) \neq 0$, $\lambda \neq 0$, and so $c(V) = W$ and c is an isomorphism on W. Therefore $\ker(c)$ is a one-dimensional subrepresentation of V that intersects W only at 0, and so is a direct summand of W. □

Theorem 21.4. *Let $\mathfrak{g}$ be a finite dimensional semisimple Lie algebra. Then any finite dimensional representation of $\mathfrak{g}$ is completely reducible. More strongly, any subrepresentation is a direct summand.*

Proof. Let (V, ρ) be a finite dimensional representation. If it is simple, we are done. Else, let W be a proper non-zero subrepresentation of V. Consider $\mathrm{Hom}(V, W)$ as a representation of $\mathfrak{g}$, via the map μ given by

$$(\mu(x)(f))(w) = (\rho(x) \odot f)(w) - (f \odot \rho(x))(w)$$

. Let $\mathcal{V}$ and $\mathcal{W}$ be subspaces of $\mathrm{Hom}(V, w)$ consisting of maps that are scalar and 0 respectively on W. Then for any $f \in \mathcal{V}$,$x \in \mathfrak{g}$ and $w \in W$, by definition $(\mu(x)(f))(w) = 0$. Therefore $\mathcal{V}, \mathcal{W}$ are subrepresentations of $\mathrm{Hom}(V, W)$, and $\mathfrak{g}$ sends $\mathcal{V}$ to $\mathcal{W}$. Also from definition, $\mathcal{W}$ is a codimension 1 subspace of $\mathcal{V}$.

Now, applying the above lemma, there exists a one dimensional representation of $\mathcal{V}$ that is a direct summand of $\mathcal{W}$. Say the representation is generated by f; WLOG $f|_W = id_W$. Then since $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$, $\mathfrak{g}$ acts trivially on $\mathbb{C}f$. Therefore $\ker(f)$ is a subrepresentation of V, and since f is identity on W, $\ker(f) \cap W = (0)$. Therefore by a dimension argument, $\ker(f)$ is a direct summand of W. □

22 Semisimple Groups and Algebras

Definition 22.1. *An algebraic group G is called semisimple if it has no closed connected normal subgroup except e.*

Theorem 22.2. *A connected algebraic group G is semisimple iff $\mathfrak{g}$ is semisimple.*

23 Hilbert-Mumford Theorem

The following is a fairly elementary proof of the Hilbert-Mumford theorem, which uses the following lemma:

Lemma 23.1. *Let $m_{i,j}$ for $1 \leq i \leq r$ and $1 \leq j \leq n$ be integers satisfying the following property: If $b_1, \dots, b_r$ are integers, not all zero, such that*

$$\sum_{i=1}^{r} b_i m_{i,j} = 0 \quad \forall 1 \leq j \leq n,$$

then some two of the b_i have opposite sign. Then, there exist integers c_j such that

$$\sum_{j=1}^{n} m_{i,j} c_j > 0 \quad \forall 1 \leq i \leq r.$$

Proof. It is enough to prove the statement for rationals c_i, and just scale later. Also, the given condition holds for rationals b_i, by scaling. It translates to the kernel of the linear map

$$M : (b_1, \dots, b_r) \to (\sum_{i=1}^{r} b_i m_{i,1}, \dots, \sum_{i=1}^{r} b_i m_{i,n})$$

intersecting the cone of non-negative entries in $\mathbb{Q}^r$ only at the origin. Thus, due to density of $\mathbb{Q}$ in $\mathbb{R}$, it intersects non-negative cone of $\mathbb{R}^r$ only at the origin. Consider the map M^t. Note that $\ker(M)$ and $\mathrm{Im}(M^t)$ are orthogonal complements in $\mathbb{R}^r$.

We show the following: If K is a subspace of $\mathbb{R}^r$ that intersects the non-negative cone only at origin, then $K^{\perp}$ intersects the interior of the cone. Suppose K has co-dimension $k \geq 2$. Consider the image of the non-negative cone in the projection $\mathbb{R}^r/K \cong \mathbb{R}^k$; call it D. Clearly D is closed. Further, $\mathbb{R}^k \setminus \{0\}$ is connected since $k \geq 2$, so D and $-D$ cannot cover $\mathbb{R}^k \setminus \{0\}$, so there is a non-zero vector $v \in \mathbb{R}^k$ such that $\mathbb{R}v \cap D \setminus \{0\} = \varnothing$. Thus if we add the pull-back of v to K, we get a subspace of one higher dimension, which also intersects the non-negative cone only at 0. Keep doing this until we get K having codimension 1. Then, suppose K is the hyperplane $\sum_{i=1}^{r} \lambda_i x_i = 0$. Since K intersects non-negative cone only at the origin, all the λ_i must be non-zero and have the same sign, WLOG all positive. Then $(\lambda_1, \dots, \lambda_r) \in K^{\perp}$ is a vector that is in the interior of the cone, as required.

Applying the above property to $\ker(M)$, we get that some vector in $\mathrm{Im}(M^t)$ lies in interior of the non-negative cone. By scaling and using density of rationals, we can assume this vector $(c_1, \dots, c_n)$ to have integer entries. This vector satisfies the precise condition that we want. □

The Hilbert-Mumford theorem gives a nice characterization of unstable points in terms of one parameter subgroups.

Theorem 23.2. *Let G be a reductive group acting linearly on a vector space V. Let $v \in V$ be G-unstable, i.e., the closure of the orbit $\overline{G \cdot v}$ contains 0. Then there exists a one-parameter subgroup $\lambda : \mathbb{C}^* \to G$ such that $\lim_{t \to 0} \lambda(t) \cdot v = 0$.*

Before we prove this theorem, we would need another lemma relating unstable points in G to unstable points in the maximal torus T:

Lemma 23.3. *Let $T \leq G$ be a maximal torus in G, and suppose $v \in V$ is G-unstable. Then the orbit $G \cdot v$ contains a vector u which is T-unstable.*

Proof. We skip the proof here, but the key idea is: In a reductive group, the set of elements conjugate to the maximal torus form an open dense subset of G. □

Now we can finally prove the theorem.

Proof. Moving to the torus, we can assume v is T-unstable. Hence V can be split into different weight classes. More concretely, we can write any $v \in V$ as $\sum_{i=1}^{r} v_i$ where T acts by scalar multiplication on each v_i. Suppose

$$(t_1, t_2, \ldots, t_n) \cdot v_i = t_1^{m_{i,1}} \cdots t_n^{m_{i,n}} v_i.$$

for all $1 \leq i \leq r$ and $(t_1, \ldots, t_n) \in T$. □

24 Exercises

1. Use Valiant's construction to construct a matrix M whose determinant is $b^2 - 4ac$.

2. Consider the unit cube with 8 vertices V, 12 edges E and 6 faces F. Let G be the group of its symmetries. List the elements of G (by their action on the faces) as a product $S_1 \times S_2$ with S_1 as the stabilizer of a face and S_2 as a list of coset representatives.

3. For the action of G above, what is the number of orbits for its action on $F \times F \times F$? How would it change if we add the additional commuting action of S_3, the symmetric group?

4. Recall the cycle description of a single bijection f on a finite set S. We would like to list all bijections g which commute with f. Use the cycle description of f to do this. List this set for $f = (123)(45)(67)$.

5. Consider the group Z_6 acting on $\mathbb{C}^6$ by circular shift. Let $v = [1, -1, 0, 0, 0, 0]$. Find the smallest subspace $M \subseteq \mathbb{C}^6$ such that M is Z_6-invariant. Consider $\mathbb{Z}_3 \subseteq \mathbb{Z}_6$. Split M into $\mathbb{Z}_3$ invariant subspaces.

6. Let $x_1, \ldots, x_n$ be indeterminates and let $G = S_n$ act by permutations. Let $s_1, \ldots, s_n$ be the coefficients of $\prod_i (X - x_i)$. Show that any polynomial $p \in \mathbb{C}[x_1, \ldots, x_n]$ which is symmetric, .i.e., $p^G = p$, is a polynomial in $s_1, \ldots, s_n$.

7. Recall the Reynolds operator Π^G from $R \to R^G$. For the ring $\mathbb{C}[x_1, \ldots, x_n]$ above, show by first principles that $\Pi^G(ff') = f\Pi^G(f')$, if $f^G = f$. Exhibit an element f' where $\Pi^G(f') = 0$.

8. Let $G = S_3$. Then there are only three non-isomorphic and irreducible G-modules M_3 of dimension 1, M_{21} of dimension 2 and M_{111} of dimension 1. The first is generated by v_3 with $\sigma v = v$. M_{111} is generated by v_{111} such that $\sigma v_{111} = (-1)^{sign(\sigma)} v_{111}$. Let $M_{21} = \{w \in \mathbb{C}^3 | w(1) + w(2) + w(3) = 0\|$. Then M_{21} is this third module. Consider $V = \mathbb{C}[x_1, x_2, x_3]_2$, polynomials of degree 2 under the action of G on the variables. Decompose V into sum of subspaces $n_3 M_3 \oplus m_{21} M_{21} \oplus n_{111} M_{111}$.

9. Consider the surface S given by the equation $f \equiv x^2 + y^2 - z^2 = 0$. Let $p = (0, 0, 0)$ and $q = (0, 1, 1)$ on S. Compute the normal space and the tangent space at p and q. Which of these points is smooth?

10. Consider the curve defined by f and $g \equiv x^2 + z^2 - y^2 = 0$. Again, consider the points p and q and compute the normal and the tangent spaces. What is the dimension of the normal space? Why is it wrong?

11. Construct a polynomial $h \in \sqrt{(f, g)}$ which will repair the above situation.

12. Compute the resultant of $f \equiv x^2 + y^2 - 2$ and $g \equiv y - x^2$, by eliminating y. Explain the outcome. What happens when we eliminate x?

13. Consider f as above and $h \equiv x + y - 1$. Explain.

14. Compute the locus (a, b) of the translations of the unit circle which osculate with the parabola $y - x^2 = 0$.

15. Recall the action of $\mathbb{Z}_n$ on $V = \mathbb{C}^n$ by shifting. Decompose V into irreducible subspaces. Describe $\mathbb{C}[V]^G$ and show how orbits can be separated.

16. Let X be the vector space of 2×2-matrices, with coordinate vectors (X_{ij}). Consider the action $\rho : GL_2 \to End(X)$, given by $(g, x) \to gxg^{-1}$. Let the coordinate vectors of G be (g_{ij}). Decompose X under this action into irreducible spaces $X = X_0 \oplus X_1$. Explicitly construct the map:

$$\rho^* : \mathbb{C}[X] \to \mathbb{C}[G] \otimes \mathbb{C}[X]$$

Verify that $X_{11} + X_{22}$ is an invariant.

17. Corresponding to the map above, compute the bilinear Lie algebra map:

$$\rho_1 : gl_2 \times X \to X$$

Verify that X_0 and X_1 above continue to be invariant. Compute the stabilizers of various points.

18. Let $V = \mathbb{C}^{3\times 4}$, the space of 3×4-matrices, and $G = GL_4$ and $G' = SL_4$ act by left multiplication. Let v_1, v_2 and v be the points below:

$$v_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} v_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix} v = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Compute the G and G' orbits of all points. Recall the relation $\sim$ on V where we say $v \sim v'$ if their orbit closures intersect. Compute $[w]_\sim$ for $\sim$ arising from G and G' and for $w = v_1, v_2, v$. Recall that $[w]_\sim$ is also the equivalence class of separation by invariants. Construct suitable invariants to distinguish between equivalence classes.

19. Let $V = \mathbb{C}^{m \times n}$, the space of all $m \times n$-matrices with $m \leq n$. Let us consider the action of V under $G = GL_m$ and $G' = SL_m$. For a point v, let $A(v) = \{c \in \mathbb{C}^n | vc = 0\}$, the space of right annihilators. Clearly, if $v \in O(v')$ then $A(v) = A(v')$. Does $A(v)$ determine $[v]_\sim$?

20. Recall that for an action of reductive group G on V, we say that v is stable if its orbit is closed. Consider the diagonal action of G on $W = V \oplus V$. What would be stable or unstable points of W in terms of those of V? What about $\mathbb{C}[W]^G$ in terms of $\mathbb{C}[V]^G$? Analyse the case when $V = Mat_n$, the space of $n \times n$-matrices and $G = GL_n$.

21. Let T be a 1-dimensional torus acting on V through ρ. Then we have $V = \oplus V_i$ with $\rho(t)v_i = t^i v_i$ for some $i \in \mathbb{Z}$. Describe the orbits of a typical point v. When is it closed? What is $\mathbb{C}[V]$ and $\mathbb{C}[V]^T$? Analyse $[v]_\sim$.

22. Recall that for V acted upon by a reductive group G, we have $v' \in \overline{O(v)}$ iff there is a 1-parameter subgroup $\lambda(t) : \mathbb{C}^* \to G$ such that $\lim_{t\to 0} \lambda(t)v \in \overline{O(v')}$. Let $X = \{x_1, x_2, x_3\}$ and $V = Sym^d(X)$, the space of polynomials of degree d in X. What is the dimension of V? Let us now analyse V under the action of $G = SL_3$. We know that the maximal torus T in G is isomorphic to $T = \{diag(\mu_1, \mu_2, \mu_3) | \mu_1\mu_2\mu_3 = 1\}$. A 1-paramater subgroup $\lambda : \mathbb{C}^* \to G$ is then given by, upto conjugation, by $\lambda(t) = \{(t^{d_1}, t^{d_2}, t^{d_3} | d_1 + d_2 + d_3 = 0\}$. Using this, classify all polynomials $p \in V$ such that $0 \in \overline{O(v)}$.

23. Consider $G = GL_4(\mathbb{C})$ and the 1-parameter subgroup $\lambda(t)$ and consider $\mathcal{K}$ as given below and compute $\hat{\mathcal{K}}$.

$$\lambda(t) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & t \end{bmatrix}, \ \mathcal{K} = \begin{bmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & c & d \end{bmatrix}$$

Next, let A be as shown below and $\mathcal{K}' = A\mathcal{K}A^{-1}$ as shown below. Compute $\hat{\mathcal{K}}'$.

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \ \mathcal{K}' = \begin{bmatrix} a & b & c & d-a \\ c & d & 0 & -c \\ 0 & 0 & a & b \\ 0 & 0 & c & d \end{bmatrix}$$

What if A above is replaced by A'?

$$A' = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

24. Consider a set of indeterminates $\{x, y, z\}$ and $X = \mathbb{C} \cdot \{x, y, z\}$. Let GL_3 acts on X in the natural way, and let $V = Sym^4(X)$. Consider $f = ((x-z)^2 + y^2 + z^2)^2 \in V$. Compute $\mathcal{G}_f$.

Consider next the 1-PS $\lambda(t) \subseteq GL(X)$ given by $\lambda(x) = x, \lambda(y) = y$ and $\lambda(z) = tz$, as shown below. Compute $\hat{f}$ and the tangent of approach h. Verify that $\mathcal{G}_{\hat{f}} \subseteq (\mathcal{H}_f)_{\overline{h}} \subseteq \mathcal{G}_f$.

25 Solutions

1.

2. (Keerthana) Given that S_1 is the stabiliser of the face, we have 8 elements of symmetry for the face $4 -> $ rotation and 1 reflection.

$$\Rightarrow 8 \text{ in total}$$

Given that S_2 is the list of coset representatives. This will be same as number of faces = 6.

$$\Rightarrow \text{There are } 8 \times 6 \text{ elements in } S_! \times S_2$$

where S_1 is everything that keeps the face intact and S_2 is total number of positions for faces.

3.

4. (Keerthana) Given that $fg = gf \Rightarrow f = gfg^{-1}$. let the set be $x_1x_2x_3x_4x_5x_6x_7$

$$(x_1, x_2, x_3)(x_4x_5)(x_6x_7) \equiv (g(x_1), g(x_2), g(x_3))(g(x_4)g(x_5))(g(x_6)g(x_7))$$

$$\Rightarrow \underbrace{(x_1, x_2, x_3) = (g(x_1), g(x_2), g(x_3))}_{\text{3 possibilities of g}}$$

<u>**Case1:**</u>

$$(x_4, x_5) = (g(x_4), g(x_5)) \qquad (x_6, x_7) = (g(x_6), g(x_7))$$

This gives us two possibilities

<u>**Case2:**</u>

$$(x_4, x_5) = (g(x_6), g(x_7)) \qquad (x_6, x_7) = (g(x_4), g(x_5))$$

This gives us two possibilities

$\Rightarrow Total = 3 \times (2 \times 2 + 2 \times 2) = 24\, possibilities\ of\ g$

5. (Keerthana) Small subspace M that contains $(1, -1, 0, 0, 0, 0)$ and it is Z_6 invariant.

$$(-1, 0, 0, 0, 0, 1)$$
$$(0, 0, 0, 0, 1, -1)$$
$$\vdots$$

all these elements should be there.
$\Rightarrow$ The subspace is going to be the set of all vectors with sum of coordinates 0.
Now split these in such a way that they are Z_3 invariant.

Suppose $(x_1, x_2, x_3, x_4, x_5, x_6) = K(x_3, x_4, x_5, x_6, x_1, x_2)$
The K is the cube root of unity

$\Rightarrow\ K = 1 \qquad or \qquad K = \omega \qquad or \qquad K = \omega^2$

<u>$K = 1$**:**</u>
$x_1 = x_3 = x_5$
$x_2 = x_4 = x_6$
$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 0$
$\Rightarrow\ x_1 = -x_2$
$\Rightarrow$ They are going to be of the form $(x, -x, x, -x, x, -x)$

<u>$K = \omega$**:**</u>
$x_1 = x_3\omega$
$x_3 = x_5\omega$
$x_2 = x_4\omega$
$x_4 = x_6\omega$
$(\omega^2 x, \omega^2 y, \omega x, \omega y, x, y)$

<u>$K = \omega^2$**:**</u>
This case gives $(\omega x, \omega y, \omega^2 x, \omega^2 y, x, y)$

6. (Shantanu) We will prove this by double induction w.r.t. number of variables n, and for a fixed n, with respect to the degree.
   The case $n = 1$ is trivial because every polynomial is symmetric, and the only elementary symmetric polynomials is x_1. Assume now that the theorem has been proved for all polynomials for $m < n$ variables and all symmetric polynomials in n variables with degree $< d$. It is enough to prove the result for homogeneous symmetric polynomials of degree d. Decompose a symmetric polynomial P into two parts: $P = P_{bad} + x_1 x_2 \cdots x_n Q$, i.e., P_{bad} consists of all monomials that don't contain all the x_i (call these bad monomials), and we have factored an $x_1 x_2 \cdots x_n$ from the rest of the monomials. By induction hypothesis on degree, $Q \in \mathbb{C}[s_1, s_2, \ldots, s_n]$, and $x_1 x_2 \cdots x_n = s_n$ is also elementary. Thus it is enough to focus on P_{bad}.

   Define $P_0(x_1, \ldots, x_{n-1}) = P(x_1, \ldots, x_{n-1}, 0) = P_{bad}(x_1, \ldots, x_{n-1}, 0)$. This is a symmetric polynomial in $n-1$ variables, so by induction hypothesis, it is equal to $T(s'_1, s'_2, \ldots, s'_{n-1})$, where s'_i is elementary symmetric in $x_1, \ldots, x_{n-1}$. Consider the polynomial $T(s_1, s_2, \ldots, s_n)$. Since this is equal to $T(s'_1, \ldots, s'_{n-1}) = P_0(x_1, \ldots, x_n)$ when $x_n = 0$, any monomial not containing x_n has the same coefficient in T and P_0. But, the coefficient of any term in P_{bad} is determined by the monomial not containing x_n which can be obtained by permutation of variables. Therefore they have the same "bad" part, so $P_{bad} - T$ is divisible by $x_1 x_2 \cdots x_n$. Now we can use induction hypothesis on degree to conclude.

7. (Keerthana) Given Π_G is $R \to R^G$
   we have that, $\Pi_G(f) = \frac{\Sigma g \cdot f}{|G|}$
   $\Rightarrow \Pi_G(ff') = \frac{\Sigma g \cdot (ff')}{|G|}$

   Given that $f^G = f$

   $\Rightarrow \Pi_G(ff') = \frac{\Sigma gf \cdot (gf')}{|G|} = \frac{\Sigma f \cdot (gf')}{|G|} = f\frac{\Sigma (gf')}{|G|} = f\Pi_G(f')$

   $\therefore$ Hence proved

8.

9. (Keerthana) Given that $f = x^2 + y^2 - z^2$ and P=$(0,0,0)$ and q=$(0,1,1)$.
   Tangent space is $\begin{bmatrix} 2x & 2y & -2z \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0$
   $\Rightarrow 2xa + 2yb - 27c = 0$
   For $(0,0,0)$ tangent space is everything.
   For $(0,1,1)$ tangent space is (a,b,c) such that $b = c$.
   Normal space is vector space generated by $(2x, 2y, -2z)$
   for $(0,0,0)$ it is just the point $(0,0,0)$ - 0 dimensional
   for $(0,1,1)$ it is just the vector space formed by $(0,1,1)$ - 1 dimensional
   It is smooth at $(0,1,1)$.

10. (Keerthana) $f \equiv x^2 + y^2 - z^2 = 0andg \equiv x^2 + z^2 - y^2 = 0$
    The intersection is going to be lines $y^2 = z^2$ and $x = 0$
    The normal space is going to be the space formed by $(2x, 2y, -2z)$,$(2x, -2y, 2z)$
    for $(0,0,0)$, it is going to be 0 dimensional
    for $(0,1,1)$, it is going to be 1 dimensional
    The tangent space is going to be $2xa + 2yb - 2zc = 0$ and $2xa - 2yb + 2zc = 0$
    For $(0,0,0)$ it is going to be the entire space.
    For $(0,1,1)$ it is going to be (a,b,c) such that $b = c$.
    But the mistake is, for (0,1,1) the dimension of the normal space is 2 not 1.
    Because of the x^2 term it is superimposing $x = 0$ twice and taking a normal perpendicular to that and making it 1 dimensional.

11. (Keerthana) $h \in \sqrt{(f,g)}$
    $\sqrt{\frac{f+g}{2}} = x$
    If we take $h \equiv x = 0$ and $f \equiv x^2 + y^2 - z^2 = 0$.
    The curve intersection is still going to be the same but the normal space is going to be the space generated by $(1,0,0)$ and $(0,2,-2)$. This is going to be 2 dimensional.
    $\Rightarrow$ The issue is resolved

12. (Keerthana) Given that $f \equiv x^2 + y^2 - 2$ and $g \equiv y - x^2$
We need to eliminate y.
$\Rightarrow\ f \equiv y^2 + (x^2 - 2) andg \equiv y - x^2$

d=2,e=1,d+e=3

$$\begin{bmatrix} 1 & 0 & x^2-2 \\ 1 & -x^2 & 0 \\ 0 & 1 & -x^2 \end{bmatrix} = x^4 + x^2 - 2$$

This is the same as substituting $y = x^2$(same as g) in f.

13. (Keerthana) Given that $f \equiv x^2 + y^2 - 2 \equiv y^2 + (x^2 - 2)$ and $h \equiv x + y - 1 \equiv y + (x - 1)$, d+e=3

$$\begin{bmatrix} 1 & 0 & x^2-2 \\ 1 & x-1 & 0 \\ 0 & 1 & x-1 \end{bmatrix} = (x-1)^2 + (x^2 - 2)$$

This is the same as replacing $y \equiv -\underbrace{(x-1)}_{\text{same as g}} inf$

14. (Keerthana) x coordinate of the point of intersection of $y = x^2$ and $(x-a)^2 + (y-b)^2 = 1$ will be the root of the equation $(x-a)^2 + (x^2-b)^2 = 1$
$\Rightarrow\ x^4 - 2bx^2 + x^2 - 2ax + a^2 + b^2 - 1 = 0$
$\Rightarrow\ x^4 + (1-2b)x^2 - 2ax + a^2 + b^2 - 1 = 0$
Now this has 2 equal roots and 2 imaginary roots. let c,c be the repeated roots, since circle is touching the parabola.
sum of the roots=0 $\Rightarrow$ imaginary roots are $-c + idand - c - id$
Product of the roots $\Rightarrow a^2 + b^2 = c^2(c^2 + d^2) \to 1$
and $2a = cc(-c+id) + cc(-c-id) + (-c+id)(-c-id)c + (-c+id)(-c-id)c$
$\Rightarrow 2a = c^2(-2c) + 2c(c^2 + d^2)$
$\Rightarrow 2a = 2cd^2$
$\Rightarrow a = cd^2$
Now $1 - 2b = cc + c(-c+id) + c(-c-id) + c(-c+id) + c(-c-id) + (-c-id)(-c+id)$
$\Rightarrow 1 - 2b = c^2 + c(-2c) + c(-2c) + c^2 + d^2$
$\Rightarrow 1 - 2b = c^2 - 4c^2 + c^2 + d^2$
$\Rightarrow 1 - 2b = d^2 - 2c^2$
substituting $d^2 = a/c$ in equation 1 and 3
$\Rightarrow a^2 + b^2 - 1 = c^2(c^2 + a/c) \to 4\ and\ 1 - 2b = a/c - 2c^2$
$\Rightarrow 1/2 - b = a/2c - c^2$
$\Rightarrow c^2 = (1/2 - b) - (a/2c) \to 5$
By substituting 5 in 4 gives
$\Rightarrow a^2 + b^2 - 1 = c^2((1/2 - b) + a/2c)$
$\Rightarrow a^2 + b^2 - 1 = c^2(1/2 - b) + ac/2$
$\Rightarrow c^2(2b-1) - ac + 2(a^2 + b^2 - 1) = 0 \to 6$
and equation 5 is $2c^3 = (1/2 - b)2c - a$
$\Rightarrow 2c^3 = (1-2b)c - a$
$\Rightarrow 2c^3 - (1-2b)c + a = 0 \to 7$

From 6 we have $c^2 = \frac{ac - 2(a^2+b^2-1}{2b-1}$

substitute in 7 gives $2c(\frac{ac-2(a^2+b^2-1}{2b-1}) - (1-2b)c + a = 0$
$\Rightarrow 2ac^2 - 4c(a^2 + b^2 - 1) + (2b-1)^2c + a(2b-1) = 0$
$\Rightarrow 2ac^2 + (4b^2 - 4b + 1 - 4a^2 - 4b^2 + 4)c + a(2b-1) = 0$
$\Rightarrow 2ac^2 + (5 - 4b - 4a^2)c + a(2b-1) = 0$

again substitute $c^2 = \frac{ac-2(a^2+b^2-1)}{(2b-1)}$

$\Rightarrow 2a\frac{ac-2(a^2+b^2-1)}{(2b-1)} + (5 - 4b - 4a^2)c + a(2b-1) = 0$

$\Rightarrow 2ac^2 - 4a(a^2 + b^2 - 1 + (2b-1)(5 - 4b - 4a^2)c + a(2b-1)^2 = 0$

$\Rightarrow c(2a^2 + (2b-1)(5-4b-4a^2)) = 4a(a^2+b^2-1-a(2b-1)^2$

$\Rightarrow c = \frac{4a(a^2+b^2-1)-a(2b-1)^2}{2a^2(2b-1)(5-4b-4a^2)}$

substitution in equation 6 that is $c^2(2b-1) - ac + 2(a^2+b^2-1) = 0$ gives

$\Rightarrow c(2b-1)(\frac{4a(a^2+b^2-1)-a(2b-1)^2}{2a^2+(2b-1)(5-4b-4a^2)}) - ac + 2(a^2+b^2-1) = 0$

$\Rightarrow c(\frac{(2b-1)(4a(a^2+b^2-1)-a(2b-1)^2)}{2a^2+(2b-1)(5-4b-4a^2)}) - a) = -2(a^2+b^2-1)$

$\Rightarrow c = \frac{-2(a^2+b^2-1)(2a^2+(2b-1)(5-4b-4a^2))}{(2b-1)(4a(a^2+b^2-1)-a(2b-1)^2)-a(2a^2+(2b-1)(5-4b-4a^2))}$

Equating both c's gives $\Rightarrow \frac{4a(a^2+b^2-1)-a(2b-1)^2}{2a^2(2b-1)(5-4b-4a^2)} = \frac{-2(a^2+b^2-1)(2a^2+(2b-1)(5-4b-4a^2))}{(2b-1)(4a(a^2+b^2-1)-a(2b-1)^2)-a(2a^2+(2b-1)(5-4b-4a^2))}$

15. (Shantanu) Since $\mathbb{Z}_n$ is abelian, all irreducible subspaces are one-dimensional. For $0 \leq i \leq n-1$, consider the vector $v_i = (1, w^i, w^2 i, \ldots, w^{(n-1)i})$, where w is the primitive n-th root of unity. Suppose $\mathbb{Z}_n$ is generated by a, then
$$a \cdot v_i = (w^{(n-1)i}, 1, \ldots, w^{(n-2)i}) = w^{(n-1)i} v_i$$
Hence each v_i is an eigenvector with a different eigenvalue, so the irreducible subspaces of $\mathbb{C}^n$ are linear subspaces generated by the v_i.

The invariant subspace $\mathbb{C}[V]^G$ will be polynomials that are invariant under cyclic shifts, i.e., polynomials $P \in \mathbb{C}[V]$ such that $P(x_1, x_2, \ldots, x_n) = P(x_n, x_1, \ldots, x_{n-1})$. This will form a finitely generated algebra since $\mathbb{Z}_n$ finite implies $\mathbb{Z}_n$ reductive. Now, $\mathbb{Z}_n$ finite also implies each orbit is finite, so it is closed; therefore two orbit closures are disjoint iff the orbits are disjoint. Therefore by 17.16, any two disjoint orbits can be separated.

16. (Shantanu) We have the canonical decomposition into $X_0 = X^G$ and X_1. Here X_0 is the space of invariants under ρ. $x \in X_0 \iff \rho(g)x = x$ for all $g \in G \iff gxg^{-1} = x$ for all $g \in G \iff x$ commutes with all of $GL_2 \iff x$ is a scalar diagonal matrix.

Its complement X_1 is generated by elements of the form $x - \rho(g)x$ for $x \in X$ and $g \in G$. If we let $x = yg$, we see that X_1 is generated by matrices of the form $yg - gy$ for $y \in X$ and $g \in G$.

Now we construct the map ρ^*. If $g = \begin{bmatrix} g_1 & g_2 \\ g_3 & g_4 \end{bmatrix}$, then

$$\begin{aligned}
\rho^*(X_{11}) = g^{-1}X_{11}g &= \frac{1}{g_1g_4 - g_2g_3}(g_1g_4X_{11} + g_2g_4X_{12} - g_1g_3X_{21} - g_2g_3X_{22}) \\
\rho^*(X_{12}) = g^{-1}X_{12}g &= \frac{1}{g_1g_4 - g_2g_3}(g_3g_4X_{11} + g_4^2X_{12} - g_3^2X_{21} - g_3g_4X_{22}) \\
\rho^*(X_{21}) = g^{-1}X_{21}g &= \frac{1}{g_1g_4 - g_2g_3}(-g_1g_2X_{11} - g_2^2X_{12} + g_1^2X_{21} + g_1g_2X_{22}) \\
\rho^*(X_{22}) = g^{-1}X_{22}g &= \frac{1}{g_1g_4 - g_2g_3}(-g_2g_3X_{11} - g_2g_4X_{12} + g_1g_3X_{21} + g_1g_4X_{22})
\end{aligned}$$

This directly shows $\rho^*(X_{11} + X_{22}) = X_{11} + X_{22}$, i.e., it is invariant.

17. (Shantanu) Let $h \in gl_2$, then $I + \epsilon h \in G$ if $\epsilon^2 = 0$. Therefore

$$\begin{aligned}
\rho(I + \epsilon h)(x) &= (I + \epsilon h)x(I + \epsilon h)^{-1} \\
&= (I + \epsilon h)x(I - \epsilon h) \\
&= x + \epsilon(hx - xh)
\end{aligned}$$

Hence $\rho_1(h, x) = hx - xh$.

If $x \in X_0$, then x is a scalar so it commutes with every matrix $\implies \rho_1(x, h) = 0 \in X_0$, so X_0 is invariant under ρ_1.

X_1 is generated by elements of the form $yg - gy$ for $y \in X$ and $g \in G$. Thus it is sufficient to prove that $\rho_1(yg - gy, h) \in X_1$ for all $h \in gl_2$. But this is true because

$$h(yg - gy) - (yg - gy)h = hyg - hgy - ygh + gyh = (hyg - ygh) + (gyh - hgy) \in X_1$$

18.

19.

20.

21. (Shantanu) If $v = \cdots + v_{-1} + v_0 + v_1 + \cdots$, then $\rho(t)v = \cdots + t^{-1}v_{-1} + v_0 + tv_1 + \cdots$. $v \in V_0$, then every orbit is constant, so it is closed. If $v_i = 0$ for all $i < 0$ but $v \notin V_0$, then $\rho(t)v \to v_0 \notin O(v)$ as $t \to 0$, so orbit is not closed. Similarly if $v_i = 0$ for all $i > 0$ but $v \notin V_0$, then $\rho(t^{-1})v \to v_0 \notin O(v)$ as $t \to 0$, so orbit is not closed. Finally, if v has both positive and negative parts, then for any non-zero integer k, $\lim\limits_{t\to 0} \rho(t^k)v$ does not exist, so orbit is closed by the 1-parameter subgroup characterization.

Also note that by the above paragraph, if v has only positive or only negative parts, then $\overline{O(v)} = O(v) \cup \{v_0\}$. So $v \sim w$ iff they are in the same orbit, or if both of them only have positive or negative parts, and their projections on V_0 are the same.

Let $\{u_i^{(1)}, \ldots, u_i^{(d_i)}\}$ be a basis of V_i. Then $\mathbb{C}[V]$ is polynomials on $u_i^{(j)}$. Since ρ just scales each variable, a polynomial is invariant iff every monomial is invariant. A monomial $\prod (u_i^{(j)})^{a_{i,j}}$ is invariant iff $\sum i a_{i,j} = 0$. Thus $\mathbb{C}[v]^T$ is the $\mathbb{C}$-vector space of all such monomials.

22.

23. (Keerthana) Given that,

$$\lambda(t) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & t \end{bmatrix} \; and\, \mathcal{K} = \begin{bmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & c & d \end{bmatrix}$$

$$\lambda(t)\mathcal{K}\lambda(t)^{-1} = \begin{bmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & ct & dt \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1/t \end{bmatrix} = \begin{bmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & a & b/t \\ 0 & 0 & ct & d \end{bmatrix}$$

$$\lambda(t)\mathcal{K}\lambda(t)^{-1} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & b \\ 0 & 0 & 0 & 0 \end{bmatrix} t^{-1} + \begin{bmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & d \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & c & 0 \end{bmatrix} t$$

The possible leading coefficients are $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & b \\ 0 & 0 & 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & d \end{bmatrix}$

Then $\hat{\mathcal{K}}$ will be $\begin{bmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & 0 & d \end{bmatrix}$

Then,

$$\lambda(t)\mathcal{K}'\lambda(t)^{-1} = \begin{bmatrix} a & b & c & (d-a)/t \\ c & d & 0 & -c/t \\ 0 & 0 & a & b/t \\ 0 & 0 & ct & d \end{bmatrix}$$

$$\lambda(t)\mathcal{K}'\lambda(t)^{-1} = \begin{bmatrix} 0 & 0 & 0 & d-a \\ 0 & 0 & 0 & -c \\ 0 & 0 & 0 & b \\ 0 & 0 & 0 & 0 \end{bmatrix} t^{-1} + \begin{bmatrix} a & b & c & 0 \\ c & d & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & d \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & c & 0 \end{bmatrix} t$$

The leading coefficients could be $\begin{bmatrix} 0 & 0 & 0 & d-a \\ 0 & 0 & 0 & -c \\ 0 & 0 & 0 & b \\ 0 & 0 & 0 & 0 \end{bmatrix}$ or $\begin{bmatrix} a & b & c & 0 \\ c & d & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & d \end{bmatrix}$

$$\mathcal{K}' = \begin{bmatrix} a & b & c & e \\ c & d & 0 & f \\ 0 & 0 & 0 & g \\ 0 & 0 & 0 & d \end{bmatrix}$$

24.

References

[Art11] M. Artin. *Algebra.* Pearson Education, 2011.

[AW09] Scott Aaronson and Avi Wigderson. Algebrization: A New Barrier in Complexity Theory. *ACM Transactions on Computation Theory (TOCT)*, 1(1):1–54, 2009.

[BGS75] Theodore Baker, John Gill, and Robert Solovay. Relativizations of the P=?NP question. *SIAM Journal on computing*, 4(4):431–442, 1975.

[MS01] Ketan D Mulmuley and Milind Sohoni. Geometric Complexity Theory I: An Approach to the P vs. NP and Related Problems. *SIAM Journal on Computing*, 31(2):496–526, 2001.

[RR94] Alexander A Razborov and Steven Rudich. Natural Proofs. In *Proceedings of the twenty-sixth annual ACM symposium on Theory of computing*, pages 204–213, 1994.

[Sha92] Adi Shamir. IP= PSPACE. *Journal of the ACM (JACM)*, 39(4):869–877, 1992.

[Val79] Leslie G Valiant. Completeness classes in algebra. In *Proceedings of the eleventh annual ACM symposium on Theory of computing*, pages 249–261, 1979.