Modern Algebra Lecture Notes: Rings and fields set 7, revision 2

Kevin Broughan

University of Waikato, Hamilton, New Zealand

May 20, 2010

Units in $R[x]$

Let R be an integral domain. Units in $R[x]$ are constant polynomials with values which are units in R: to see this let $f(x) \in R[x]$ be a unit so there exists a $g(x)$ such that $f(x).g(x) = 1$.

Let $f(x) = a_0 + \cdots + a_n x^n,\ g(x) = b_0 + \cdots + b_m x^m$ where $a_n \neq 0,\ b_m \neq 0$. Then

$$f(x).g(x) = 1 = (a_0 + \cdots)(b_0 + \cdots) = a_0 b_0 + \cdots + a_n b_m x^{n+m}.$$

Equating coefficients, $a_0 b_0 = 1$, and if $n + m \geq 1$, $a_n b_m = 0$ which is impossible in an ID.

Therefore a_0 is a unit and $n + m = 0$ so $n = 0$ and we can write $f(x) = a_0$. $\square$

Example

(1) In $\mathbb{Z}[x]$ units are $f(x) = 1,\ g(x) = -1$.
(2) In $\mathbb{R}[x]$ units are constant, non-zero polynomials $f_a(x) = a$ since $a.(1/a) = 1$ in $\mathbb{R}$.

What do elements of $F[x]/\langle f(x)\rangle$ look like ?

Assume $f(x)$ is not constant. We have seen that when $f(x)$ is irreducible over F that $F[x]/\langle f(x)\rangle$ is a field. What is it like ? What do the elements look like ? What are the algebraic operations "+" and ".", given a non-zero element $[g(x)]$ how do we find its inverse ?

Recall $\langle f(x)\rangle = A$ is an ideal so

$$[g(x)] = [h(x)] \iff g(x) - h(x) \in A \iff f(x) \mid g(x) - h(x).$$

So given $g(x)$ write first $g(x) = f(x).q(x) + r(x)$. If $r(x) = 0$ then $g(x) \in A$ so $g(x) - 0 \in A$ so $[g(x)] = [0]$. If $\deg r(x) < \deg f(x)$ then $g(x) - r(x) = f(x).q(x) \in A$ so $[g(x)] = [r(x)]$. That is we can find an element of degree less than $\deg f(x)$ representing the same coset as $g(x)$.

If already $\deg g(x) < \deg f(x)$ then $r(x) = g(x)$ so $[g(x)] = [r(x)]$. In other words we can use polynomials of degree less than $\deg f(x)$ as representatives of elements of the field. Conversely if two of these lower degree polynomials have $r_1(x) \neq r_2(x)$ then their classes cannot be equal because $[r_1(x)] = [r_2(x)]$ implies $f(x) \mid r_1(x) - r_2(x)$ so $\deg f(x) \leq \max\{\deg r_1(x), \deg r_2(x)\} < \deg f(x)$, which is false.

Algebraic operations in $F[x]/\langle f(x)\rangle$: summary

Representatives

Hence, if the degree of $f(x)$ is $n \geq 1$, elements of the quotient field look like $[r(x)]$ where $r(x) = a_0 + \cdots + a_{n-1}x^{n-1}$ where the a_i can be chosen arbitrarily from F giving in each case a unique element. That's what $F[x]/\langle f(x)\rangle$ looks like, namely $[r(x)]$'s. Later we will see a nice vector space description.

Addition

To add classes we, as usual add representatives:
$[r_1(x)] + [r_2(x)] = [r_1(x) + r_2(x)]$ and the sum on the right is just the polynomial with the coefficients of $r_1(x)$, $r_2(x)$ added in F.

Multiplication

To multiply classes we multiply the representatives and then take the remainder after division by $f(x)$: $[r_1(x)][r_2(x)] = [r_1(x).r_2(x)] = [r_1(x).r_2(x) \bmod f(x)]$.

Inverse

To get the inverse of a non-zero class represented by $r(x)$ we use the extended Euclidean algorithm in $F[x]$ to find polynomials $h(x)$, $k(x)$ such that $h(x).r(x) + k(x).f(x) = \gcd(r(x), f(x)) = 1$, and then $[h(x)][r(x)] + 0 = [1]$ so $[h(x) \bmod f(x)]$ is the inverse for $[r(x)]$.

Algebraic operations in $F[x]/\langle f(x)\rangle$: Example

Definitions

Let $F = \mathbb{Q}$, $f(x) = x^3 + 2$, $r_1(x) = x^2 + x + 1$, $r_2(x) = 2x^2 - 3$ so $F[x]/\langle f(x)\rangle = \mathbb{Q}[x]/\langle x^3 + 2\rangle$. Note that $f(x)$ is irreducible over $\mathbb{Z}$ hence $\mathbb{Q}$ by Eisenstein's criteria with $p = 2$.

Addition

Easy: $[x^2 + x + 1] + [2x^2 + 0x - 3] = [3x^2 + x - 2]$

Multiplication

Harder:
$lhs = [x^2+x+1][2x^2-3] = [(x^2+x+1)(2x^2-3)] = [2x^4+2x^3-x^2-3x-3]$.

$$\text{But } 2x^4 + 2x^3 - x^2 - 3x - 3 = (2x + 2)(x^3 + 2) - x^2 - 7x - 7$$

so $lhs = [-x^2 - 7x - 7]$ which is the product.

Getting the inverse

Inverse

Hardest: Use the Euclidean algorithm in $\mathbb{Q}[x]$ (it's the same as in $\mathbb{Z}$) to get

$$\gcd(x^3+2, r_1(x) = 1 = \frac{1}{3}(x^3+2) + \left(\frac{1-x}{3}\right)(x^2+x+1)$$

so the inverse of $[r_1(x)]$ is $[\frac{1-x}{3}]$. To check
$(x^2+x+1)(\frac{1-x}{3}) = 1/3 - x^3/3 \sim 1/3 - (x^3+2)/3 + 2/3 = 1$ as required.

That GCD

That the GCD is 1 follows from the irreducibility of $f(x)$ in $F[x]$, since if the GCD was non-trivial, i.e. a polynomial of degree 1 or more, then since it would also divide $r_1(x)$ it would have degree less than that of $f(x)$, so would be a non-trivial factor of $f(x)$, which unfortunately (or fortunately!) is irreducible. Its fun to think that once you have the equation

$$h(x).f(x) + k(x).r(x) = 1$$

you can't have any common root for $f(x)$ and $r(x)$ in **any** extension field of F! More later about this phenomena.

Introduction

Groups, rings and fields are not enough it seems to meet all the needs for mathematical structures of an algebraic kind. We need **vectors** of finite and infinite dimensions, and we need ways to multiply vectors, such as complex multiplication, quarternionic multiplication, Lie brackets etc These structures rely on group properties for addition, fields for their "scalars" and rings for their multiplication (when it is associative). Structures with vector multiplication are called **algebras**.

Then we have mappings between and within these structures, giving groups of isomorphisms and rings of homomorphisms and "endomorphisms", often modelled by matrix rings. In this paper we will keep most structures quite concrete, and not go to this higher level.

Definition

A vector space V over a field $\mathbb{K}$ has a binary operation "+" giving it an Abelian group structure, and a binary operation "." from $\mathbb{K} \times V \to V$, called scalar multiplication, which satisfies, for all $\alpha,\ \beta \in \mathbb{K}$ and $x, y \in V$:

$$\alpha.(x+y) = \alpha.x + \alpha.y,\ (\alpha+\beta).x = \alpha.x + \beta.x,\ \alpha.(\beta.x) = (\alpha\beta).x,\ 1.x = x.$$

Examples

(1) $\mathbb{K} = \mathbb{R},\ V = \mathbb{R}^n$, the real vector space of dimension n.

(2) $\mathbb{K} = \mathbb{C},\ V = \mathbb{C}^n$, the complex vector space of dimension n.

(3) $\mathbb{K} = \mathbb{Z}/5\mathbb{Z},\ V = \mathbb{Z}/5\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$, a two dimensional vector space over a field with 5 elements.

(4) $\mathbb{K} = \mathbb{Q},\ V = \mathbb{Q}[x]/\langle x^5 + 3\rangle$, a vector space of dimension 5 over $\mathbb{Q}$ which is itself a field, so vectors can be multiplied.

Definitions

A set of vectors $\{x_1, \ldots, x_m\}$ in V is called linearly independent if whenever $\sum_{1 \leq i \leq m} \alpha_i . x_i = 0$ we must have $\alpha_i = 0,\ 1 \leq i \leq m$. These vectors are then all non-zero and distinct. No x_i can be expressed as a linear combination of the other x_j's.

A set of vectors $\{x_1, \ldots, x_m\}$ in V is said to span V if every element $v \in V$ can be expressed as a sum $v = \sum_{1 \leq i \leq m} \alpha_i . x_i$.

A set of vectors $\{x_1, \ldots, x_m\}$ in V is said to be a basis for V if it is both linearly independent and spans V.

A vector subspace of V is a subset W closed under the operations of V, i.e. if $x, y \in W \implies x + y \in W$ and $\alpha . x \in W$ for all $\alpha \in \mathbb{K}$.

If a vector space has a finite basis, its size is denoted by $[V : \mathbb{K}]$ or $\dim_{\mathbb{K}} V$ and called the dimension of the vector space.

Examples

(1) $\mathbb{R}^n$ has a basis $\{e_i : 1 \leq i \leq n\}$ where e_i has the component 1 in the i'th slot or coordinate, and zero elsewhere. Hence $[\mathbb{R}^n : \mathbb{R}] = n$.

(2) The set of polynomials of degree at most n over $\mathbb{Q}$ has a basis $\{1, x, x^2, \ldots, x^n\}$.

(3) The set of all polynomials over a field F has infinite dimension and a basis $\{1, x, x^2, \ldots, x^i, \ldots\}$.

Basis existence theorem

Every vector space has a basis. This is a deep fact and requires the logical axiom of choice or something equivalent like Zorn's lemma. It is only needed for infinite dimensional spaces.

Given a finite set of non-zero vectors $X = \{x_1, \ldots, x_m\}$ in V, the set of all linear combinations $\sum_{1 \leq i \leq m} \alpha_i . x_i$ forms a subspace W of V, and the set of vectors spans W.

We can replace X by a linearly independent set of vectors: $\{x_1\}$ is linearly independent. If it spans W we are done. Then consider $\{x_1, x_2\}$. If it is linearly dependent we must have a relation $\alpha_1 x_1 + \alpha_2 x_2 = 0$ where we must have $\alpha_2 \neq 0$, else this would give a linear dependence equation for $\{x_1\}$. Solve the equation for x_2 to show that the span of $\{x_1\}$ is equal to the span of $\{x_1, x_2\}$ so we can remove x_2.

Proceed inductively retaining or removing vectors until we are left with a subset of vectors which is linearly independent and has the same span as X.

Statement

Let V over $\mathbb{K}$ have a basis of size $n \in \mathbb{N}$. Then every other basis has this same size, so $n = [V : \mathbb{K}]$.

Proof

Let $B = \{b_1, \ldots, b_m\}$ be the given basis and suppose $X = \{x_1, \ldots, x_n\}$ is a linearly independent subset which spans V. Suppose, to get a contradiction $m < n$. (If $m > n$ exchange the roles of X and B.)

Since B spans V we can write $x_1 = \beta_1 b_1 + \ldots \beta_m b_m$ with at least one of the scalars non-zero, since $x_1 \neq 0$. Suppose $\beta_j \neq 0$. Rearrange the equation to express b_j as a linear combination of x_1 and the remaining b_i's. If $B_1 = B \cup \{x_1\} \setminus \{b_j\}$, then B_1 spans V.

Completion of the Dimension Theorem 29 Proof

Keep going, replacing another b_i by x_2, another by x_3 etc, all the while keeping the same span. Note that there will always be a b_j to remove, since at stage $k+1$,

$$x_{k+1} = \alpha_1 x_1 + \cdots + \alpha_k x_k + \text{ a sum of terms } \beta_i b_i$$

and at least one β_i must be non-zero, since otherwise x_{k+1} will be a sum of the other x_i's, and this is impossible, since they are linearly independent.

Having replaced m of the b_i there will be none left, and the spanning set will be $B_m = \{x_1, \ldots, x_m\}$. Then $x_{m+1} = \alpha_1 x_1 + \cdots + \alpha_m x_m$ which implies $0 = \alpha_1 x_1 + \cdots + (-1)x_{m+1} + 0.x_{m+2} + \cdots + 0.x_n$ a non-trivial linear dependence relation among the vectors in X, which is impossible. Hence $m = n$.$\square$