An Introduction to Invariant Theory

Harm Derksen, University of Michigan

Optimization, Complexity and Invariant Theory
Institute for Advanced Study, June 4, 2018
Plan of the Talk

- applications of invariants
Plan of the Talk

▶ applications of invariants
▶ a classical, motivating example: binary forms
Plan of the Talk

▶ applications of invariants
▶ a classical, motivating example: binary forms
▶ polynomial rings ideals
Plan of the Talk

▶ applications of invariants
▶ a classical, motivating example: binary forms
▶ polynomial rings ideals
▶ group representations and invariant rings

Harm Derksen, University of Michigan
An Introduction to Invariant Theory
Plan of the Talk

- applications of invariants
- a classical, motivating example: binary forms
- polynomial rings ideals
- group representations and invariant rings
- Hilbert’s Finiteness Theorem
Plan of the Talk

- applications of invariants
- a classical, motivating example: binary forms
- polynomial rings ideals
- group representations and invariant rings
- Hilbert’s Finiteness Theorem
- the null cone and the Hilbert-Mumford criterion
Plan of the Talk

- applications of invariants
- a classical, motivating example: binary forms
- polynomial rings ideals
- group representations and invariant rings
- Hilbert's Finiteness Theorem
- the null cone and the Hilbert-Mumford criterion
- degree bounds for invariants
Plan of the Talk

▶ applications of invariants
▶ a classical, motivating example: binary forms
▶ polynomial rings ideals
▶ group representations and invariant rings
▶ Hilbert’s Finiteness Theorem
▶ the null cone and the Hilbert-Mumford criterion
▶ degree bounds for invariants
▶ polarization of invariants and Weyl’s Theorem
Plan of the Talk

▶ applications of invariants
▶ a classical, motivating example : binary forms
▶ polynomial rings ideals
▶ group representations and invariant rings
▶ Hilbert’s Finiteness Theorem
▶ the null cone and the Hilbert-Mumford criterion
▶ degree bounds for invariants
▶ polarization of invariants and Weyl’s Theorem
▶ Invariant Theory for other fields
Applications of Invariants

Definition

An invariant is a quantity or expression that stays the same under certain operations.
Applications of Invariants

Definition

An *invariant* is a quantity or expression that stays the same under certain operations.

The total energy in a physical system is an *invariant* as the system evolves over time.
Applications of Invariants

Definition

an *invariant* is a quantity or expression that stays the same under certain operations

the total energy in a physical system is an *invariant* as the system evolves over time

*loop invariants* can be used to prove the correctness of an algorithm although the number of iterations in a loop may vary, the loop invariant tell us to say something about the variables after the iterations
Applications of Invariants

*Knot invariants* (such as the Jones polynomial) can be used to distinguish knots.

Knot invariants remain unchanged under Reidemeister moves.
Applications of Invariants

*Knot invariants* (such as the Jones polynomial) can be used to distinguish knots.

Knot invariants remain unchanged under Reidemeister moves.

(co-)homology groups are invariants of topological manifolds.
In invariant theory we restrict ourselves to

- invariants that are polynomial functions on a vector space
in invariant theory we restrict ourselves to

- invariants that are polynomial functions on a vector space
- invariants that remain unchanged under *group symmetries* such as rotations, permutations etc.
In invariant theory we restrict ourselves to

- invariants that are polynomial functions on a vector space
- invariants that remain unchanged under group symmetries such as rotations, permutations etc.

we start with a motivating example from 19th century invariant theory
Classical Invariant Theory: Binary Forms

A *binary form* of degree 2 is a polynomial

\[ p(z, w) = p_1 z^2 + p_2 zw + p_3 w^2 \]

with \( p_1, p_2, p_3 \in \mathbb{C} \)
a binary form of degree 2 is a polynomial

\[ p(z, w) = p_1 z^2 + p_2 zw + p_3 w^2 \]

with \( p_1, p_2, p_3 \in \mathbb{C} \)

\[ \text{SL}_2 = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : ad - bc = 1 \right\} \]

is the group of \( 2 \times 2 \) matrices with determinant one

a matrix \( A \in \text{SL}_2 \) gives a linear change of coordinates in \( \mathbb{C}^2 \)
a binary form of degree 2 is a polynomial

\[ p(z, w) = p_1 z^2 + p_2 zw + p_3 w^2 \]

with \( p_1, p_2, p_3 \in \mathbb{C} \)

\[ \text{SL}_2 = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : ad - bc = 1 \right\} \]

is the group of 2 \( \times \) 2 matrices with determinant one

a matrix \( A \in \text{SL}_2 \) gives a linear change of coordinates in \( \mathbb{C}^2 \)

the group \( \text{SL}_2 \) acts on (the coefficients of) binary forms:
we make the substitution \((z, w) \mapsto (az + cw, bz + dw)\) and get another polynomial

\[ p'(z, w) = p(az + cw, bz + dw) = p'_1 z^2 + p'_2 zw + p'_3 w^2 \]
where

\[ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}_2, \]

\[ \begin{pmatrix} p'_1 \\ p'_2 \\ p'_3 \end{pmatrix} = M_A \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} \]

and

\[ M_A = \begin{pmatrix} a^2 & ab & b^2 \\ ac & ad + bc & bd \\ c^2 & cd & d^2 \end{pmatrix} \]
the polynomial \( f(x_1, x_2, x_3) = x_2^2 - 4x_1x_3 \in \mathbb{C}[x_1, x_2, x_3] \) (the discriminant) can be viewed as a function from \( \mathbb{C}^3 \) to \( \mathbb{C} \) and an easy calculation shows that

\[
f \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = p_2^2 - 4p_1p_3 = (p_2')^2 - 4p_1'p_3' = f \begin{pmatrix} p_1' \\ p_2' \\ p_3' \end{pmatrix}
\]
the polynomial \( f(x_1, x_2, x_3) = x_2^2 - 4x_1x_3 \in \mathbb{C}[x_1, x_2, x_3] \) (the discriminant) can be viewed as a function from \( \mathbb{C}^3 \) to \( \mathbb{C} \) and an easy calculation shows that

\[
f \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = p_2^2 - 4p_1p_3 = (p'_2)^2 - 4p'_1p'_3 = f \begin{pmatrix} p'_1 \\ p'_2 \\ p'_3 \end{pmatrix}
\]

we say that \( f(x_1, x_2, x_3) \) is an *invariant* under the action of \( SL_2 \).
the polynomial \( f(x_1, x_2, x_3) = x_2^2 - 4x_1x_3 \in \mathbb{C}[x_1, x_2, x_3] \) (the discriminant) can be viewed as a function from \( \mathbb{C}^3 \) to \( \mathbb{C} \) and an easy calculation shows that

\[
f \left( \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} \right) = p_2^2 - 4p_1p_3 = (p_2')^2 - 4p_1'p_3' = f \left( \begin{pmatrix} p_1' \\ p_2' \\ p_3' \end{pmatrix} \right)
\]

we say that \( f(x_1, x_2, x_3) \) is an *invariant* under the action of \( \text{SL}_2 \)

\( f(x_1, x_2, x_3) \) is a *fundamental* invariant that generates all invariants: if \( h(x_1, x_2, x_3) \) is another polynomial invariant, then there exists a polynomial \( q(y) \) such that \( h(x_1, x_2, x_3) = q(f(x_1, x_2, x_3)) \)
we may identify binary forms of degree $n$ with vectors in $\mathbb{C}^{n+1}$:

$$p_1 z^n + p_2 z^{n-1} w + \cdots + p_{n+1} w^n \leftrightarrow \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_{n+1} \end{pmatrix}$$
we may identify binary forms of degree \( n \) with vectors in \( \mathbb{C}^{n+1} \):

\[
p_1 z^n + p_2 z^{n-1} w + \cdots + p_{n+1} w^n \leftrightarrow \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_{n+1} \end{pmatrix}
\]

the vector space of binary forms of degree \( n \) is an \((n + 1)\)-dimensional representation of \( \text{SL}_2 \)
polynomial invariants for binary forms of arbitrary degree were extensively studied in the 19th century by mathematicians like Boole, Sylvester, Cayley, Aronhold, Hermite, Eisenstein, Clebsch, Gordan, Lie, Klein, Capelli etc.
polynomial invariants for binary forms of arbitrary degree were extensively studied in the 19th century by mathematicians like Boole, Sylvester, Cayley, Aronhold, Hermite, Eisenstein, Clebsch, Gordan, Lie, Klein, Capelli etc.

Theorem (Gordan 1868)

for binary forms of degree \( d \) there exists a finite system of fundamental invariants that generate all invariants (i.e., every invariant is a polynomial expression in the fundamental invariants)
polynomial invariants for binary forms of arbitrary degree were extensively studied in the 19th century by mathematicians like Boole, Sylvester, Cayley, Aronhold, Hermite, Eisenstein, Clebsch, Gordan, Lie, Klein, Capelli etc.

**Theorem (Gordan 1868)**

for binary forms of degree $d$ there exists a finite system of fundamental invariants that generate all invariants (i.e., every invariant is a polynomial expression in the fundamental invariants)

one of the main objectives was to find an explicit system of fundamental invariants for binary forms up to degree $d$

(currently known for $d \leq 10$)
The Polynomial Ring

\[ x_1, x_2, \ldots, x_n \] coordinate functions on \( V = \mathbb{C}^n \)
a polynomial \( f(x_1, \ldots, x_n) \) can be viewed as function from \( V \) to \( \mathbb{C} \)
\( \mathbb{C}[x] = \mathbb{C}[x_1, \ldots, x_n] \) graded ring of polynomial functions
The Polynomial Ring

\(x_1, x_2, \ldots, x_n\) coordinate functions on \(V = \mathbb{C}^n\)
a polynomial \(f(x_1, \ldots, x_n)\) can be viewed as function from \(V\) to \(\mathbb{C}\)
\(\mathbb{C}[x] = \mathbb{C}[x_1, \ldots, x_n]\) graded ring of polynomial functions

Definition (Ideal)

A subset \(I \subseteq \mathbb{C}[x]\) is an ideal if

1. \(0 \in I\);
2. \(f(x), g(x) \in I \Rightarrow f(x) + g(x) \in I\);
3. \(f(x) \in \mathbb{C}[x], g(x) \in I \Rightarrow f(x)g(x) \in I\).
Hilbert’s Basis Theorem

the *ideal* \( (S) \) generated by a subset \( S \subseteq \mathbb{C}[x] \) is

\[
\{ a_1(x)f_1(x) + \cdots + a_r(x)f_r(x) \mid r \in \mathbb{N}, \forall i a_i(x) \in \mathbb{C}[x], f_i(x) \in S \}
\]
Hilbert’s Basis Theorem

the ideal \((S)\) generated by a subset \(S \subseteq \mathbb{C}[x]\) is

\[
\{a_1(x)f_1(x) + \cdots + a_r(x)f_r(x) \mid r \in \mathbb{N}, \forall i \; a_i(x) \in \mathbb{C}[x], f_i(x) \in S\}
\]

**Theorem (Hilbert 1890)**

*every ideal \(I \subseteq \mathbb{C}[x]\) is generated by a finite set*

*(\(\mathbb{C}[x]\) is noetherian)*
Hilbert’s Basis Theorem

the *ideal* $(S)$ generated by a subset $S \subseteq \mathbb{C}[x]$ is

$$\{a_1(x)f_1(x) + \cdots + a_r(x)f_r(x) \mid r \in \mathbb{N}, \forall i \ a_i(x) \in \mathbb{C}[x], \ f_i(x) \in S\}$$

**Theorem (Hilbert 1890)**

*every ideal $I \subseteq \mathbb{C}[x]$ is generated by a finite set*  
($\mathbb{C}[x]$ is noetherian)

*if $S \subseteq \mathbb{C}[x]$, then $(S) = (T)$ for some finite subset $T \subseteq S$*
The *ideal* \( (S) \) generated by a subset \( S \subseteq \mathbb{C}[x] \) is

\[
\{ a_1(x)f_1(x) + \cdots + a_r(x)f_r(x) \mid r \in \mathbb{N}, \forall i \ a_i(x) \in \mathbb{C}[x], f_i(x) \in S \}
\]

**Theorem (Hilbert 1890)**

*every ideal \( I \subseteq \mathbb{C}[x] \) is generated by a finite set (\( \mathbb{C}[x] \) is noetherian)*

If \( S \subseteq \mathbb{C}[x] \), then \( (S) = (T) \) for some finite subset \( T \subseteq S \)

Hilbert used this theorem to prove a his Finiteness Theorem in Invariant Theory (discussed later)
suppose $V = \mathbb{C}^n$ is a representation of a group $G$
this means that every $g \in G$ acts by some $n \times n$ matrix
$M_g : V \rightarrow V$ (so $g \cdot v = M_g v$)
suppose \( V = \mathbb{C}^n \) is a representation of a group \( G \)
this means that every \( g \in G \) acts by some \( n \times n \) matrix
\( M_g : V \to V \) (so \( g \cdot v = M_g v \)) and we have \( M_e = I \) and
\( M_{gh} = M_g M_h \)
this also implies \( M_g^{-1} = (M_g)^{-1} \)
suppose $V = \mathbb{C}^n$ is a representation of a group $G$
this means that every $g \in G$ acts my some $n \times n$ matrix
$M_g : V \to V$ (so $g \cdot v = M_g v$) and we have $M_e = I$ and
$M_{gh} = M_g M_h$
this also implies $M_g^{-1} = (M_g)^{-1}$

if $f(x) \in \mathbb{C}[x]$ and $M = (m_{i,j})$ is $n \times n$ matrix, then $v \mapsto f(Mv)$ is
a polynomial function given by the formula

$$f\left(\sum_{j=1}^{n} m_{1,j}x_j, \ldots, \sum_{j=1}^{n} m_{n,j}x_j\right)$$
$G$ acts on $\mathbb{C}[x]$ as follows:

if $g \in G$ and $f(x) \in \mathbb{C}[x]$ then define $(g \cdot f)(x) \in \mathbb{C}[x]$ by

$$(g \cdot f)(v) = f(M_{g^{-1}}v)$$

(we use $M_{g^{-1}}$ instead of $M_g$ to make it a left action)
Action of a Group $G$

$G$ acts on $\mathbb{C}[x]$ as follows:

if $g \in G$ and $f(x) \in \mathbb{C}[x]$ then define $(g \cdot f)(x) \in \mathbb{C}[x]$ by

$$(g \cdot f)(v) = f(M_g^{-1} v)$$

(we use $M_{g^{-1}}$ instead of $M_g$ to make it a left action)

$\mathbb{C}[x]$ is an $\infty$-dimensional $\mathbb{C}$-vector space

the monomials form a basis

$G$ acts by linear transformations on $\mathbb{C}[x]

$\mathbb{C}[x]$ is an $\infty$-dimensional representation of $G$
The Invariant Ring

\[ f(x) \in \mathbb{C}[x] \text{ is } G\text{-invariant if } (g \cdot f)(x) = f(x) \text{ for all } g \in G \]

\[ f(x) \in \mathbb{C}[x] \text{ is } G\text{-invariant if and only if it is constant on all } G\text{-orbits in } V \]
The Invariant Ring

\[ f(x) \in \mathbb{C}[x] \text{ is } G\text{-invariant if } (g \cdot f)(x) = f(x) \text{ for all } g \in G \]
\[ f(x) \in \mathbb{C}[x] \text{ is } G\text{-invariant if and only if it is constant on all } G\text{-orbits in } V \]

**Definition**

\( \mathbb{C}[x]^G \) is the set of all \( G\)-invariant polynomials in \( \mathbb{C}[x] \)

\( \mathbb{C}[x]^G \) is a *subalgebra*, i.e., contains \( \mathbb{C} \) and is closed under addition, subtraction and multiplication
The Invariant Ring

\( f(x) \in \mathbb{C}[x] \) is \( G \)-invariant if \( (g \cdot f)(x) = f(x) \) for all \( g \in G \)

\( f(x) \in \mathbb{C}[x] \) is \( G \)-invariant if and only if it is constant on all \( G \)-orbits in \( V \)

### Definition

\( \mathbb{C}[x]^G \) is the set of all \( G \)-invariant polynomials in \( \mathbb{C}[x] \)

\( \mathbb{C}[x]^G \) is a subalgebra, i.e., contains \( \mathbb{C} \) and is closed under addition, subtraction and multiplication

if \( f_1(x), \ldots, f_r(x) \in \mathbb{C}[x] \) then

\[
\mathbb{C}[f_1(x), \ldots, f_r(x)] := \{ p(f_1(x), \ldots, f_r(x)) \mid p(y_1, \ldots, y_r) \in \mathbb{C}[y_1, \ldots, y_r] \}
\]

is the subalgebra of \( \mathbb{C}[x] \) generated by \( f_1(x), \ldots, f_r(x) \).
The Symmetric Group

\[ G = S_n \] acts on \( V = \mathbb{C}^n \) by permuting the coordinates for \( \sigma \in S_n \). \( M_\sigma \) is the corresponding permutation matrix.

\( S_n \) acts on \( \mathbb{C}[x] \) as

\[
(\sigma \cdot f)(x_1, \ldots, x_n) = f(x_\sigma(1), \ldots, x_\sigma(n))
\]

Theorem

\( \mathbb{C}[x] \) \( S_n = \mathbb{C}[e_1(x), \ldots, e_n(x)] \)
The Symmetric Group

\( G = S_n \) acts on \( V = \mathbb{C}^n \) by permuting the coordinates for \( \sigma \in S_n \), \( M_\sigma \) is the corresponding permutation matrix.

\( S_n \) acts on \( \mathbb{C}[x] \) as

\[
(\sigma \cdot f)(x_1, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)})
\]

define the \( k \)-th elementary symmetric function as

\[
e_k(x) = \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} x_{i_1} x_{i_2} \cdots x_{i_k}
\]

for example \( e_1 = x_1 + x_2 + \cdots + x_n \) and \( e_n = x_1 x_2 \cdots x_n \)
The Symmetric Group

$G = S_n$ acts on $V = \mathbb{C}^n$ by permuting the coordinates for $\sigma \in S_n$, $M_\sigma$ is the corresponding permutation matrix $S_n$ acts on $\mathbb{C}[x]$ as

$$(\sigma \cdot f)(x_1, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

define the $k$-th elementary symmetric function as

$$e_k(x) = \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} x_{i_1} x_{i_2} \cdots x_{i_k}$$

for example $e_1 = x_1 + x_2 + \cdots + x_n$ and $e_n = x_1 x_2 \cdots x_n$

**Theorem**

$\mathbb{C}[x]^{S_n} = \mathbb{C}[e_1(x), \ldots, e_n(x)]$
assume that $G$ is (linearly) reductive, which means that every representation of $G$ is a direct sum of irreducible representations. Examples are $\text{GL}_n$, $\text{SL}_n$, $\text{O}_n$, finite groups.
assume that $G$ is (linearly) reductive, which means that every representation of $G$ is a direct sum of irreducible representations. Examples are $\text{GL}_n$, $\text{SL}_n$, $\text{O}_n$, finite groups.

**Theorem (Hilbert 1890)**

$\mathbb{C}[x]^G$ is a finitely generated algebra, i.e.,

$\mathbb{C}[x]^G = \mathbb{C}[f_1(x), \ldots, f_r(x)]$ for some $r < \infty$ and $f_1(x), \ldots, f_r(x) \in \mathbb{C}[x]^G$
Hilbert’s Finiteness Theorem

assume that $G$ is (linearly) reductive, which means that every representation of $G$ is a direct sum of irreducible representations. Examples are $\text{GL}_n$, $\text{SL}_n$, $\text{O}_n$, finite groups.

**Theorem (Hilbert 1890)**

$\mathbb{C}[x]^G$ is a finitely generated algebra, i.e.,

$\mathbb{C}[x]^G = \mathbb{C}[f_1(x), \ldots, f_r(x)]$ for some $r < \infty$ and $f_1(x), \ldots, f_r(x) \in \mathbb{C}[x]^G$.

**Proof sketch:**

$J \subseteq \mathbb{C}[x]$ ideal generated by all homogeneous, non-constant $f(x) \in \mathbb{C}[x]^G$ ($\infty$ many!)
Hilbert’s Finiteness Theorem

assume that $G$ is (linearly) reductive, which means that every representation of $G$ is a direct sum of irreducible representations. Examples are $\text{GL}_n$, $\text{SL}_n$, $\text{O}_n$, finite groups.

**Theorem (Hilbert 1890)**

$\mathbb{C}[x]^G$ is a finitely generated algebra, i.e.,

$\mathbb{C}[x]^G = \mathbb{C}[f_1(x), \ldots, f_r(x)]$ for some $r < \infty$ and $f_1(x), \ldots, f_r(x) \in \mathbb{C}[x]^G$.

**proof sketch:**

$J \subseteq \mathbb{C}[x]$ ideal generated by all homogeneous, non-constant $f(x) \in \mathbb{C}[x]^G$ ($\infty$ many!)

Basis Theorem: $J = (f_1(x), \ldots, f_r(x))$ for some $r < \infty$ and homogeneous $f_1(x), \ldots, f_r(x) \in \mathbb{C}[x]^G$.
assume that $G$ is (linearly) reductive, which means that every representation of $G$ is a direct sum of irreducible representations. Examples are $\text{GL}_n$, $\text{SL}_n$, $\text{O}_n$, finite groups.

**Theorem (Hilbert 1890)**

$\mathbb{C}[x]^G$ is a finitely generated algebra, i.e.,

$\mathbb{C}[x]^G = \mathbb{C}[f_1(x), \ldots, f_r(x)]$ for some $r < \infty$ and

$f_1(x), \ldots, f_r(x) \in \mathbb{C}[x]^G$

proof sketch:

$J \subseteq \mathbb{C}[x]$ ideal generated by all homogeneous, non-constant $f(x) \in \mathbb{C}[x]^G$ ($\infty$ many!)

Basis Theorem: $J = (f_1(x), \ldots, f_r(x))$ for some $r < \infty$ and homogeneous $f_1(x), \ldots, f_r(x) \in \mathbb{C}[x]^G$

by induction one shows that $\mathbb{C}[x]^G = \mathbb{C}[f_1(x), \ldots, f_r(x)]$
Definition

$\beta(\mathbb{C}[x]^G)$ is the smallest $d$ such that $\mathbb{C}[x]^G$ is generated by polynomials of degree $\leq d$
Degree Bounds

Definition

\(\beta(\mathbb{C}[x]^G)\) is the smallest \(d\) such that \(\mathbb{C}[x]^G\) is generated by polynomials of degree \(\leq d\)

Theorem (Jordan 1876)

For binary forms of degree \(d\) we have \(\beta(\mathbb{C}[x_1,\ldots,x_d+1]^{SL_2}) \leq d^6\)
Degree Bounds

**Definition**

\[ \beta(\mathbb{C}[x]^G) \text{ is the smallest } d \text{ such that } \mathbb{C}[x]^G \text{ is generated by polynomials of degree } \leq d \]

**Theorem (Jordan 1876)**

*for binary forms of degree \( d \) we have* \( \beta(\mathbb{C}[x_1, \ldots, x_{d+1}]^{SL_2}) \leq d^6 \)

**Theorem (Emmy Noether 1916)**

*if \( G \) is finite then* \( \beta(\mathbb{C}[x]^G) \leq |G| \)
the proof of Hilbert’s finiteness theorem does not give an algorithm for finding generators, nor does it give an upper bound for $\beta(\mathbb{C}[x]^G)$ for arbitrary $G$. 
the proof of Hilbert’s finiteness theorem does not give an algorithm for finding generators, nor does it give an upper bound for $\beta(\mathbb{C}[x]^G)$ for arbitrary $G$

so Hilbert gave another, more constructive proof in 1893 of his Finiteness Theorem using his notion of the null cone
Hilbert’s Null cone

for $v \in V$, $G \cdot v = \{g \cdot v \mid g \in G\}$ is orbit of $v$

$G \cdot v \subseteq V$ closure of the orbit

Theorem

$G \cdot v \cap G \cdot w \neq \emptyset \iff f(v) = f(w)$ for all $f(x) \in \mathbb{C}[x]^G$
Hilbert’s Null cone

for $v \in V$, $G \cdot v = \{ g \cdot v \mid g \in G \}$ is orbit of $v$
$G \cdot v \subseteq V$ closure of the orbit

**Theorem**

$G \cdot v \cap G \cdot w \neq \emptyset \iff f(v) = f(w)$ for all $f(x) \in \mathbb{C}[x]^G$

$\Rightarrow: f \in \mathbb{C}[x]^G$ is constant on $G \cdot v$ and $G \cdot w$
Hilbert’s Null cone

for \( v \in V \), \( G \cdot v = \{ g \cdot v \mid g \in G \} \) is orbit of \( v \)

\( \overline{G \cdot v} \subseteq V \) closure of the orbit

**Theorem**

\( \overline{G \cdot v} \cap \overline{G \cdot w} \neq \emptyset \iff f(v) = f(w) \) for all \( f(x) \in \mathbb{C}[x]^G \)

\( \Rightarrow: f \in \mathbb{C}[x]^G \) is constant on \( \overline{G \cdot v} \) and \( \overline{G \cdot w} \)

**Definition**

Hilbert’s Null cone:

\[
\mathcal{N} := \{ v \in V \mid 0 \in \overline{G \cdot v} \} = \\
= \{ v \in V \mid f(v) = f(0) \text{ for all } f(x) \in \mathbb{C}[x]^G \}
\]

if \( \mathbb{C}[x]^G = \mathbb{C}[f_1(x), \ldots, f_r(x)] \) with \( f_1(x), \ldots, f_r(x) \) homogeneous, non-constant, then \( \mathcal{N} = \{ v \in V \mid f_1(v) = \cdots = f_r(v) = 0 \} \)
Example: Multiplicative Group

\[ G = \mathbb{C}^*, \ V = \mathbb{C}^4 \]

for \( t \in \mathbb{C}^* \), define

\[
M_t = \begin{pmatrix}
t & 0 & 0 & 0 \\
0 & t & 0 & 0 \\
0 & 0 & t^{-1} & 0 \\
0 & 0 & 0 & t^{-1}
\end{pmatrix}
\]

\[
t \cdot \begin{pmatrix}
v_1 \\
v_2 \\
v_3 \\
v_4
\end{pmatrix} = M_t \begin{pmatrix}
v_1 \\
v_2 \\
v_3 \\
v_4
\end{pmatrix} = \begin{pmatrix}
tv_1 \\
tv_2 \\
t^{-1}v_3 \\
t^{-1}v_4
\end{pmatrix}
\]

\[ \mathcal{N} = \{ v_1 = v_2 = 0 \} \cup \{ v_3 = v_4 = 0 \} \]
Example: Multiplicative Group

\[ \mathbb{C}[x_1, x_2, x_3, x_4]^{\mathbb{C}^*} = \mathbb{C}[x_1 x_3, x_1 x_4, x_2 x_3, x_2 x_4] \]

\[ N = \{ v_1 v_3 = v_1 v_4 = v_2 v_3 = v_2 v_4 = 0 \} = \{ v_1 = v_2 = 0 \} \cup \{ v_3 = v_4 = 0 \} \]

Note that in this case, there is an algebraic relation between the generators, namely

\[ (x_1 x_3)(x_2 x_4) = (x_1 x_4)(x_2 x_3) \]
Hilbert-Mumford criterion

Definition

A one parameter subgroup (1-PSG) is a homomorphism of algebraic groups \( \lambda : \mathbb{C}^* \rightarrow G \).
**Definition**

a one parameter subgroup (1-PSG) is a homomorphism of algebraic groups $\lambda : \mathbb{C}^* \rightarrow G$

**Theorem (Hilbert-Mumford criterion)**

If $v \in V = \mathbb{C}^n$, then $v \in \mathcal{N} \iff$ there exists a 1-PSG $\lambda : \mathbb{C}^* \rightarrow G$ with $\lim_{t \to 0} \lambda(t) \cdot v = 0$
Conjugation of $n \times n$ Matrices

$V = \text{Mat}_{n,n}$, the space of $n \times n$ matrices
$G = \text{GL}_n$ (the group of invertible $n \times n$ matrices) acts by conjugation: if $A = (a_{i,j}) \in V$ and $g \in G$ then $g \cdot A = gAg^{-1}$
Conjugation of $n \times n$ Matrices

\[ V = \text{Mat}_{n,n}, \text{ the space of } n \times n \text{ matrices} \]
\[ G = \text{GL}_n \text{ (the group of invertible } n \times n \text{ matrices) acts by conjugation: if } A = (a_{i,j}) \in V \text{ and } g \in G \text{ then } g \cdot A = gAg^{-1} \]

if

\[
(\star) \quad \lambda(t) = \begin{pmatrix}
t^{k_1} \\
\vdots \\
t^{k_n}
\end{pmatrix}
\]

with $k_1 \geq k_2 \geq \cdots \geq k_n$, then

\[
\lambda(t) \cdot A = \lambda(t)A\lambda(t)^{-1} = (t^{k_i-k_j} a_{i,j}).
\]

so $\lim_{t \to 0} \lambda(t) \cdot A = 0$ if and only if $A$ is strict upper triangular.
Conjugation of $n \times n$ Matrices

$V = \text{Mat}_{n,n}$, the space of $n \times n$ matrices
$G = \text{GL}_n$ (the group of invertible $n \times n$ matrices) acts by conjugation: if $A = (a_{i,j}) \in V$ and $g \in G$ then $g \cdot A = gAg^{-1}$

if

\[(\ast) \quad \lambda(t) = \begin{pmatrix} t^{k_1} \\ \vdots \\ t^{k_n} \end{pmatrix} \]

with $k_1 \geq k_2 \geq \cdots \geq k_n$, then

\[\lambda(t) \cdot A = \lambda(t)A\lambda(t)^{-1} = (t^{k_i-k_j}a_{i,j}).\]

so $\lim_{t \to 0} \lambda(t) \cdot A = 0$ if and only if $A$ is strict upper triangular

every 1-PSG is of the form $(\ast)$ after a base change, so

$A \in \mathcal{N} \iff A$ conjugate to strict upper triang. mat. $\iff A$ is nilpotent
Conjugation of $n \times n$ Matrices

$X = (x_{i,j})$ where $x_{i,j}$ are indeterminates

$$\det(tI - X) = t^n - f_1(x)t^{n-1} + \cdots + (-1)^n f_n(x)$$

where $x = x_{1,1}, x_{1,2}, \ldots, x_{n,n}$

$f_1(A) = \text{trace}(A), f_n(A) = \det(A)$
Conjugation of $n \times n$ Matrices

$X = (x_{i,j})$ where $x_{i,j}$ are indeterminates

$$\det(tI - X) = t^n - f_1(x)t^{n-1} + \cdots + (-1)^n f_n(x)$$

where $x = x_{1,1}, x_{1,2}, \ldots, x_{n,n}$
$f_1(A) = \text{trace}(A), f_n(A) = \det(A)$

**Theorem**

$$\mathbb{C}[x]^G = \mathbb{C}[f_1(x), \ldots, f_n(x)]$$

$A \in \mathcal{N} \iff f_1(A) = \cdots = f_n(A) = 0 \iff \det(tI - A) = t^n \iff A$ nilpotent
Theorem (Hilbert 1893)

suppose $f_1(x), \ldots, f_r(x) \in \mathbb{C}[x]^G$ are homogeneous and

$\mathcal{N} = \{ v \mid f_1(v) = \cdots = f_r(v) = 0 \}$
Degree Bounds

**Theorem (Hilbert 1893)**

Suppose $f_1(x), \ldots, f_r(x) \in \mathbb{C}[x]^G$ are homogeneous and $N = \{v \mid f_1(v) = \cdots = f_r(v) = 0\}$

Then there exists finitely many homogeneous invariants $h_1(x), \ldots, h_s(x)$ such that every invariant $p(x) \in \mathbb{C}[x]^G$ is of the form

$$p(x) = a_1(x)h_1(x) + \cdots + a_s(x)h_s(x)$$

For some $a_1(x), \ldots, a_s(x) \in \mathbb{C}[f_1(x), \ldots, f_r(x)]$
Theorem (Popov 1980)

suppose $f_1(x), \ldots, f_r(x) \in \mathbb{C}[x]^G$ are homogeneous of the same degree $d$ and $\mathcal{N} = \{v \mid f_1(v) = \cdots = f_r(v) = 0\}$
Theorem (Popov 1980)

Suppose \( f_1(x), \ldots, f_r(x) \in \mathbb{C}[x]^G \) are homogeneous of the same degree \( d \) and \( \mathcal{N} = \{ v \mid f_1(v) = \cdots = f_r(v) = 0 \} \). Then there exists finitely many homogeneous invariants \( h_1(x), \ldots, h_s(x) \) of degree at most \( n(d - 1) \) such that every invariant \( p(x) \in \mathbb{C}[x]^G \) is of the form

\[
p(x) = a_1(x)h_1(x) + \cdots + a_s(x)h_s(x)
\]

for some \( a_1(x), \ldots, a_s(x) \in \mathbb{C}[f_1(x), \ldots, f_r(x)] \).
Theorem (Popov 1980)

suppose \( f_1(x), \ldots, f_r(x) \in \mathbb{C}[x]^G \) are homogeneous of the same
degree \( d \) and \( \mathcal{N} = \{ v \mid f_1(v) = \cdots = f_r(v) = 0 \} \)

then there exists finitely many homogeneous invariants
\( h_1(x), \ldots, h_s(x) \) of degree at most \( n(d - 1) \) such that every
invariant \( p(x) \in \mathbb{C}[x]^G \) is of the form

\[
p(x) = a_1(x)h_1(x) + \cdots + a_s(x)h_s(x)
\]

for some \( a_1(x), \ldots, a_s(x) \in \mathbb{C}[f_1(x), \ldots, f_r(x)] \)

in particular, \( \beta(\mathbb{C}[x]^G) \leq \max\{d, n(d - 1)\} \leq nd \)
(because \( \mathbb{C}[x]^G = \mathbb{C}[f_1(x), \ldots, f_r(x), h_1(x), \ldots, h_s(x)] \))
Theorem (D.)

Suppose \( f_1(x), \ldots, f_r(x) \in \mathbb{C}[x]^G \) are homogeneous of the degree at most \( d \) and \( \mathcal{N} = \{ v \mid f_1(v) = \cdots = f_r(v) = 0 \} \)
Theorem (D.)

Suppose $f_1(x), \ldots, f_r(x) \in \mathbb{C}[x]^G$ are homogeneous of the degree at most $d$ and $N = \{v \mid f_1(v) = \cdots = f_r(v) = 0\}$

Then we have

$$\beta(\mathbb{C}[x]^G) \leq \frac{3}{8} nd^2$$
Theorem (D.)

Suppose $f_1(x), \ldots, f_r(x) \in \mathbb{C}[x]^G$ are homogeneous of the degree at most $d$ and $\mathcal{N} = \{v \mid f_1(v) = \cdots = f_r(v) = 0\}$

then we have

$$\beta(\mathbb{C}[x]^G) \leq \frac{3}{8} nd^2$$

for binary forms of degree $n$, the null cone is defined by homogeneous invariants of degree $\leq 2n^3$, and we get

$$\beta(\mathbb{C}[x]^G) \leq \frac{3}{8} (n + 1)(2n^3)^2$$

(which is slightly worse than Jordan’s bound)
Theorem (D.)

suppose \( f_1(x), \ldots, f_r(x) \in \mathbb{C}[x]^G \) are homogeneous of the degree at most \( d \) and \( \mathcal{N} = \{ v \mid f_1(v) = \cdots = f_r(v) = 0 \} \)

then we have

\[
\beta(\mathbb{C}[x]^G) \leq \frac{3}{8} nd^2
\]

for binary forms of degree \( n \), the null cone is defined by homogeneous invariants of degree \( \leq 2n^3 \), and we get

\[
\beta(\mathbb{C}[x]^G) \leq \frac{3}{8} (n + 1)(2n^3)^2
\]

(which is slightly worse than Jordan’s bound)

if \( G \) is fixed then the bound \( \beta(\mathbb{C}[x]^G) \) is polynomial in \( n \) (the dimension of \( V \)) and the largest euclidean length among the weights appearing in the representation
Polarization

$V$ representation of $G$

$\mathbb{C}[x] = \mathbb{C}[x_1, \ldots, x_n]$ ring of polynomial functions on $V = \mathbb{C}^n$

$\mathbb{C}[x, y]$ of polynomial functions on $V \oplus V = \mathbb{C}^{2n}$

for $f(x) \in \mathbb{C}[x]^G$ we can write

$$f(x_1 + ty_1, \ldots, x_n + ty_n) = f_0(x, y) + f_1(x, y)t + \cdots + f_d(x, y)t^d$$

where $f_0(x, y), \ldots, f_d(x, y) \in \mathbb{C}[x, y]^G$
Polarization

\[ R[m] := \mathbb{C}[x_{1,1}, \ldots, x_{n,1}, \ldots, x_{1,m}, \ldots, x_{n,m}] \] is the ring of polynomial functions on \( V^m \cong \text{Mat}_{n,m} \).
Polarization

\[ R[m] := \mathbb{C}[x_{1,1}, \ldots, x_{n,1}, \ldots, x_{1,m}, \ldots, x_{n,m}] \] is the ring of polynomial functions on \( V^m \cong \text{Mat}_{n,m} \)

if \( m < s \) then we can polarize \( f(x) \in R[m]^G \) to get invariants in \( R[s]^G \)

**Theorem (Weyl)**

if \( s > n = \dim V \) then polarizing generators from \( R[n]^G \) give generators of \( R[s]^G \).
Polarization

\[ R[m] := \mathbb{C}[x_{1,1}, \ldots, x_{n,1}, \ldots, x_{1,m}, \ldots, x_{n,m}] \] is the ring of polynomial functions on \( V^m \cong \text{Mat}_{n,m} \)

if \( m < s \) then we can polarize \( f(x) \in R[m]^G \) to get invariants in \( R[s]^G \)

Theorem (Weyl)

if \( s > n = \text{dim} V \) then polarizing generators from \( R[n]^G \) give generators of \( R[s]^G \).

in particular, \( \beta(R[s]^G) = \beta(R[n]^G) \)
instead of $\mathbb{C}$, we can take any algebraically closed field of characteristic 0
Instead of $\mathbb{C}$, we can take any algebraically closed field of characteristic 0. The degree bounds are valid for arbitrary fields of characteristic 0. We need “algebraically closed” to make geometric statements about the null cone, orbits, etc.
instead of $\mathbb{C}$, we can take any algebraically closed field of characteristic 0

the degree bounds are valid for arbitrary fields of characteristic 0
we need “algebraically closed” to make geometric statements about the null cone, orbits, etc.

most statements are either false, or more difficult to prove in positive characteristic
Weyl’s theorem is false in positive characteristic
Weyl’s theorem is false in positive characteristic

Noether’s bound ($\beta(\mathbb{C}[x]^G) \leq |G|$ for finite $G$) is wrong in positive characteristic
Other Fields

Weyl’s theorem is false in positive characteristic

Noether’s bound \( (\beta(\mathbb{C}[x]^G) \leq |G| \) for finite \( G \) \) is wrong in positive characteristic

invariant rings of reductive groups are also finitely generated in positive characteristic, but the proof is harder (using theorems of Nagata and Haboush) and many of the geometric statements about the null cone etc. are still true