# Algorithms in Invariant Theory

With 0 Figures

Anakin Dey

Last Edited on 5/7/25 at 15:27

# Contents

# Preface

A core interest of mine (at the time of writing this) is algorithms in the context of commutative algebra and algebraic geometry. As such, Bernd Sturmfel's text *Algorithms in Invariant Theory* [Str08] is a good resource for first learning this stuff. Perhaps in the future I'll include notes and some source code

# Chapter 1

# Introduction

## 1.1 Symmetric Polynomials

**Solution: [Str08] 1.1.5:** From [Page 23, 2.11'][Mac98], we have the following identities

$$
\begin{aligned}
1 \cdot \sigma_1 &= p_1, \\
2 \cdot \sigma_2 &= p_1\sigma_1 - p_2, \\
3 \cdot \sigma_3 &= p_1\sigma_2 - p_2\sigma_1 + p_3, \\
&\vdots \\
k \cdot \sigma_k &= \sum_{r=1}^{k} (-1)^{r-1} p_r \sigma_{k-r}.
\end{aligned}
$$

Treating the $\sigma_i$ as indeterminants, we can re-express the above system of equations:

$$
\begin{aligned}
p_1 &= 1 \cdot \sigma_1, \\
p_2 &= p_1\sigma_1 - 2 \cdot \sigma_2, \\
p_3 &= p_2\sigma_1 - p_1\sigma_2 + 3 \cdot \sigma_3, \\
&\vdots \\
p_k &= \left( \sum_{r=1}^{k-1} (-1)^{r-1} p_{k-r} \sigma_r \right) + (-1)^{k-1} \cdot k\sigma_k.
\end{aligned}
$$

Consider $\sigma_1, -\sigma_2, \sigma_3, \ldots, (-1)^n\sigma_{n-1}, \sigma_n$ as indeterminants. Thus, we obtain the following matrix equation:

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ p_1 & 2 & 0 & \cdots & 0 \\ p_2 & p_1 & 3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ p_{k-1} & p_{k-2} & \cdots & \cdots & (-1)^k \cdot k \end{pmatrix} \begin{pmatrix} \sigma_1 \\ -\sigma_2 \\ \vdots \\ (-1)^k \sigma_{k-1} \\ \sigma_k \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_{k-1} \\ p_k \end{pmatrix}.$$

Then, Cramer's rule yields that

$$\begin{aligned}
\sigma_k &= \det \begin{pmatrix} 1 & 0 & 0 & \cdots & p_1 \\ p_1 & 2 & 0 & \cdots & p_2 \\ p_2 & p_1 & 3 & \cdots & p_3 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ p_{k-1} & p_{k-2} & \cdots & \cdots & p_k \end{pmatrix} \det \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ p_1 & 2 & 0 & \cdots & 0 \\ p_2 & p_1 & 3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ p_{k-1} & p_{k-2} & \cdots & \cdots & (-1)^k \cdot k \end{pmatrix}^{-1} \\
&= \frac{(-1)^k}{k!} \det \begin{pmatrix} 1 & 0 & 0 & \cdots & p_1 \\ p_1 & 2 & 0 & \cdots & p_2 \\ p_2 & p_1 & 3 & \cdots & p_3 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ p_{k-1} & p_{k-2} & \cdots & \cdots & p_k \end{pmatrix} \\
&= \frac{(-1)^k}{k!} \cdot (-1)^k \det \begin{pmatrix} p_1 & 1 & 0 & \cdots & 0 \\ p_2 & p_1 & 2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ p_{k-1} & p_{k-2} & \cdots & p_1 & k-1 \\ p_k & p_{k-1} & \cdots & \cdots & p_1 \end{pmatrix} = \frac{1}{k!} \det \begin{pmatrix} p_1 & 1 & 0 & \cdots & 0 \\ p_2 & p_1 & 2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ p_{k-1} & p_{k-2} & \cdots & p_1 & k-1 \\ p_k & p_{k-1} & \cdots & \cdots & p_1 \end{pmatrix}.
\end{aligned}$$

□

## 1.2 Gröbner Bases

**Lemma 1.2.1**:

Let $R = \mathbb{C}[x_1, \dots, x_n]$. Then with the usual grading, let $H(R,z) \coloneqq \sum_{d=0}^{\infty} \dim_{\mathbb{C}}(R_d) z^d$. We have that

$$H(R,z) \coloneqq \sum_{d=0}^{\infty} \dim_{\mathbb{C}}(R_d) z^d = \sum_{d=0}^{\infty} \binom{d+n-1}{n-1} = \frac{1}{(1-z)^n}.$$

**Proof**: To see that $H(R,z) = \sum_{d=0}^{\infty} \binom{d+n-1}{n-1}$, just count the number of monomials of degree $d$ in $n$ variables. The value $\binom{d+n-1}{n-1}$ is the number of non-negative integer solutions to $a_1 + \cdots + a_n = d$. Each solution corresponds to a monomial $x_1^{a_1} \cdots x_n^{a_n}$. Then to see that $H(R,z) = \frac{1}{(1-z)^n}$, consider the product of infinite sums $(1 + z + z^2 + \cdots) \cdots (1 + z + z^2 + \cdots)$ a total of $n$-times. Then the coefficient of $z^d$ again corresponds to the number of such non-negative integer solutions. Since $\frac{1}{1-z} = 1 + z + z^2 + \cdots$, we obtain the desired equality. □

**Lemma 1.2.2**:

For $1 \leq k \leq n$, we have that

$$h_k(x_k, \dots, x_n) + \sum_{i=1}^{k} (-1)^i h_{k-i}(x_k, \dots, x_n) \sigma_i(x_1, \dots, x_n) = 0.$$

**Proof**: Using the generating functions for the $h_i$ and $\sigma_i$, we have that the above expression is the coefficient of $t^k$ in the product

$$\prod_{i=k}^{n} (1 - x_i t)^{-1} \cdot \prod_{i=1}^{n} (1 - x_i t) = \prod_{i=1}^{k-1} (1 - x_i t).$$

However, the right-hand side of this has degree $k-1$ in $t$. Thus, the coefficient of $t^k$ is indeed 0. □

**Solution: [Str08] 1.2.1**: Let $M$ be a set of monomial generators for $\operatorname{init}(I)$ and let $m$ be minimally nonstandard. Since $m$ is a monomial and in $\operatorname{init}(I)$, we have that $m' \mid m$ for some monomial $m' \in M$. However, note that $m' \in \operatorname{init}(I)$ and furthermore by the fact that $m$ is minimally nonstandard, we cannot have that $m'$ strictly divides $m$. Thus, $m' = m$ and $m \in M$. Then by Dickson's Lemma, we have that $M$ is a finite set and thus there are finitely many minimally nonstandard monomials. □

**Solution: [Str08] 1.2.2**: This is [CLO15, Chapter 2, §7, Theorem 5]. □

**Solution: [Str08] 1.2.3**: This is [CLO15, Chapter 3, §1, Theorem 2]. □

**Solution: [Str08] 1.2.4:** This is following [Rob85] and [GP07]. Let $\geq$ be a monomial ordering on $\mathbb{C}[x_1,\dots,x_n]$. This is equivalent to a total semigroup ordering $\geq$ on $\mathbb{Z}^n$. Such a semigroup ordering gives a unique total ordering on $\mathbb{Q}^n$. To see this, for $\overline{q}=(q_1,\dots,q_n)\in\mathbb{Q}^n$, let $m\in\mathbb{Z}$ such that $m\cdot q_i\in\mathbb{Z}$ for all $i$. Then say that $\overline{q}\geq 0$ if and only if $m\cdot\overline{q}\geq 0$ where the latter ordering is in $\mathbb{Z}^n$.

Let $V\subseteq\mathbb{Q}^n$ be a $\mathbb{Q}$-vector space with $\dim_{\mathbb{Q}}(V)=r$. Then let

$$V_0 := \{\, z\in\mathbb{R}^n \mid \forall\varepsilon>0, \exists\, z_+(\varepsilon), z_-(\varepsilon)\in V\cap B_\varepsilon(z) \text{ such that } z_+(\varepsilon)>0, z_-(\varepsilon)<0\,\}.$$

Then $V_0$ is clearly a $\mathbb{R}$-subspace of $\mathbb{R}^n$. With the ordering $\geq$ on $\mathbb{Q}^n$, we can define $V_+$ and $V_-$ depending on if $\overline{q}\geq 0$ or $\overline{q}<0$. We define a map $\pi\colon V\to\{-1,1\}$, where $V$ has the Euclidean topology and $\{-1,1\}$ has the discrete topology. Let $\pi(q)=1$ if there exists an open ball $U_\varepsilon(q)$ such that $U_\varepsilon(q)\cap V\subseteq V_+$ and $\pi(q)=-1$ if there exists an open ball $U_\varepsilon(q)$ such that $U_\varepsilon(q)\cap V\subseteq V_-$. Then $\pi$ is continuous. Recall that topological vector spaces over $\mathbb{R}$ are connected. Thus, we cannot have that $\dim_{\mathbb{R}} V_0 < r-1$ as if it were, then $V_{\mathbb{R}}\setminus V_0$ would be connected **⟨⟨ why is this bad? ⟩⟩** . Then suppose that $\dim_{\mathbb{R}} V_0 = r$. Then we have an ordered basis $e_1,\dots,e_r$ such that $e_i>0$ for all $i$. But then the linear combinations of the $e_i$ with positive coefficients are a subspace of $V_+$ which is a contradiction **⟨⟨ why? ⟩⟩** .

To construct the first row of the matrix, start with $V=\mathbb{Q}^n$ and consider the obtained $V_0$. Then the dimension 1 subspace orthogonal to $V_0$ in $\mathbb{R}^n$ defines the first row of $A$. We can continue this construction inductively to obtain the full matrix $A$. □

**Solution: [Str08] 1.2.6:** First, we recall some definitions. The $S$-polynomial of $f$ and $g$ is

$$S(f,g) := \frac{\operatorname{lcm}(\mathrm{LM}(f),\mathrm{LM}(g))}{\mathrm{LT}(f)} f - \frac{\operatorname{lcm}(\mathrm{LM}(f),\mathrm{LM}(g))}{\mathrm{LT}(g)} g.$$

For a set of polynomials $\mathcal{F}=\{f_1,\dots,f_t\}\subseteq k[x_1,\dots,x_n]$, we write $f\to_{\mathcal{F}}0$ if there exists $a_1,\dots,a_t\in k[x_1,\dots,x_n]$ such that $a_1f_1+\cdots+a_tf_t=0$. Then [CLO15, Chapter 2, §9, Theorem 3] says that a basis $\mathcal{F}=\{f_1,\dots f_t\}$ is a Gröbner basis for $G$ if and only if $S(f_i,f_j)\to_{\mathcal{F}}0$ for all $i\neq j$. But [CLO15, Chapter 2, §9, Proposition 4] says that for $f,g\in\mathcal{F}$ with relatively prime initial monomials, we have that $S(f,g)\to_{\mathcal{F}}0$. This proves the claim. □

### 1.3 What is Invariant Theory?

**Solution: [Str08] 1.3.1**: Let $\Gamma$ be a finite group. Consider $f(x) = \prod_{g\in\Gamma} g\cdot x$. Then $f$ is well defined as $\Gamma$ is finite, invariant under the action of $\Gamma$, and of degree $|\Gamma| > 0$.

Now suppose $\Gamma$ is the subgroup of matrices $\lambda I_n$ for $\lambda \in \mathbb{C}^\times$. Then for any polynomial $f(\overline{x}) = \sum_I \overline{a}^I \overline{x}^I \in \mathbb{C}[\overline{x}]^\Gamma$ and for any such $\lambda I_n \in \Gamma$, we have that

$$\sum_I \overline{a}^I \overline{x}^I = f(\overline{x}) = \lambda I_n \cdot f(\overline{x}) = \sum_I \overline{a}^I \lambda^{|I|} \overline{x}^I.$$

Then comparing coefficients, we deduce that $f(\overline{x})$ is fixed if and only if $f(\overline{x}) \in \mathbb{C}$. Thus, $\mathbb{C}[\overline{x}]^\Gamma = \mathbb{C}$. □

**Solution: [Str08] 1.3.3**: Fix $a_1, \ldots, a_n \in \mathbb{Z}$ and let $\Gamma = \{\operatorname{diag}(t^{a_1,\ldots,t^{a_n}}) \mid t \in \mathbb{C}^\times\}$. For $d \in \Gamma$ and a monomial $x_1^{\nu_1}\cdots x_n^{\nu_n}$, we have that $d\cdot x_1^{\nu_1}\cdots x_n^{\nu_n} = t^{a_1\nu_1+\cdots+a_n\nu_n} x_1^{\nu_1}\cdots x_n^{\nu_1}$. Thus, we want to determine the set of fixed exponent vectors

$$\mathcal{M} = \{(\nu_1,\ldots,\nu_n) \in \mathbb{Z}^n \mid \nu_1,\ldots,\nu_n \geq 0, a_1\nu_1 + \cdots + a_n\nu_n = 0\}.$$

This is exactly the object of student in §1.4, and in particular is solved by [Str08, Algorithm 1.4.5]. **⟨⟨ Is there a more direct way to see this? ⟩⟩** □

**Solution: [Str08] 1.3.4**: Recall that $\mathrm{GL}_n(\mathbb{C})$ is an affine algebraic subvariety of $\mathbb{A}_{\mathbb{C}}^{n^2+1}$. Consider the subspace of matrices in $\mathrm{GL}_n(\mathbb{C})$ which have distinct eigenvalues. Note that this is a Zariski open, and thus dense, subspace of $\mathrm{GL}_n(\mathbb{C})$. Indeed, let $A \in \mathrm{GL}_n(\mathbb{C})$ have eigenvalues $\lambda_1,\ldots,\lambda_n$. Then the discriminant of the characteristic polynomial $p_A$ of $A$ is $D(p_A) = \prod_{1\leq i<j\leq n}(\lambda_i - \lambda_j)$. Recall that the discriminant of a degree $d$ polynomial $f(x)$ is $\frac{(-1)^{\binom{d}{2}}}{\mathrm{LC}(f)} \operatorname{Res}_x(f(x), f'(x))$. For the characteristic polynomials, this is all expressible in terms of the entries of the matrix. Thus, the subspace of matrices in $\mathrm{GL}_n(\mathbb{C})$ is a dense open subset of $\mathrm{GL}_n(\mathbb{C})$. As this is an infinite set of matrices, any polynomial invariant on this dense subset must be invariant everywhere.

Let $f(\overline{X}) \in \mathbb{C}[\overline{x}]^{\mathrm{GL}_n(\mathbb{C})}$, where $\overline{X}$ is a matrix of indeterminates. Let $A$ have distinct eigenvalues, and so $A$ is diagonalizable so that there exists a matrix $M \in \mathrm{GL}_n(\mathbb{C})$ such $A = MDM^{-1}$ for some diagonal matrix $D$. In particular, the entries of $D$ are the eigenvalues of $A$. Thus, $f(A) = f(MDM^{-1}) = f(D)$. Furthermore, we may conjugate $A$ by permutation matrices to reorder the eigenvalues. Thus, $f$ must be a *symmetric* polynomial in the eigenvalues of $A$, denote these by $e_i := e_i(\lambda_1,\ldots,\lambda_n)$. Recall that via the characteristic polynomial, we can express these $e_i$ in terms of the entries of $A$ in general so that each $e_i \in \mathbb{C}[\overline{X}]$ and it makes sense to write that $\mathbb{C}[e_1,\ldots,e_n] \subseteq \mathbb{C}[\overline{X}]$. Thus, $\mathbb{C}[\overline{X}]^{\mathrm{GL}_n(\mathbb{C})} \subseteq \mathbb{C}[e_1,\ldots,e_n]$. Denote by $f_i(\overline{X})$ the coefficient of $t^i$ in $\det(tI_n - \overline{X})$. Then by noting that the characteristic polynomial is fixed under conjugation, and by comparing coefficients, we see that each $f_i$ is also fixed under conjugation. Thus, we overall have that each $e_i$ is fixed under conjugation and overall, $\mathbb{C}[\overline{X}]^{\mathrm{GL}_n(\mathbb{C})} = \mathbb{C}[e_1,\ldots,e_n]$. □

## 1.4 Torus Invariants and Integer Programming

**Solution: [Str08] 1.4.3**: With the addition of slack variables, we can without loss of generality compute a Hilbert basis for the monoid

$$\mathcal{M}'_{\mathcal{A}} = \{ \overline{\mu} \in \mathbb{Z}^d \mid \mathcal{A} \cdot \overline{\mu} = \overline{0} \} .$$

At a high level, we may use [Str08, Algorithm 1.4.5] multiple times to compute the Hilbert basis for $\mathcal{M}'_{\mathcal{A}}$. Of course if $\overline{\mu} = \overline{0}$ then $\mathcal{A} \cdot \overline{0} = \overline{0}$. Then for the nonzero case, we may divide $\mathbb{Z}^d$ into https://en.wikipedia.org/wiki/Orthant and apply [**Algorithm 1.4.5**] to each orthant. Then we can take the union over all the orthants of the Hilbert bases for each orthant to get a Hilbert basis for the whole space. This is still minimal because when defining a Hilbert basis, we care about *non-negative* integer linear combinations. □

# Bibliography


[CLO15] D.A. Cox, J. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Undergraduate Texts in Mathematics. Springer International Publishing, 2015. ISBN: 9783319167213. URL: https://books.google.com/books?id=yL7yCAAAQBAJ.

[GP07] G. Greuel and G. Pfister. *A Singular introduction to commutative algebra. With contributions by Olaf Bachmann, Christoph Lossen and Hans Sch"onemann. 2nd extended ed.* English. 2nd ed. Berlin: Springer, 2007, p. xx 689. URL: http://link.springer.com/book/10.1007%2F978-3-540-73542-7.

[Mac98] Ian Grant Macdonald. *Symmetric functions and Hall polynomials*. Oxford university press, 1998.

[Rob85] Lorenzo Robbiano. "Term Orderings on the Polynominal Ring." In: vol. 204. Apr. 1985, pp. 513–517. ISBN: 978-3-540-15984-1. DOI: 10.1007/3-540-15984-3_321.

[Str08] Bernd Strumfels. *Algorithms in Invariant Theory*. Springer Vienna, 2008. ISBN: 9783211774175. DOI: 10.1007/978-3-211-77417-5. URL: http://dx.doi.org/10.1007/978-3-211-77417-5.