Groebner Bases

Ryan Greenup

April 14, 2021

Groebner bases is one of the main practical tools for solving systems of polynomial equations.

1 Summary

Much of the theory of Groebner Basis is buried under needless amounts of abstract algebra, this is for the most part unnecessary and if I were to begin this investigation again I would first implement Buchberger's Algorithm, manually, using *Sympy* by referring to:

- Lectures 14 and 15 of Andreas Schulz OCW Integer Programming Course [4]
- Chapters 1-2 of *Ideals, Varieties and Algorithms* [9]
- Lecture 15 of Judy Holdenner's course on Algebraic Geometry [18]
- Lecture 14 of Pablo Parrilo's course on Algebraic Techniques [22]
- The *Sympy* source code for:
  - `polys.roebnertools` [29]
  - `solvers.polysys` [30]
- The *Sympy* documentation for *Polynomial Manipulation* [28]

Unfortunately this was not an option for me as these resources were not known to me until very late in the investigation. I hope that this report can serve as a guide for others who pick up this topic such that they can:

- Come to grips with the core concepts and practical applications quickly without wasting time on abstract algebra that is poorly explained [1]
- Identify useful resources that are written well and written with accessibility in mind

[1] In the absence of better materials a lot of time was wasted (yes, wasted, not spent) on complex algebraic concepts when all I needed was an algorithm to experiment with, an algorithm that the complex texts would not provide.

- Avoid material that serves as, for lack of a better word, as a red herring.

Although `sympy` is probably not the best tool for studying commutative algebra specifically (and the implementation may not be battle tested either, see e.g. [33]), the simple and accessible nature of `sympy` made it's documentation by far the most valuable resource for grappling with this topic.

An extension to this investigation would be to:

- Try and implement Buchberger's Algorithm from scratch using functions and iterations in *Python* in order to return a Reduced Groebner Basis
  - See Definition 4 of [9, §7]
- Try to demonstrate, in good detail, the relationship between the Euclidean Algorithm and Buchberger's Algorithm
  - See [9, p. 95]
- Try to implement Buchberger's Algorithm using *Normal Selection Strategy* [15, §3.1.2], see also [29, 23].

1.1 Further Resources

The following resources may be useful as reference material, but I would advice against using these as any sort of primary material, in order of recommendation (but not necessarily relevance)

1. Judson, T. W., & Open Textbook Library, Abstract algebra theory and applications [17]
2. Howlett, R., An undergraduate course in Abstract Algebra: Course notes for MATH3002 [25]
3. Lee, G., Abstract Algebra [12]
4. Grillet, P. A., Abstract Algebra [13]
5. Hibi, T., Grobner Bases: Statistics and Software Systems [15]
6. Adams, W. W., & Loustaunau, P., An introduction to Gröbner bases [2]
7. Nicodemi, O., Sutherland, M. A., & Towsley, G. W., An introduction to abstract algebra with notes to the future teacher [21]

Further material that I haven't had a chance to look throughA: includes:

- Becker, T., Weispfenning, V., & Kredel, H., Gröbner bases: a computational approach to commutative algebra [6]

2 Introduction

A Groebner Basis is a set of polynomials that spans the solution space of another set of polynomials, they are of interest to us because they are useful for solving systems of polynomial equations and provide a generalized theory that shows the relationships between:

- Polynomial Long Division with multiple variables and divisors, see e.g. [9, §3]
- The Division Algorithm see e.g. [9, §3] and [21]
- The LCM and GCD [9, §2.6]
- The Euclidean Algorithm and Gaussian Elimination
  - Both of which provide output that are special cases of Groebner Bases.

The theory of Groebner Bases even provides a framework to re-express the Fundamental Theorem of Algebra [23] .

3 Polynomials

Let K be some field (typically $\mathbb{Q}, \mathbb{R}, \mathbb{C}$, see §9.1.1 for more information)

3.1 Monomials

A *monomial* in the variables $x_1, x_2, \ldots x_n$ is given by: [15, p. 3]

$$\prod_{i=1}^{n} [x_i^{a_i}] = x_1^{1_1} \cdot x_2^{a_2} \cdot x_3^{a_3} \ldots x_n^{a_n} \quad : a \in \mathbb{Z}^+$$

Note however that a must be a non-negative integer [10, p. 48]

3.1.1 Degree

The degree is given by the sum of the exponents, so:

$$\deg\left(\prod_{i=1}^{n} [x_i^{a_i}]\right) = \sum_{i=1}^{n} [a_i]$$

3.1.2 Terms

A term is a monomial with a non-zero coefficient, so for example:

$$17 \cdot x_1^3 \cdot x_2^5 \cdot x_3^{13}$$

Is a term with degree 21 $(3 + 5 + 13)$ and a coefficient of 17.

3.1.3 Polynomials

A polynomial is a finite sum of terms, the degree of which is defined to be the maximum degree of any of the terms.

Exception The polynomial:

$$f = 0$$

Has an undefined degree. Terms only have a **non-zero** coefficient, hence 0 doesn't have any terms and so the definition of degree doesn't work for it.

Whereas $f = c, \quad \exists c \in \mathbb{C}$ does have 1 term, for which the degree is 0.

Support of a polynomial The support of a polynomial f is the set of monomials appearing in f, e.g. for the following 6th degree polynomial in 2 variables, the support of that polynomial is given by:

$$f(x) = x^2 + 3x^3 + 4y \implies \operatorname{supp}(f) = \left\{x^2, 3x^3, 4y\right\}$$

The initial of the support $\operatorname{in}_{\prec}(f)$ is the polynomial with the highest ranking with respect to some ordering of the monomials (see §) [15, p. 1.1.5].

Other Terminology The following terms are commonly used: [9, §2.2]

- The $\operatorname{multidegree}(f)$, is the largest power of any variable of any monomial in a polynomial
- The Leading coefficient $\operatorname{LC}(f)$ is the term corresponding to the monomial containing the variable that corresponds to the multidegree
- The Leading monomial $\operatorname{LM}(f)$ is the monomial corresponding containing the variable that corresponds to the multidegree
- The Leading term $\operatorname{LT}(f)$ is the product of the leading coefficient and the leading monomial

So for example, in the polynomial:

$$f = 4x^2y^2 + 3x^3 + 7xy$$

- The initial is $4x^2y^2$
- The Leading Coefficient is 3
- The Leading Monomial is x^3
- The Leading Term [2] is x^3

[2] This also lines up with `sympy`'s `LT()` function, beware not to confuse the initial with the leading term,

Homogenous Polynomial If all terms of a polynomial have an equal degree (say $\exists q \in \mathbb{N}$ Then that polynomial is said to be a *homogenous polynomiial of degree* q, e.g.

$$x_1^3 \cdot x_2^4 \cdot x_3^2 + x_1^6 \cdot x_5^2 \cdot x_7$$

is a homogenous polynomial of degree 7.

The Polynomial Ring The Rings, Vectors and Polynomials

Let $K[x_1, x_2, x_3, \ldots x_n] = K[\mathbf{X}_n]$ denote the set of all polynomials in the variables $x_1, x_2, x_3, \ldots x_n$ with coefficients in some field K.

If f and g are polynomials from $K[x_1, x_2, x_3, \ldots x_n]$ with addition and multiplication defined in the ordinary way (i.e. just normal algebra), then $K[x_1, x_2, x_3, \ldots x_n]$ forms an algebraic structure known as a Ring.

Readers may be familiar with the axioms of a vector space, for which the set of polynomials in $K[x_1]$ of degree less than n also satisfies [19, §4.4], a ring structure is much the same concept, it's a set with specific characteristics. One of the main differences is that while a vector space requires a scalar multiplicative identity, a ring structure does not.

On the other hand not all vector spaces are necessarily rings because they are not necessarily closed under multiplication (although defining multiplication by element-wise product would remedy this), see §9.1.1 for more information.

4 Ideals and Varieties

4.1 Affine Space

The affine n-space of some field K is given by: [9, §1.1]

$$K^n = \left\{(a_1, a_2, a_3, \ldots, n) \mid a_i \in K, \forall i \in \mathbb{Z}^+\right\}$$

For example if K was given by $\mathbb{R}$ the resulting affine n-space would be the *Cartesian Plane.*

4.2 Zero Point

The zero-point of some function $f \in K[\mathbf{X}_n]$ is a point in K^n: [15]

$$f(a_1, a_2, a_3 \ldots a_n) = 0.$$

In the broader context of equations rather than specifically functions, these points are often referred to as roots.

These points are often referred to as roots [17, §17.2], however this is usually in the context of equations more broadly rather than functions specifically. [1]

different algorithms or ways to calculate an S-polynomial seem to use either and it doesn't matter, I'm not sure why yet, but I am sure that there is a difference between the initial monomial and the leading term.

4.3 Variety

Consider a set of functions $F = \{f_1, f_2, f_3, \dots f_s\}$, the variety of this set of functions is denoted $\mathbf{V}(F)$ and is the set of all zero-points of all the functions:

$$\boldsymbol{V}(F) = \left\{(a_1, a_2, a_3, \dots a_n) \in K^n \mid f_i(a_1, a_2, a_3, \dots a_n), \forall i \in \mathbb{Z}^+ < s\right\}$$

The convention is that all functions in F are set to be equal to 0, and if this convention is taken, the variety of that set is the set of solutions corresponding to that set of equations.

4.3.1 Example

Consider for example the set $\{-y + x^2 - 1, -y + 1\}$, the solution to this system can be found by substitution:

$$\begin{aligned} -y + x^2 - 1 = 0 &= -y + 1 \\ x^2 - 1 &= y = 1 \\ x^2 &= 2 \\ x &= \pm\sqrt{2} \end{aligned}$$

and so:

$$\boldsymbol{V}\left(\left\{-y + x^2 - 1, -y + 1\right\}\right) = \left\{\left(-\sqrt{2}, 1\right), \left(\sqrt{2}, 1\right)\right\}$$

4.4 Ideals

Ideals are a set with a particularly convenient property, given functions $f, g \in K[\mathbf{X}_n]$, a subring $I \subset K[\mathbf{X}]$ is said to be an ideal if it is closed under addition and admits other functions under multiplication: [15, §1.1.3]

1. $f \in I \wedge g \in I \implies f + g \in I$
2. $f \in I \wedge g \in k[\mathbf{X}] \implies gf \in I$

So for example, $\{0\}$ is an ideal of the polynomial ring in all variables, and as a matter of fact $0 \in I$ for all ideals of polynomial rings in all variables.

A subring is a subset that is itself a ring, so I would be a subset that is closed under addition and multiplication and contains an additive identity (i.e. $0 \in I$). [3] As a matter of fact it can be shown that:

- $0 \in I$
- $\{0\}$ is an ideal

for all ideals in all variables and that is an ideal (because otherwise the result would not be admitted to I).

[3] It would also be sufficient to show that the I is closed under both addition and subtraction [17, §16.1]

4.4.1 Example

Let $R = \mathbb{Z}$ and $I = 2\mathbb{Z}$, the set of $\mathbb{Z}$ is a commutative ring with unity, $2\mathbb{Z} \subset \mathbb{Z}$ is:

1. $2\mathbb{Z} \neq \emptyset$
2. closed under multiplication and addition
3. admits any other integer under multiplication (i.e. even $\times$ anything is even)

4.5 Ideals and Varieties

If we have a variety of $V \subset K^n$, we denote, $I(V)$ as the set of all polynomials $f_i \in k[\mathbf{X}]$: [15, §1.1.3]

$$f_i(a_1, a_2, a_3, \ldots a_n) = 0, \quad \forall (a_1, a_2, a_3, \ldots a_n) \in V.$$

this set of functions satisfies the properties of an ideal and is known as the ideal of V [9].

In other words, the ideal of the variety of a set of functions, $I(\mathbf{V}(F))$, is the set of, polynomials, that have the same zero-points as the simultaneous zero points of all functions in F.

4.6 Generating Ideals

The ideal generated by F is:

$$\langle F \rangle = \left\{ p_1 f_1 + p_2 f_2 + p_3 f_3 + \ldots p_n f_n \mid f_i \in F, p_i \in K[\boldsymbol{X}], \forall i \in \mathbb{Z}^+ \right\}$$

Such a set satisfies the properties of an ideal and is a subset of the functions that share the zero-points with F: [9, p. 34]

$$\langle F \rangle \subseteq I(\boldsymbol{V}(F))$$

$\langle F \rangle$ is the set of all the linear combinations of elements in F with polynomials in $K[\mathbf{X}_n]$, another way to phrase it would be that $\langle F \rangle$ is the set of polynomial consequences of F [9, p. 30].

If some **finite** set of polynomials F, can generate an ideal I, it is said that I is finitely generated and that F is a basis for I. Every ideal in $K[\mathbf{X}_n]$ is finitely generated [9, p. 77], this is known as *Hilbert's Basis Theorem*, this is important because it means we if we had an algorithm that involved taking different polynomials from such a basis, that algorithm would eventually end.

If two sets are bases of the same ideal, they will have the same variety, i.e. if two sets can generate the same set of functions, they'll have the same solutions (assuming that the set of functions is an ideal), this also implies

4.6.1 Initial Ideal

The initial ideal:

$$\langle \mathrm{in}_{\prec}(I) \rangle = \langle \{ \mathrm{in}_{\prec}(f) : 0 \neq f \in I \} \rangle$$

is generated by infinitely many monomials, namely the initial monomials, for the infinitely many polynomials in the ideal I. [15, §1.1.5]

It's common also to see a similar definition for the ideal generated by the leading terms is denoted $\langle \mathrm{LT}(f) \rangle$ [9, §2.5].

4.6.2 Comparison with Linear Algebra

If S is some set of vectors and every vector in a vector subspace V can be written as a linear combination of the elements of S is is said that S spans V, so for example $S = \{\langle 1, 0\rangle, \langle 0, 1\rangle\}$ spans $\mathbb{R}^2$ or $S = \{1, x, x^2\}$ spans P_2.

Ideals for rings are similar in nature to vector subspaces and normal subgroups. It's worth drawing attention to the fact that that the term basis in the context of an ideal (which could be more accurately called a generating set [26]) is quite different from a linear basis [9, p. 35].

In linear algebra a basis spans and is linearly independent, the basis of an ideal however only spans, there is no independence, for example:

$$\begin{aligned} f_1(x, y) &= y \qquad \vec{v}_1 = \langle 0, 1\rangle \\ f_2(x, y) &= x \qquad \vec{v}_2 = \langle 1, 0\rangle \end{aligned}$$

Linear independence is generally satisfied if linear combination is equal to zero, only if the multiplying terms are zero, i.e. f_1 and f_2 are linearly independent only if:

$$\begin{aligned} 0 &= a\langle 0, 1\rangle + b\langle 1, 0\rangle, \quad \forall a, b \in \mathbb{R} \\ &= \langle a, b\rangle \\ &\implies a = b = 0 \end{aligned}$$

This clearly doesn't work for polynomials, however, because setting $g_i = x$ and $g_j = -y$ satisfies such an equation.

$$0 = g_i y + g_j x, \quad \not\implies g_i = g_j = 0, \quad \forall g_i g_j \in k[\mathbf{X}]$$

So linear independence doesn't have a lot of meaning with polynomials, it's only the spanning property that is meaningful.

5 Initials and Leading Monomials

5.1 Monomial Ordering

Monomials are ordered by degree, e.g. $x \prec x^2$ or $xyz \prec x^2yz$, however in many variables it isn't always clear which order should be chosen, for example the following monomials have the same degree and if they are ordered by the value on first variable:

$$xy^3 \prec x^2yz$$

If however they are ordered by trying to minimize the last variable:

$$x^2yz \prec xy^3$$

Recall from polynomial long division that the first term in a polynomial important to the algorithm, for a similar reason it is necessary to decide before hand on an ordering, and generally in this report the lexicographic order (i.e. alphabetical) will be used.

This isn't as important as many texts make it out to be and so further discussion appears further below in §9.2.

6 Groebner Bases

A finite subset G of an ideal I is a Grobner Basis, (with respect to some term order $\prec$, if: [7, 15]

$$\{\text{in}_\prec(g) \mid g \in G\}$$

generates $\{\text{in}_\prec(I)\}$

It's common also to see this definition reformulated with respect to leading terms as opposed to initial monomials, in which case G is said to be a Groebner Bases if: [9, p. 2.5]

$$\text{LT}(I) = \langle \text{LT}(g_1), \text{LT}(g_2), \text{LT}(g_3), \ldots \text{LT}(g_n) \rangle$$

there are many such generating sets, we can add any element to G to get another Groebner Basis, so in practice we may be more concerned with reduced Groebner Basis. Note also that even though the leading term is different from the initial monomial, either can be used to define a Groebner Bases, however it is not yet clear to me if the Groebner Bases will depend on the monomial ordering $\prec$ only if the initial is used to define it.

The variety of a set of functions depends only on the ideal of F, if two sets generate the same ideal they have the same variety and if G is a Grobner Basis for F, then $V(G) = V(F)$.

The reason we care about a Groebner Bases more generally is because the set tends to provide more information of the solution space.

7 Buchberger's Criterion

G is a Groebner basis, if and only if, every S-polynomial formed by any two pairs from G has a remainder of 0, where the S-polynomial is given by: [9, §2.6]

$$S(f,g) = \text{lcm}(\text{LM}(f), \text{LM}(g)) \times \left(\frac{f}{\text{LT}(f)} - \frac{g}{\text{LT}(g)}\right)$$

The remainder that we are concerned with is:

$$r = \overline{S(f,g)}^{G} = S(f,g) \mod \prod_{g \in G} (G)$$

8 Bucherger's Algorithm

Buchberger's Algorithm takes a set of polynomials, F and eventually returns another set G which is a Groebner Bases.

To do this the algorithm tests every pair of polynomials in F with the criterion above, if the remainder for any pair is non zero, it is placed into F as another polynomial. Once every combination has been considered, the original set F will be a Groebner Basis.

8.1 Reduced Groebner Basis

A reduced Groebner Basis is a Groebner Basis that has needless polynomials discarded, I have not had time to investigate these yet.

8.2 Examples

for examples of Buchberger's Algorithm, refer to the attached *Jupyter Notebook*, this is quite sparse as resources to understand the algorithm were discovered quite late in the investigation as was the realisation that use `sympy` had a significant amount of documentation on the algorithm.

9 Abstract Algebra

The following are concepts that are *nice to have* in understanding the topic, but are not strictly necessary to get a broad understanding of the topic.

They were needlessly investigated early on because accessible resources (e.g. [9, 4, 29]) had not yet been discovered.

9.1 Background

9.1.1 Algebra

Relations A relation on a set A is a subset R of the Cartesian product:

$$A \times A = \{(a,b) : \quad a,b \in A\}$$

If $(a,b) \in R$ it is written that $a \ R \ b$.

Example The example most relevant to the theory of Groebner bases [4] is the $<$ relation. If we had the set $A = \{1,\ 2,\ 3\}$

[4] Relevant because we need to decide on an ordering relation in order to use Buchberger's algorithm, which is needed to find a Groebner Basis.

The cartesian product would be:

$$A \times A = \left\{ \begin{array}{l} (1,1), (1,\ 2), (1,3), \\ (2,1), (2,2), (2,3), \\ (3,\ 1), (3,2), (3,3) \end{array} \right\}$$

The set corresponding to the relation < would be:
$\{(1,2),\ (1,3),\ (2,3)\}$
and so it is said that:

- $1 < 2$
- $1 < 3$
- $2 < 3$

Types of Relations

- **Reflexive** relations are relations where
  - $\forall\, a \in A, a\ \ R\ \ a$
- **Symmetric** relations are such that
  - $\forall\, a, b \in A, a\ R\ b \Rightarrow b\ R\ a$
- **Transitive** relations are such that
  - $a\ R\ b \wedge\ \ b\ R\ c \Rightarrow\ \ a\ R\ c$
    - $\forall\, a, b, c \in A$

If all of these are satisfied, the relation is said to be *an equivalence relation.*

Why? Although this might seem needlessly pedantic, the algorithm we hope to use to find solutions to systems of polynomial equations, Buchberger's Algorithm, require us to decide on a way to order polynomials, for example in a quadratic equation it's pretty straight forward:

$$f(x) = ax^2 + bx + c$$

But for multiple variables it gets confusing, for example we could order the terms by degree, but if multiple terms are of the same degree then we could make sure that the left most variable has an exponent that is descending, or, we could try and make sure that the right most term is ascending:

$$\begin{aligned} f\,(w,x,y,z) &= wz + xy && (1) \\ &= xy + wz && (2) \end{aligned}$$

This is already pretty confusing so having a firm definition of ordering is important.

Congruence

Equivalence Classes The set of all elements of A that satisfy the relation for a is said to be the /equivalence class of a with respect to R:

$$[a]_R = [a] = \{b \in A : b\ R\ a\}$$

So returning to the example from §9.1.1, we would have:

- $[1]_< = \emptyset$
- $[2]_< = \{1\}$
- $[3]_< = \{1, 2\}$

Congruence Modulo n It is said that a and b are *congruent modulo* n if $n \mid (a - b)$ and it is written:

$$a \equiv b \pmod{n}$$

It is common to see $\mod$ used as an operator:

$$a \mod b = r$$

The congruence class of a modulo n is expressed $[a]_n$ and is the equivalence class of a whereby the relation is congruence in modulo n:

$$[a]_n = \{b \in \mathbb{Z} : b \equiv a \pmod{\text{n}}\}$$

1. Example Clock time is a congruence class, for example 11 O'clock + 3 hours = 2 PM:

   $$[11]_{12} + [3]_{12} = [2]_{12}$$

   Another example could be binary:

   $$[1]_2 + [3]_2 = [0]_2$$

   See also [25, §4c]

2. Congruence generalised with Groups If G is a group and H a subgroup, if we have:

   $$a^{-1}b \in H$$

   then it is said:

   $$a \equiv b \pmod{H}$$

   the use of "≡" is appropriate because the relationship is:

- reflexive
- symmetric
- transitive

and is hence an equiv class.

consider for example:

$$12\mathbb{Z} \leqslant \mathbb{Z}$$

so we have 5-17 $\in$ 12 Z

So we write:

$$5 \equiv 17 \pmod{12\mathbb{Z}}$$

See [12, §3.7].

3. Congruence Modulo an Ideal Congruence can be extended to an ideal on any ring structure, that's why we needed to generalise this structure, in order to use these theoryies.

   congruence modulo an ideal is

   If I is an ideal in a ring R

$$a \equiv b \pmod{I} \iff a - b \in I$$

   The use of justified because this is an equivalence relation

   The equivalence class is the set of all elements that satisfy that relation for a:

$$\forall a \in A,$$
$$[a]_R = [a] = \{b \in A : bra\}.$$

   So in the context of congruence:

$$a \in G$$
$$[a] = \{b \in G : b \equiv a \pmod{H}\}.$$

   if we wanted to find b :

$$\begin{aligned} b &\equiv a \pmod{H} \\ a^{-1}b &\in H \\ a^{-1}b &= h, \quad \exists h \in H \\ b &= ah. \end{aligned}$$

So we have:

$$[a] = \{ah : h \in H\}.$$

This is known as the left coset [17, §6.1]. The left cosets of H in G partition G: [12, §3.3]

(a) Each $a \in G$ is in onlyone left coset, which is aH
(b) $aH \cap bH = \emptyset$ or $aH = bH$

This can be used to show:

$$H \leq G \implies |H| \mid |G|.$$

this is known as Lagranges Theorem. [12, §3.7]

(a) Normal Subgroups
A normal subgroup is a subgroup $N \leq G$:

$$aN = Na \quad \forall a \in G.$$

This is not so strict as to require all elements be commutative (although commutative groups are of course normal)

4. Congruence Classes for Polynomials If f and g are in an ideal I, then [9, p. 240]:

$$f - g \in I \implies f \equiv g \pmod{I}$$

Groups A set G is a group, if there in a binary operation, $\star$, defined on that set such that:

1. The binary operation is closed on the set

$$a, b \in G \implies a \star b \in G$$

2. The binary operation is associative

$$a, b, c \in G \implies a \star (b \star c) = (a \star b) \star c$$

3. There is an element that doesn't do anything under the binary operation, this is known as an identity element, for example 1 is an identity element to the multiplication operation.

$$\exists e \in G :$$
$$a \star e = e \star a = a$$

4. Every element has an inverse

$$\forall a \in G,\ \ \exists a^{-1} \in G : \\ a \star a^{-1} = e$$

   - For operations that are additive in nature, it is common to use the notation: $-a$ [12, §3.3]

5. If the binary operation is also commutative, the group is said to be abelian:

$$\forall a, b \in G, \\ a \star b = b \star a \iff G \text{ is abelian.}$$

Example An example of a group is a set of all matrices of a given size under addition, this can be seen because:

1. Adding matrices gives back matrices of the same size,
2. Introducing brackets in addition doesn't change the result
3. A matrix with all 0's is an identity
4. Any matrix $\mathbf{A}$ has an inverse (namely $-\mathbf{A}$)

This example would also be an abelian group because addition is commutative.

Note that if the operation was matrix multiplication, $\cdot$ (denoted as `%*%` in ***R*** [31]), only square matrices with a non-zero determinant (e.g. $|\mathbf{A}| \neq 0$) could be a group. This is because the matrix would need to be invertible. [5]

But Why? The reason groups are interesting is because many natural structures can be described by a set and a binary operation, obvious examples are sets of numbers, vectors, matrices and equations, but more generally Group theory can be used to describe puzzles like *Rubik's Cube* [16], chemical structures [14] and has been used in the theory of quantum mechanics [32]. [6]

Rings Examples, equivalence class ring [17, Ch. 3] see also §2.4 of nicodemii [21, §2.4]

Rings are an abelian group under addition $+$, with a second binary operation that corresponds to multiplication $\times$, this operatuion must be closed, associative and distributive, but there is no need for an inverse or identity [12, §8.1]. So a ring structure is a set $\mathcal{R}$, with two closed binary operations, that satisfies the following axioms of a ring [21, §§2.4-2.6]:

[5] although the element-wise product, $\odot$, would not present this issue.

[6] See generally this [3] *Stack Exchange Discussion.*

1. Associativity of Addition

   $(\forall a, b, c \in \mathcal{R})\ (a + b) + c = a + (b + c)$

1. Commutativity of Addition

   $(\forall a, b \in \mathcal{R})\ a + b = b + a$

2. Additive Elements Exist

   $(\forall a \in \mathcal{R}) \wedge (\exists 0 \in \mathcal{R})\ a + 0 = 0 + a = a$

3. Additive Inverse Exists

   $(\forall a \in \mathcal{R}) \wedge (\exists b \in \mathcal{R})\ a + b = b + a = 0$

   - This can be equivalently expressed:

   $(\forall a \in \mathcal{R}) \wedge (\exists\, (-a) \in \mathcal{R})\ a + (-a) = (-a) + a = 0$

4. Associativity of Multiplication

   $(\forall a, b, c, \in \mathcal{R})\ (a \cdot b) \cdot c = a \cdot (b \cdot c)$

5. Distributivity of Multiplication over Addition

   - $(\forall a, b, c, \in \mathcal{R})\ (a \cdot (b + c) = (a \cdot b) + (a \cdot c))$, AND
   - $(\forall a, b, c, \in \mathcal{R})\ (a + b) \cdot c = (a \cdot c) + (b \cdot c)$

Further Axioms Other conditions that correspond to different classes of rings are:

1. Commutativity of Multiplication

   - A ring that satisfies this property is called a **commutative ring**
     $(\forall a, b \in \mathcal{R})\ a \cdot b = b \cdot a$

2. Existence of a Multiplicative Identity Element (A ring with Unity)

   - A ring that satisfies this property is called a **ring with identity** or

   equivalently a **ring with unity** (the multiplicative identity, often denoted by 1, is called the **unity** of the ring.

   $(\exists 1 \in \mathcal{R})\ (\forall a \in \mathcal{R})\ 1 \cdot a = a \cdot 1 = a$

Example An obvious example of a ring is the set of all integers $\mathbb{Z}$ with the ordinary meaning of addition and multiplication. A more insightful example would be a congruence class, for example $\mathbb{Z}_{12}$, this satisfies the axioms of a ring, but some values are zero divisors. If two elements of a ring multiply to give 0, those values are said to be zero divisors, for example 3 and 4 are zero divisors in Z_{12}:

$$[3]_{12} \times [4]_{12} = [0]_{12}$$

An element that has an inverse is said to be a unit, for example:

$$[2]_9 \times [5]_9 = [1]_9$$

An element can't both be a unit and a zero divisor, because one multiplies to give 0 and the other to give 1, however, in many algebraic structures (e.g. $\mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$) every element has a multiplicative inverse, and this motivates the next algebraic structure.

Integral Domains An integral domain is a commutative ring with identity that has no zero divisors.

Example The obvious example of an integral domain is $\mathbb{Z}$, but any $\mathbb{Z}_p$ where p is a prime number, will also be an integral domain.

Another example is the set of all polynomials with real coefficients, this will be explored in greater detail below, but for the moment observe that this algebraic structure conforms to the axioms of a ring and has no zero divisors.

It can be clearly seen that the set of polynomials has no zero divisors because:

$$f \times g = 0 \tag{3}$$
$$\implies f = 0, \vee g = 0 \tag{4}$$
$$\tag{5}$$

in either case f or g is not a non-zero divisor.

Note however that not every element of the polynomials has an inverse, for example the function $f(x) = x$ would have an inverse $f^{-1}(x) = \frac{1}{x}$, but this is not a polynomial.

This leads to the final algebraic structure that will be considered here. [7]

Fields A field is a commutative ring with identity in which all non-zero elements are units.

Because every element of a field is a unit, it implies that every element is not a zero-divisor, and so hence a field is:

[7] There are other algebraic structures that could be interesting, for example polynomials can also be considered as vectors, see e.g. [19], as a matter of fact all vector spaces are rings if multiplication is defined element-wise by the *Hadamard product* ($\odot$), this could be an interesting relationship to investigate further.

- a special case of an integral domain, which is in turn
- a special case of a ring, which is in turn
- a special case of a group.

Rings and Integral Domains It seems that the reason the theory of Groebner Bases is concerned with the ring of polynomials over a field is related to the irreducibility of the polynomial, see generally [11].

Note also that the Ring of polynomials over an integral domain (a property satisfied also by a field) is more accurately an integral domain [27, 24], not merely a ring.

Why aren't Polynomials fields A field is an integral domain, for which every element has an inverse, so consider some function, say $g(x) = x$, if the set of polynomials was a field, there would have to exist some $f(x)$ such that:

$$x \cdot f(x) = 1$$

however if we evaluate this at $x = 0$

$$0 \cdot f(0) = 1$$

well... this clearly doesn't work, so it's clear that this $f(x)$ doesn't exist and so the set of polynomials is not a field, see generally [8]

One might wonder if there's a good reason why $f(x) = \frac{1}{x}$ isn't considered a polynomial, notwithstanding the fact that it doesn't quite fix this example with 0:

- All polynomials over the real numbers are continuous, that would make this membership inconvenient.
  - On the other hand there are discontinuities of arbitrary polynomials over certain fields, what's a good example of a such a field though?

The easy and uninformative answer is that $\frac{1}{x}$ does not have positive indices, outright violating the definition.

9.1.2 Vector Spaces

The ring of polynomials over a field K:

$$K\left[x_1, x_2, x_3, \ldots, x_n\right]$$

is a n-vector space with a basis given by the set of all power products:

$$\left\{x_1^{\beta_1}, x_2^{\beta_2}, x_3^{\beta_3}, \ldots x_n^{\beta_n}\right\}$$

Basis A basis is a set of vectors that [5, p. 39] are:

- Linearly independent
- Spans an n-dimensional vector space??

Linear Independence a set of vectors are linearly independent if:

$$a_1 v_1 + a_2 v_2 + a_3 v_3 \ldots = 0 \iff a_1 = a_2 = a_3 = \ldots = a_m$$

Span The span of a set of vectors, is the set of all possible linear combinations of those vectors.

So for example:

$$\begin{aligned} \mathbb{R}^2 &= \text{span}\left(\{(0,1),(1,0)\}\right) && (6)\\ &= \text{span}\left(\{(0,2),(2,0)\}\right) && (7)\\ &= \text{span}\left(\{(1,1),(1,-1)\}\right) && (8)\\ & && (9) \end{aligned}$$

To visualize this in $\mathbb{R}^2$, imagine that by varying the scaling value of each vector, any point on $\mathbb{R}^2$ can be reached.

Vectors A ring with unity is a vector space, however a vector space only needs to be closed under scalar multiplication. This means vector spaces are not necessarily rings unless the multiplication operation is closed, an example of a closed vector multiplication is element-wise multiplication, this is known as the hadamard product (think like mutliplying 'numpy' arrays.)

9.2 Monomial Orders

groebner bases of a system of equations A partial order on a set is a relation R:

- xRx
  - reflexivity
- $xRy \wedge yRx \implies x = y$
  - Antisymmetry
- $xRy \wedge yRz \implies xRz$
  - Transitivity

So for example, the set of integers has $\leq$ as a relation such that $n_1 \in \mathbb{Z}$:

- $n \leq n$
- $n_1 \leq n_2 \wedge n_2 \leq n_1 \implies n_1 = n_2$
- $n_1 \leq n_2 \wedge n_2 \leq n_3 \implies n_1 \leq n_3$

A partially ordered set is one with a relation that is a partial order.

- partial order
  - a relation
- partially ordered set
  - a set

A total order is a partial order such that $\forall x, y$ either xRy or yRx, the obvious example is $<$, consider for example $\mathbb{C}$, this has a partial order if the the modulus is considered, it's only a partial order because, e.g. $|i+i| = |i-i|$. not all sets will have a partial ordering, e.g. the somewhat contrived example has no (at least obvious) partial order.

$$\left\{\square, \triangle, \sqrt{-1}x^{e^x}\right\}.$$

$k\left[\mathbf{X}\right]$ is a polynomial ring in n variables and $\mathcal{M}_n$ is the set of amonomials in the variables $x_1, x_2, x_3, \ldots x_n$.

A monomial order on $k\left[\mathbf{X}\right]$ is a total order on $\prec$ on $\mathcal{M}_n$:

1. $i \prec u, \quad \forall 1 \in u \in \mathcal{M}_n$
2. $u, v \in \mathcal{M}_n \wedge u \prec v \implies uw \prec vw, \forall w \in \mathcal{M}_n$

Lexical monomial order Let:

$$u = x_1^{a_1} x_2^{a_2} x_3^{a_3} \ldots x_n^{a_n}$$
$$v = x_1^{b_1} x_2^{b_2} x_3^{b_3} \ldots x_n^{b_n}.$$

The lexicographic order on $k\left(\mathbf{X}\right)$ is given by the total order $<_{\text{lex}}$ on $\mathcal{M}_n$ by setting:

$$u <_{\text{lex}} v.$$

if:

1. $\sum_{i=1}^{n}\left[a_i\right] \leq \sum_{i=1}^{n}\left[b_i\right]$
2. the leftmost non-zero term in the following vector is positive:
   - b_1-a_1, b_2-a_2, b_3-a_3 ...b_n-a_n

Reverse lexicographic is:

1. $\sum_{i=1}^{n}\left[a_i\right] \leq \sum_{i=1}^{n}\left[b_i\right]$
2. the **rightmost** non-zero term in the following vector is **negative**:
   - b_1-a_1, b_2-a_2, b_3-a_3 ...b_n-a_n

These should be combined into one statement ↑
So for example consider:

$$x_1x_4 - x_2x_3.$$

by lexicographic we have

$$x_2x_3 \prec x_1x_4.$$

because the leftmost entry is positive in the vector described before:

$$\langle 1, -1, -1, 1 \rangle .$$

by reverse lexicographic we have

$$x_1x_4 \prec x_2x_3.$$

because the **rightmost** entry is **negative** in the vector described before:

$$\langle -1, 1, 1, -1 \rangle .$$

This may be discussed more in the org mode note.
an interesting property that comes back in the buchberger algorithm and polynomial long division is:

$$\mathrm{in}_\prec (f \cdot g) = \mathrm{in}_\prec (f) \, \mathrm{in}_\prec (g) .$$

Colloquial

Lexicographic The highest variable is so expensive that it makes the entire monomial expensive.

Reverse Lexicographic The lowest variable is so chap that it makes the entire monomial cheap.

9.3 Dickson's Lemma

9.3.1 Divisors

For *monomials*:

- $u = \prod_{i=1}^{n} [x_i^a i]$ $\quad a \in \mathbb{Z}^+$
- $v = \prod_{i=1}^{n} \left[x_i^b i\right]$ $\quad b \in \mathbb{Z}^+$

u is said to divide v if $a_i \leq b_i \quad \forall i \in [1, n]$

Example Consider:

- $u = x^2y^3z^5$
- $v = x^1y^2z^3$

In this case $v \mid u$ because:

$$1 < 2$$
$$2 < 3$$
$$3 < 5$$

$$\frac{u}{v} = \frac{x^2}{x^1} \cdot \frac{y^3}{y^2} \cdot \frac{z^5}{z^3}.$$

9.3.2 Minimal Element

let $\mathcal{M}_n$ be the set of a all monomials in the variables $x_1, x_2, x_3, \ldots x_n$ and $M \subset \mathcal{M}_n$ be a nonempty subset thereof.

The following condition describes a minimal element $u \in M$:

$$(v \in M \wedge v \mid u) \implies v = u$$

In other words, u is a minimal element if the only way that $v \mid u$ is if $v = u$.

Example Consider $\mathcal{M}_2$:

$$\begin{aligned} \mathcal{M}_2 = \{ & xy, xy^2, & xy^3, \ldots & \quad (10) \\ & & x^2y, x^2y^2, x^2y^3, \ldots & \quad (11) \\ & & x^3y, x^3y^2, x^3y^3, \ldots & \quad (12) \\ & & \vdots & \quad (13) \\ \} & & & \quad (14) \end{aligned}$$

and let's have the subset $M = \{x^2y, x^2y^2, x^2y^3 \ldots\}$, the minimum elements are:

$$\left\{x^2y\right\}$$

clearly $|M| = \infty$, however this number of mi $\implies$ $umelementswillalwaysbefinite, thisisknownas$**Dic**

9.3.3 Dickson's Lemma

/Dickson's Lemma is the main result needed to prove the termination of Buchberger's algorithm for computing Groebner basis of polynomial ideals/ [20].

Let

- $\mathcal{M}_n$ be the set of all monomials in variables $x_1, x_2, x_3 \dots x_n$.
- M be a nonempty subset of $\mathcal{M}_n$

The set of minimal elements of a nonempty subset $M \subset \mathcal{M}_n$ is at most finite.

This intuitively makes sense, I can't have an infinite number of minimums, otherwise they wouldn't be minimums, the proof is very difficult though.

In one Variable By definition, a monomial is raised to the power of a non-zero integer, in a single variable monomial the smallest index will correspond to the minimal element (by the definition of the minimal element) and hence the existence of a minimum element in $\mathrm{Z}^{+} implies the existence of a minimum element in M \subset \mathcal{M}_n$.

In Two Variables Assume that there is an infinite number of minimal elements:

$$u_1 = x^{a_1} y^{b_1} \tag{15}$$
$$u_2 = x^{a_2} y^{b_2} \tag{16}$$
$$u_3 = x^{a_3} y^{b_3} \tag{17}$$
$$u_4 = x^{a_4} y^{b_4} \tag{18}$$
$$u_5 = x^{a_5} y^{b_5} \tag{19}$$
$$\dots$$

Let's order the values by the first exponential such that $a_1 \leq a_2 \leq a_3 \dots$.

If $a_i = a_{i+1}$, then either:

- $u_1 = u_{i+1}$
  - We can't have this because set's do not contain repeated elements.
- $y^{b_i} \neq y^{b_{i+1}}$
  - But this would mean that either u_i or $u_{i+1\}}$ is not a minimal element, so this can't occur either.

This means that each a_i must be different and so:

$$a_i < a_2 < a_3 \dots \tag{20}$$

If $u_i | u_{i+1}$ one of them is not a minimal element and so we must have $b_i > b_{i+1}$, hence $b_i > b_2 > b_3 \dots$.

This means that b_i represents an upper bound for the number of different minimal elements, hence the number of minimal elements must be finite.

In n variables **INDUCTION** If the number of minimal elements is finite for $M_n \subset \mathcal{M}_n$ we would expect M_{n+1} to be finite as well, adding an extra variable should not make the number of minimal elements infinite because the integers in the index will still behave as an upper bound.

I need to formalise this as per [15, §1.1.2].

References

[1] Ilya (https://math.stackexchange.com/users/5887/ilya). *Root or Zero...Which to Use When?* eprint: https://math.stackexchange.com/q/82645. URL: https://math.stackexchange.com/q/82645 (cit. on p. 5).

[2] William W. Adams and Philippe Loustaunau. *An Introduction to Gröbner Bases*. Graduate Studies in Mathematics v. 3. Providence, R.I: American Mathematical Society, 1994. 289 pp. ISBN: 978-0-8218-3804-4 (cit. on p. 2).

[3] Alex B. *Why Should We Care about Groups at All?* eprint: https://math.stackexchange.com/q/19328. URL: https://math.stackexchange.com/q/19328 (cit. on p. 15).

[4] Andreas Schulz. *Integer Programming and Combinatorial Optimization*. MIT OpenCourseWare. URL: https://ocw.mit.edu/courses/sloan-school-of-management/15-083j-integer-programming-and-combinatorial-optimization-fall-2009/ (visited on 04/08/2021) (cit. on pp. 1, 10).

[5] Sheldon Axler. *Linear Algebra Done Right*. New York: Springer, 2014. ISBN: 978-3-319-11079-0 (cit. on p. 18).

[6] Thomas Becker, Volker Weispfenning, and Heinz Kredel. *Gröbner Bases: A Computational Approach to Commutative Algebra*. Graduate Texts in Mathematics 141. New York: Springer-Verlag, 1993. 574 pp. ISBN: 978-0-387-97971-7 978-3-540-97971-5 (cit. on p. 2).

[7] Bernd Sturmfels, director. *Introduction to Grobner Bases - Prof. Bernd Sturmfels*. Jan. 21, 2017. URL: https://www.youtube.com/watch?v=TNO5WuxuNak (visited on 04/12/2021) (cit. on p. 9).

[8] Bill Dubuque. *Abstract Algebra - Why Can't the Polynomial Ring Be a Field?* Mathematics Stack Exchange. URL: https://math.stackexchange.com/a/2523 (visited on 04/12/2021) (cit. on p. 18).

[9] David A. Cox, John B. Little, and Donal O'Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra: With 91 Illustrations*. 2nd ed. Undergraduate Texts in Mathematics. New York: Springer, 1997. 536 pp. ISBN: 978-0-387-94680-1 (cit. on pp. 1–5, 7–10, 14).

[10] E. H. Connell. *Elements of Abstract and Linear Algebra*. University of Miami, Mar. 30, 2001. URL: http://www.math.miami.edu/~ec/book/ (cit. on p. 3).

[11] *Equivalence of Definitions of Irreducible Polynomial over Field - ProofWiki*. URL: https://proofwiki.org/wiki/Equivalence_of_Definitions_of_Irreducible_Polynomial_over_Field (visited on 04/12/2021) (cit. on p. 18).

[12] Gregory Lee. *Abstract Algebra*. New York, NY: Springer Berlin Heidelberg, 2018. ISBN: 978-3-319-77648-4 (cit. on pp. 2, 13–15).

[13] Pierre A. Grillet. *Abstract Algebra*. 2. ed. Graduate Texts in Mathematics 242. New York, NY: Springer, 2007. 669 pp. ISBN: 978-0-387-71568-1 978-0-387-71567-4 (cit. on p. 2).

[14] *Group Theory and Its Application to Chemistry*. Chemistry LibreTexts. Oct. 2, 2013. URL: https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Supplemental_Modules_(Physical_and_Theoretical_Chemistry)/Group_Theory/Group_Theory_and_its_Application_to_Chemistry (visited on 04/10/2021) (cit. on p. 15).

[15] Takayuki Hibi. *Grobner Bases: Statistics and Software Systems*. 2014. ISBN: 978-4-431-54574-3. URL: http://www.vlebooks.com/vleweb/product/openreader?id=none&isbn=9784431545743 (visited on 03/03/2021) (cit. on pp. 2–9, 24).

[16] David Joyner. *Adventures in Group Theory: Rubik's Cube, Merlin's Machine, and Other Mathematical Toys*. Baltimore: Johns Hopkins University Press, 2002. 262 pp. ISBN: 978-0-8018-6945-7 978-0-8018-6947-1 (cit. on p. 15).

[17] Thomas W Judson and Open Textbook Library. *Abstract Algebra Theory and Applications*. 2016. ISBN: 978-1-944325-02-2. URL: https://open.umn.edu/opentextbooks/textbooks/217,%20http://abstract.pugetsound.edu/ (visited on 04/10/2021) (cit. on pp. 2, 5, 6, 14, 15).

[18] Judy Holdener. *Algebraic Geometry*. 2013. URL: http://pi.math.cornell.edu/~dmehrle/notes/old/alggeo/ (cit. on p. 1).

[19] Ron Larson and Bruce H. Edwards. *Elementary Linear Algebra*. 2nd ed. Lexington, Mass: D.C. Heath, 1991. 592 pp. ISBN: 978-0-669-24592-9 (cit. on pp. 5, 17).

[20] F. J. Martın-Mateos et al. "A Formal Proof of Dickson's Lemma in ACL2". In: *Logic for Programming, Artificial Intelligence, and Reasoning*. Ed. by Moshe Y. Vardi and Andrei Voronkov. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2003, pp. 49–58. ISBN: 978-3-540-39813-4. DOI: 10.1007/978-3-540-39813-4_3 (cit. on p. 22).

[21] Olympia Nicodemi, Melissa A. Sutherland, and Gary W. Towsley. *An Introduction to Abstract Algebra with Notes to the Future Teacher*. Upper Saddle River, NJ: Pearson Prentice Hall, 2007. 436 pp. ISBN: 978-0-13-101963-8 (cit. on pp. 2, 3, 15).

[22] Pablo Parrilo. *Algebraic Techniques and Semidefinite Optimization*. MIT OpenCourseWare. URL: https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-972-algebraic-techniques-and-semidefinite-optimization-spring-2006/ (visited on 04/08/2021) (cit. on p. 1).

[23] Prof. Bernd Sturmfels, director. *Introduction to Grobner Bases - Prof. Bernd Sturmfels*. Jan. 20, 2017. URL: https://www.youtube.com/watch?v=TNO5WuxuNak (visited on 04/12/2021) (cit. on pp. 2, 3).

[24] *Ring of Polynomial Forms Is Integral Domain - ProofWiki*. URL: https://proofwiki.org/wiki/Ring_of_Polynomial_Forms_is_Integral_Domain (visited on 04/12/2021) (cit. on p. 18).

[25] Robert Howlett. *An Undergraduate Course in Abstract Algebra: Course Notes for MATH3002*. URL: https://www.maths.usyd.edu.au/u/bobh/UoS/ (cit. on pp. 2, 12).

[26] Bernd Sturmfels, ed. *Solving Systems of Polynomial Equations*. Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics no. 97. Providence, R.I: Published for the Conference Board of the Mathematical Sciences by the American Mathematical Society, 2002. 152 pp. ISBN: 978-0-8218-3251-6 (cit. on p. 8).

[27] SymPy Development Team. *Basic Functionality of the Module — SymPy 1.8 Documentation*. Sympy Documentation. URL: https://docs.sympy.org/latest/modules/polys/basics.html (visited on 04/11/2021) (cit. on p. 18).

[28] SymPy Development Team. *Gröbner Bases and Their Applications — Polynomials Manipulation Module v1.0 Documentation*. Sympy Documentation. Apr. 9, 2021. URL: https://mattpap.github.io/masters-thesis/html/src/groebner.html (visited on 04/11/2021) (cit. on p. 1).

[29] SymPy Development Team. *Sympy.Polys.Groebnertools — SymPy 1.4 Documentation*. Sympy Documentation. URL: http://www.caacle.com/sympy-docs-html-1.4/_modules/sympy/polys/groebnertools.html (visited on 04/12/2021) (cit. on pp. 1, 2, 10).

[30] SymPy Development Team. *Sympy.Solvers.Polysys — SymPy 1.4 Documentation*. Sympy Documentation. URL: http://man.hubwiz.com/docset/SymPy.docset/Contents/Resources/Documents/_modules/sympy/solvers/polysys.html (visited on 04/11/2021) (cit. on p. 1).

[31] R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2020. URL: http://www.R-project.org/ (cit. on p. 15).

[32] Michael Tinkham. *Group Theory and Quantum Mechanics*. Mineola, N.Y: Dover Publications, 2003. 340 pp. ISBN: 978-0-486-43247-2 (cit. on p. 15).

[33] *Wrong Groebner Basis · Issue #11623 · Sympy/Sympy*. GitHub. URL: https://github.com/sympy/sympy/issues/11623 (visited on 04/11/2021) (cit. on p. 2).