Division rings and theory of equations

by Vivek Mukundan
Vivekananda College, Chennai
Guide : Professor B.Sury

This was the second part of a report on the project done under the JNCASR Summer Fellowship in May-June 2006 in I.S.I.Bangalore

In the first report, I had described the structure theory of semisimple rings and modules and its applications to matrix groups. As mentioned there, the structure theory isolates the division rings as basic objects of study from which semisimple rings are built. Here, I start by studying some aspects of division algebras. After discussing a number of interesting characterizations of commutativity due to Herstein, Jacobson and to Wedderburn etc., the Cartan-Brauer-Hua theorem and the fundamental Skolem-Noether theorem are proved. Following this, the theory of equations over division rings is studied. Results due to Gordon, Motzkin, Bray, Whaples and Niven are demonstrated. Finally, Vandermonde matrices over division rings are discussed culminating in a beautiful result of Lam on their invertibility whose proof uses the ideas above.

§ Division rings and commutativity theorems

Recall that a division ring D is a (not necessarily commutative) ring with unity in which the set D^* of non-zero elements is a group under the multiplication of D. Of course, all fields are division rings. The most familiar example of a division ring which is not a field is that of Hamilton's real quaternions

$$\mathbf{H} = \{a_0 + a_1 i + a_2 j + a_3 k : a_i \in \mathbf{R}\}.$$

Note in this example, $\mathbf{H}$ contains $\mathbf{R}$ as constant quaternions a_0. Thus, $\mathbf{H}$ contains the field $\mathbf{R}$ as a subring which is contained in its center; this is referred to as an $\mathbf{R}$-algebra. In general, if D is a division ring, its center is a field k, and D is a simple (as a ring) k-algebra. One calls D central simple over a field K if its center is K. In this section, we start by studying various properties of division rings which force commutativity. Following that, we discuss the famous result of Frobenius which classifies all the possible division rings with center $\mathbb{R}$. Then, we also prove the important Skolem-Noether theorem.

Examples :

(i) *General quaternion algebra.*
If $a, b \in \mathbf{Q}^*$, the generalized quaternion algebra $D(a,b)$ is defined as follows. Consider formal symbols i, j, k with $i^2 = a, j^2 = b, ij = k = -ji$. One can consider the $\mathbf{Q}$-algebra generated by i, j; that is,

$$D(a,b) = \{a_0 + a_i i + a_2 j + a_3 k\}.$$

The multiplication is dictated by the multiplication of the symbols i, j above. Note in particular that $k^2 = -ab$. Then $D(a,b)$ is a division algebra if, and only if, the equation $ax^2 + by^2 = 1$ has no solution $x, y \in \mathbf{Q}$.

(ii) *Cyclic algebras.*
Let K/F be a cyclic (Galois) extension. Let $G(K/F)$ be the Galois group of K/F and let σ be a generator. Put $s =$ order of σ. Fix $a \in F^*$ and a symbol x. We define

$$D = K.1 \oplus K.x \oplus \cdots \oplus K.x^{s-1}$$

with multiplication described by

$$x^s = a, \quad x.t = \sigma(t)x \ \forall \ t \in K.$$

Then D is an $F-$algebra of dimension s^2 and $F \subseteq Z(D)$. Such an algebra D is called the cyclic algebra associated with σ and a and it is denoted by $(K/F, \sigma, a)$.

$H = (\mathbb{C}/\mathbb{R}, \sigma, -1)$, where σ is complex conjugation, the usual Hamilton quaternion algebra is an example of a cyclic algebra.

Remarks :
Consider the map θ from the division algebra $\mathbf{H}$ of real quaternions to the ring $M_2(\mathbf{C})$ of 2×2 matrices over the complex numbers, defined as follows :

$$a_0 + a_1 i + a_2 j + a_3 k \mapsto \begin{pmatrix} a_0 + a_1 i & a_2 + a_3 i \\ -a_2 + a_3 i & a_0 - a_1 i \end{pmatrix}.$$

This is a ring homomorphism. Note also that every *non-zero* element of $\mathbf{H}$ maps to an invertible matrix. In this manner, $\mathbf{H}$ can be viewed as a subring of $M_2(\mathbf{C})$. Thus, $\mathbf{H}$ is a real form of the complex matrix ring and this aspect has far-reaching generalizations. The theory of Brauer groups of fields (that is of various forms of simple algebras over the field which become isomorphic to a matrix algebra over the algebraic closure of the field) is a major subject of study by itself.

Remarks :
There is no division algebra D which is finite-dimensional as a vector space over $\mathbb{C}$ (or more generally, an algebraically closed field) other than $\mathbb{C}$ itself. The reason is that each element of D outside $\mathbb{C}$ would give a proper finite extension field of $\mathbb{C}$, an impossibility. The following beautiful result of Wedderburn shows a similar fact holds over finite fields also.

Wedderburn's "Little" Theorem
Let D be a finite division ring. Then D is a field. In particular, any finite subring of a division ring is a field.
Proof.
Consider the center F of D; this is a finite field and has cardinality a power of a prime, say $|F| = q = p^d$. Let $n = dim_F D$. We need to show that $n = 1$. Suppose $n > 1$; then D^* is a finite nonabelian group. Look at its class equation. Firstly, we note that if $a \in D$, then its centralizer $C_D(a)$ in D is an F-vector subspace of D; in fact, it is clearly a division ring itself. If $r(a) = dim_F C_D(a)$, then by transitivity of the dimension, we have $r(a)|n$. In

other words, since

$$|F^*| = q - 1, |D^*| = q^n - 1, |C_D(a)^*| = q^{r(a)} - 1$$

the class equation is of the form

$$q^n - 1 = q - 1 + \sum \frac{q^n - 1}{q^{r(a)-1}}$$

where the sum is over various non-singleton conjugacy classes in D^*. Now, one knows that for every natural number m, the cyclotomic polynomial

$\Phi_m(x) = \prod_{(l,m)=1}(x - e^{2i\pi l/m})$ is an irreducible integral polynomial and one has the factorization

$$x^m - 1 = \prod_{l|m} \Phi_l(x).$$

In other words, when $l|m, l < m$, then $\Phi_m(x)$ divides $\frac{x^m-1}{x^l-1}$. Applying this to the class equation, we obtain that $\Phi_n(q)$ divides each term of the sum and hence the term $q-1$ also. However,

$$|\Phi_n(q)| = \prod_{(l,n)=1} |q - e^{2i\pi l/n}| > (q-1)^{\phi(n)} \geq q - 1$$

which contradicts the fact that $\Phi_n(q)$ divides $q-1$. We have used here the inequality $|q - e^{2i\pi l/n}| > q - 1$ which is valid for any $(l,n) = 1$ as $n > 1, q \geq 2$. This contradiction proves that the class equation cannot have any term in the sum and hence D must be commutative.

Remark :
It is an easy exercise in group theory to show that any finite subgroup of K^* for any field K, is cyclic. This is no longer true if one has a division ring instead of a field. For example, note that the quaternion group $\{\pm 1, \pm i, \pm j, \pm k\}$ is a non-cyclic subgroup of $\mathbf{H}^*$. Interestingly, such examples cannot arise if one works with division algebras in positive characteristic, as seen in the following result.

Corollary to Wedderburn's little theorem :
Let D be a division algebra of prime characteristic p (that is, D contains $\mathbf{F}_p$ as a subring of its center. If G is a finite subgroup of D^*, then G is cyclic.

Proof :
Consider the subring

$$K = \{\sum \alpha_i g_i : \alpha_i \in F_p, g_i \in G\}.$$

As K is a finite subring of D, it is a field. Since $G \subseteq K^*$, is a finite subgroup. G is cyclic.

A proposition for membership in center :
Let D be a division Ring. If $y \in D$ commutes with all $ab - ba \forall a, b \in D$ (set of all additive commutators), then $y \in Z(D)$. In particular, if all additive commutators are central, then D is a field.
Proof :
Note that $a(ab - ba)$ is again an additive commutator, namely, $a.ab - ab.a$. Since y commutes with additive commutators

$$x.(xy - yx).y = xy(xy - yx) = yx.(xy - yx).$$

Thus $(xy - yx)^2 = 0 \Rightarrow xy = yx \forall x \in D$.

Proposition :
If D is a division ring, then the division ring R generated by $Z(D)$ and all additive commutators is the whole of D.
Proof :
Let $x \in Z(D)$. Then $\exists y \in D$, such that $xy \neq yx$.Therefore $x(xy - yx) \in R^*$ and $xy - yx \in R^*$.Thus $x \in R^*$.

Definition :
Let D be any ring. Then an additive subgroup of D is said to be a *Lie Ideal* if it is invariant under all inner derivations of D.

Cartan-Brauer-Hua theorem (additive version) :
Let $K \subset D$ be a subring such that K is a Lie ideal in D. Let $charK \neq 2$. Then $K \subseteq Z(D)$.
Proof :
Let $d \in D \backslash K$, $a \in K$. Then $da - ad \in K$(an Lie Ideal). Therefore $d(da - ad) - (da - ad)d \in K$. $\Rightarrow d^2a - 2dad + ad^2 \in K$. Also $d^2a - ad^2 \in K$. Thus, by adding we get $2d(da - ad) \in K$. Thus $da = ad$. For if not, then we would have $d \in K$, a contradiction.
Now let $a, b \in K^*$ and $d \in D^* \backslash K$.Then $db \in D \backslash K$. Thus db commutes with

a. that is $dba = adb$. But, $da = ad$. Therefore, $d(ba - ab) = 0$. This implies $ba = ba$. Hence $K \subseteq Z(D)$.

Herstein's Lemma :
Let D be a division ring of characteristic $p > 0$, $a \in D$, be any noncentral torsion element, which is algebraic over F_p. Then, there exists an element (even a commutator) $y \in D^*$, such that $yay^{-1} = a^i \neq a$.
Proof :
Let $K = F_p[a]$, a finite field. Therefore $a^{p^n} = a$ for some $n > 0$. Consider the inner derivation $ad\,(a) \neq 0$, as $a \notin D$. Also $ad\,(a)$ acts $K-$linearly on D.
Claim : $ad\,(a)$ has a nonzero eigen-vector.
Now $ad\,(a) = L_a - R_a$ and L_a, R_a commute. Therefore, $ad\,(a)^{p^n}\,(x) = (L_a^{p^n} - R_a^{p^n})\,(x) = a^{p^n}x - xa^{p^n} = ax - xa = (ad\,(a))(x)$. Thus $ad(a)^{p^n} = ad(a)$ in $End_K(D)$. Since

$$t^{p^n} = t \prod_{b \in K^*} (t - b),$$

we get

$$0 = ad(a)^{p^n} - ad(a) = ad(a) \prod_{b \in K^*} (ad(a) - b).$$

As ad is not identically zero, some $ad(a) - b$ is not 1-1; let us say, $ad(a_0) - b_0$ for some $b_0 \in K^*$. Therefore, there exists $x \in D^*$ such that $(ad(a) - b_0)(x) = 0$. So $ax - xa = b_0x$; that is, $xax^{-1} = a - b_0 (\neq a) \in \mathbf{F}_p[a]^* = K^*$, which is a cyclic group. Therefore $xax^{-1} = a^i \neq a$. Let $y = ax - xa (= b_0x \neq 0)$. Then $yay^{-1} = a^i$.

Jacobson's commutativity theorem :
Let D be a division ring. For all $a, b \in D$, Suppose there exists a $n(a,b) > 1$ such that $(ab - ba)^{n(a,b)} = ab - ba$. Then D is a field.
Proof :
Suppose not. Therefore $ab - ba \notin Z(D)$.Then $ab - ba$ has finite order. Consider $z \in Z(D)$; then $z(ab - ba) = azb - zba$ also has finite order. Therefore $(ab - ba)^r = (z(ab - ba))^r = 1$ for some r. This implies $z^r = 1$ that is char $D > 0$. Observe that $ab - ba$ is algebraic over F_p. (the polynomial being $t^{n(a,b)} - t = 0$ for a,b).
By Herstein's lemma, there exists an additive commutator y such that

$$y(ab - ba)y^{-1} = (ab - ba)^i \neq ab - ba \cdots\cdots (i)$$

Since $< y >$ normalizes $< ab - ba >$ and both are finite groups, so is $< y >< ab - ba >$. But then it has to be cyclic, a contradiction of (i).

Frobenius's theorem on real division algebras :
Let D be a algebraic division algebra over $\mathbf{R}$. Then, it is either $\mathbf{R}$, $\mathbf{C}$,or $\mathbf{H}$.
Proof :
Assume without loss of generality that $dim_{\mathbf{R}} D \geq 2$. Then for any $\alpha \in D \backslash \mathbf{R}$, we have $\mathbf{R}[\alpha] \cong \mathbf{C}$ as it is a finite nontrivial extension of $\mathbf{R}$. Fix any such copy of $\mathbf{C}$ inside D and view D as a left $\mathbf{C}$-vector space. Also, let D^+, D^- denote the eigensubspaces of D for multiplication by i. That is,

$$D^{\pm} = \{d \in D : di = \pm id\}.$$

Then D^+ and D^- are subspaces and $D = D^+ \oplus D^-$. If $d^+ \in D^+$, then d^+ commutes with $\mathbf{C}$ and so $\mathbf{C}[d^+] \cong \mathbf{C}$, the copy we started with; therefore $D^+ = \mathbf{C}$. If $D^- = \{0\}$, then $D \cong \mathbf{C}$.
If $z \in D^- \backslash \{0\}$, then $x \rightarrow xz$ is a $\mathbf{C}$-linear isomorphism from D^+ to D^-; therefore, $dim_{\mathbf{R}} D = 4$. As z is algebraic over $\mathbf{R}$, the field $\mathbf{R}[z]$ has degree 2 over $\mathbf{R}$. Note that $zi = -iz$ implies that z^2 commutes with i which means that $z^2 \in D^+ = \mathbf{C}$, the copy we started with. So

$$z^2 \in (\mathbf{R} + \mathbf{R}z) \cap \mathbf{C} = \mathbf{R}.$$

Writing $z^2 = \pm r^2$ with $r \in \mathbf{R}$, we must have the minus sign since $z \notin \mathbf{R}$. Thus $z^2 = -r^2$. Putting $j = \frac{z}{r}$, we have $j^2 = -1, ij = -ji$ since $j \in D^-$. Also, note that

$$D = D^+ \oplus D^+ z = \mathbf{C} \oplus \mathbf{C}j = \mathbf{R} \oplus \mathbf{R}i \oplus \mathbf{R}j \oplus \mathbf{R}ij.$$

So D is a copy of $\mathbf{H}$, the real quaternions.

A beautiful result on division algebras is the Cartan-Brauer-Hua theorem. The following result which is of independent interest, is also useful in proving the Cartan-Brauer-Hua theorem as well as in proving that D^* cannot be nilpotent unless it is abelian.

Lemma :
Let D be a division Ring. If $a \in D$ commutes with all the commutators $xyx^{-1}y^{-1}$, then $a \in Z(D)$.
Proof :

Suppose $a \notin Z(D)$. Then there exists an element b such that $ab \neq ba$. Then $b \neq 0, -1$ and so $b^{-1}, (b+1)^{-1}$ exist and we have

$$1 - aba^{-1}b^{-1} = 1 + aba^{-1} - aba^{-1} - aba^{-1}b^{-1}$$

$$= a(b+1)a^{-1} - aba^{-1}b^{-1}(b+1)$$

$$= (a(b+1)a^{-1}(b+1)^{-1} - aba^{-1}b^{-1})(b+1).$$

Now as a commutes with all commutators, and thus, it commutes with the left hand side as well as with the two terms within the first bracket on the right hand side. Therefore a commutes with $b+1$, and hence with b itself, which is a contradiction.

Cartan-Brauer-Hua theorem (multiplicative version) :
Let $A \subseteq D$ be a division subring stable under all inner conjugations of D. Then $A = D$ or $A \subseteq Z(D)$.
Proof :
Assume $A \neq D$. Let $a \in A^*$, $b \in D \backslash A$. Using the identity in the lemma we have $a^{-1} - ba^{-1}b^{-1} = ((b+1)a^{-1}(b+1)^{-1} - ba^{-1}b^{-1})(b+1)$. Since the left hand side is in A and $(b+1)a^{-1}(b+1)^{-1} - ba^{-1}b^{-1} \in A$, and since $b+1 \notin A$, we have that the left hand side must be 0. Thus $ab = ba$. Let now $a' \in A$. Then $a'b \in D \backslash A$.
So $a.a'b = a'b.a = a'.ba = a'.ab$; that is, $aa' = a'a$. Therefore $A \subseteq Z(D)$.

Corollary :
(i) *Let D be a division ring and assume $d \in D \backslash Z(D)$. Then D is generated by the conjugates of d.*
(ii)*If D is a noncommutative division ring, then it is generated as a division ring by all $xyx^{-1}y^{-1}$.*

Theorem (nilpotence implies abelian) :
Let D be a division ring and

$$\{1\} \subseteq G_1 \subseteq G_2 \subseteq \cdots$$

be the upper central series of D^*; that is,

$$G_1 = Z(D^*), G_{i+1}/G_i = Z(D^*/G_i), \cdots$$

Then

$$G_1 = G_2 = \cdots\cdots$$

Hence D^* is nilpotent if and only if D is a field.

Proof :

We shall use the Carter-Brauer-Hua theorem. Let $a \in G_2 \backslash G_1$. So $axa^{-1}x^{-1} \in G_1 \forall x \in D^*$. So $a \notin G_1 = Z(D)$. Therefore there exists $b \in D^*$ such that $ab \neq ba$. From the identity

$$1 - aba^{-1}b^{-1} = \{a(b+1)a^{-1}(b+1)^{-1} - aba^{-1}b^{-1}\}(b+1)$$

we have

$$(a(b+1)a^{-1}(b+1)^{-1} - aba^{-1}b^{-1})(b+1) \in Z(D).$$

Since $Z(D)$ is a field, we have $b+1 \in Z(D)$, which is a contradiction.

Remarks and definitions :

Let $F = Z(D)$ be the center of a division ring D. If $f(t) \in F[t]$, and if $a \in D$ is a root of f, then so is any conjugate of a. Also note that if $a \in D$ is algebraic over F (that is, it satisfies a nonzero polynomial over F), then so are all its conjugates and they have the same minimum polynomial over F, which is called the *minimum polynomial* of the conjugacy class. In fact, if D is algebraic over F, and $a \in D$, then any other root of the minimal polynomial of a must be conjugate to a in D ! This follows from the following very important and widely used result :

Skolem-Noether theorem :

Let A be a central simple algebra over K. Let B be a simple K algebra. Let $\sigma, \tau : B \to A$ be two algebra homomorphisms. Then there exists an inner automorphism $Int(a)$ of A such that $\tau = Int(a) \circ \sigma$.

Proof :

Consider the case $A = End(V)$ for a K vector space V. Then V is also a $A-$module. Via σ and τ we can view V as a B module in two ways. Call them V_σ and V_τ. Since all B simple modules have to be isomorphic by Schur's lemma, there exists a B-isomorphism $f : V_\tau \to V_\sigma$ that is $\forall b \in B, x \in V, f(\tau(b)x) = \sigma(b)(f(x))$. Hence $\tau(b) = f^{-1}\sigma(b)f$. Since $f \in A$, we have the result in this case.

In the general case we consider $B \otimes_K A^o$ for B and $A \otimes_K A^{op}$ for A where A^{op} denotes the opposite algebra of A. Consider $\sigma \otimes id, \tau \otimes id$ from $B \otimes_K A^{op} \to A \otimes_K A^{op} \cong End_K(A)$.

The last isomorphism is seen as follows. For $a \in A, b \in A^{op}$, the map $\phi : A \to A$ given by $\phi(x) = axb$ is an endomorphism of the K vector space A.

Now the map $a \otimes b \to \phi$ is the required isomorphism.
By the first case there exists an $\alpha \in A \otimes_K A^{op}$ such that

$$(\sigma \otimes id)(x) = \alpha(\tau \otimes id)(x)\alpha^{-1} \forall x \in B \otimes A^{op} \cdots (i)$$

Hence $\alpha \in C_{A \otimes A^{op}}(1 \otimes A^{op}) = A \otimes 1$.
Writing $\alpha = a \otimes 1$ and applying (i) to $x = c \otimes 1$, we get $\sigma(c) = a\tau(c)a^{-1}$. Note that the first case applies because $B \otimes A^o$ is simple for the general reason that whenever X is a central simple algebra and Y is simple over K, then $X \otimes_K Y$ is simple.

§ Polynomials over division algebras

Over a field K, a non-zero polynomial f can have at the most deg f roots; this is easy to see using the remainder theorem. However, already over $\mathbf{H}$, we see that i, j, k etc. are all roots of the polynomial $t^2 + 1$. Moreover, our familiar intuitions from equations over fields often fails in many other ways. For example, over a field, a polynomial with a factor of the form $t - a$ evidently vanishes when evaluated at a. However, look at the polynomial $(t - i)(t - j) = t^2 - (i + j)t + ij$ over the Hamilton quaternion division algebra $\mathbf{H}$. The value at i is

$$i^2 - (i + j)i + ij = ij - ji \neq 0!$$

Note however that the value at j is 0. A careful look at this aspect reveals the following. In the above, we are writing polynomials in the form $c_0 + c_1t + c_2t^2 + \cdots + c_nt^n$ with the constants c_i on the left and the powers of the variable t on the right. When we specialize a value of t, obviously the value depends on whether the variable has appeared to the left or to the right. As we shall see, if we write polynomials in the above familiar form with the coefficients to the left, various facts like remainder theorem hold good when we look at 'right' remainders. If we consider a polynomial of the form $g(t)(t - a)$ (with the same convention of writing coefficients on the left), it will turn out that the polynomial vanishes when evaluated at a. Similarly, if we were to write polynomials with the coefficients on the right, we would have a 'left' remainder theorem etc. In this section, we study polynomial equations over noncommutative division rings and describe the various points of departure from equations over fields. The results are surprising and interesting. Without further ado, let us discuss these aspects now.

Firstly, here is a curious characterization of division rings using linear equations. Note that a general linear equation over a noncommutative ring is of the form $\sum_{i=1}^{r} a_i x b_i = c$. The equation which has been studied is the equation of the form $ax - xb = c$. The following result characterizes division rings in terms of solutions of this type of equations.

Let R be a ring with unity. Assume that the equation $ax - xb = c$ is solvable in x whenever $a \neq b$. Then R is a division ring. Further, if each such equation *has a unique solution, then R is a field.*

Proof :

For the first statement, consider $a \in R, a \neq 0, b = 0$ and $c = 1$. Then the equation reads $ax = 1$. Let $x = a_1$ be a solution Then $aa_1 = 1$. Note that $a_1 \neq 0$ and so we also have a_2 such that $a_1a_2 = 1$. Thus $a = a(a_1a_2) = (aa_1)a_2 = a_2$. Thus each $a \in R$ is invertible and hence R is a division ring. For the second statement, let us suppose R to be a noncommutative ring which admits a solution for each equation of the form $ax - xb = c$ with $a \neq b$. If $ab \neq ba$ for some a, b, then the equation $abx - xba = 0$ has two different solutions 0 and a. So the above type of equations over R can have unique solutions only if R is commutative.

Open question :

Is there a division ring which is not a field and admits solutions for every equation of the form $ax - xb = c$ with $a \neq b$?

Let R be any ring and D be a division ring contained in R. If $f(t) = \sum a_it^i \in R[t]$, we define the value $f(r) := \sum a_ir^i$. We call $r \in R$ a right root of $f(t) = \sum a_it^i$, if $f(r) = 0$.

Note that $\sum_i a_ir^i$ may not be equal to $\sum_i r^ia_i$.

In particular, if $f(t) = g(t)h(t)$, then we may not have $f(r) = g(r)h(r)$; that is, the 'evaluation map' may not be a homomorphism. However, sticking to this 'evaluation' map, an easy observation is the right factor theorem stated next. Following that is a key lemma which tells us what one can say about a root of $g(t)h(t)$ which is not a root of $h(t)$.

Right factor theorem :

Let R be any ring and $r \in R$ is a root of $f(t)$ if and only if $t - r$ is a right divisor of $f(t)$ in $R[t]$. The set of polynomials having r as a root is the left ideal $R[t](t - r)$.

Lemma :

Let $f(t) = g(t)h(t) \in D[t]$. If $d \in D$ is such that $h(d) = a \neq 0$, then $f(d) = g(ada^{-1})h(d)$. Consequently, if d is a root of f but not a root of h, then ada^{-1} is a root of g.

Proof:

Let $g(t) = \sum_{i=1}^m b_it^i$. Then $f(t) = \sum_{i=1}^m b_ih(t)t^i$. So

$$f(d) = \sum_{i=1}^m b_ih(d)d^i = \sum_{i=1}^m b_iad^i = \sum_{i=1}^m b_iad^ia^{-1}a = \sum_{i=1}^m b_i(ada^{-1})^ia = g(ada^{-1})h(d).$$

Corollary :
If $f(t) = (t - a_1)(t - a_2)\cdots(t - a_n)$ where $a_i \in D$. Then every root of f is conjugate to one of the a_i's.

As we observed earlier, $x^2 + 1 = 0$ has infinite roots over the division ring $\mathbf{H}$ (division ring of real quaternions). The following result is a generalisation of the familiar result that over a field, a polynomial has at the most its degree number of roots :

Theorem (**Gordon-Motzkin**) :
Let D be a division ring. If $f(t) \in D[t]$, then all its roots lie in at most n conjugacy classes where $\deg f = n$.
Proof:
We proceed by induction on n. For $n = 1$ it is obvious. Suppose $n \geq 2$ and let r be a root of f.Then by the proposition above $f(t) = g(t)(t-r)$ for some $g(t) \in D[t]$ of degree $n - 1$. If $s \neq r$ is another root of f, then by the lemma, s is conjugate to a root of g which in turn lies in one of the $n - 1$ conjugacy classes by the induction hypothesis. Thus by induction, we have the result.

Lemma :
Let A be an algebraic conjugacy class in D(over F) with minimum polynomial $f(t) \in F[t]$. If a polynomial $h(t) \in D[t]\backslash\{0\}$ vanishes identically on A, then $\deg h \geq \deg f$.
Proof :
Suppose not. Pick a polynomial $h(t) = t^m + a_1 t^{m-1} + \cdots + a_m$ such that $h(A) = 0$ and $m < \deg f$. Since $h(t) \notin F[t]$, some $a_i \notin F$. Therefore there exists an element $b \in D^*$ so that $ba_i \neq a_i b$. Clearly now $a^m + a_1 a^{m-1} + \cdots + a_m = 0$ $\forall a \in A$. Conjugating by b, we have

$$(bab^{-1})^m + (ba_1b^{-1})(bab^{-1})^{m-1} + \cdots + (ba_mb^{-1}) = 0 \ \forall a \in A \cdots (I)$$

But since $bab^{-1} \in A$,

$$(bab^{-1})^m + a_1(bab^{-1})^{m-1} + \cdots + a_m = 0 \ \forall a \in A \cdots (II)$$

From (I) and (II), we get that $\sum (ba_ib^{-1} - a_i)t^{m-i}$ vanishes on A; this contradicts the choice of m and proves the lemma.

Corollary :
If $h(t) \in D[t]$ vanishes on A if and only if $h(t) \in D[t]f(t)$.

Proof :
If $f(a) = 0 \forall a \in A$, then $f(t) \in D[t](t-a)$. If $h(t) \in D[t]f(t)$, clearly $h(t) \in D[t](t-a)$ and therefore $h(A) = 0$. Conversely, if $h(A) = 0$, $h(t) \neq 0$, then by division algorithm, $h(t) = g(t)f(t)+r(t)$, with $\deg r(t) < \deg f(t)$, where $f(t)$ is the minimum polynomial of the conjugacy class A. Since $h(A) = f(A) = 0$, we have $r(t) = 0$. Thus, from the lemma above we have $\deg h > \deg f$. Thus $h(t) = g(t)f(t) \Rightarrow h(t) \in D[t]f(t)$.

Corollary :
Let D be an infinite division ring. Then if $h(t) \in D[t]$ is such that $h(d) = 0 \forall d \in D$, then $h(t) = 0$, the zero polynomial.
Proof :
Suppose not. Pick a monic polynomial h of least degree such that $h(d) = 0$ $\forall d \in D$. Let $h(t) = t^m + a_1 t^{m-1} + \cdots + a_m$. We get that all $a_i \in Z(D) = F$ (similar to the argument used in the lemma). Therefore, $h(t) \in F[t]$. Since $h(F) = 0$, F is finite. Now $h(D) = 0 \Rightarrow D$ is algebraic over F, and so D is commutative. This is a contradiction.

Theorem(**Dickson**) :
Let $a, b \in D$ be algebraic over F. Then a is conjugate to b in D if and only if they have the same minimum polynomial.
Proof :
Clearly, if a, b are conjugates, then they have the same minimum polynomial. Conversely, suppose f is the common minimum polynomial for a and b. Regarding f as an element of $F(a)[t]$, $f(t) = g(t)(t-a) = (t-a)g(t)$ by the remainder theorem over fields. Since $\deg g < \deg f$ and f is the minimum polynomial of b, by the lemma there exists some conjugate xbx^{-1} of b in D such that $g(xbx^{-1}) \neq 0$. But $f(t) \in F[t]$ is the minimum polynomial of b, and so $f(aba^{-1}) = 0$ since $f(xbx^{-1}) = 0$, $g(xbx^{-1}) \neq 0$, therefore some conjugate of xbx^{-1} is a zero of $t-a$ by the lemma prior to the Gordon-Motzkin theorem. Thus a is conjugate to b.

Theorem (**Wedderburn**) :
Let A be a conjugacy class which is algebraic over F. Let f denote its minimum polynomial over F and suppose $n = \deg f$. Then there exist $a_1, a_2, \cdots, a_n \in A$ such that $f(t) = (t-a_n)\cdots(t-a_1)$. Moreover $a_1 \in A$ can be chosen arbitrarily. Further the decomposition of f can be cyclically permuted.
Proof :
Let $a_1 \in A$ be arbitrary. Since $f(a_1) = 0$,Therefore $f(t) = g(t)(t-a_1)$

for some $g(t)$. If $A = \{a_1\}$, then $a_1 \in F$ and therefore $f(t) = (t - a_1)$. If $A \neq \{a_1\}$, there exists a conjugate a_2' of a_1 such that $a_2' \neq a_1$. Since $f(a_2') = 0$. Therefore g vanishes at some conjugate a_2 of a_2'. Therefore we can write $g(t) = g_2(t)(t - a_2)$. that is $f(t) = g_2(t)(t - a_2)(t - a_1)$. Proceeding this way, we get $f(t) = g_r(t)(t - a_r) \cdots (t - a_1)$ with r maximum possible. This implies by the above discussion that $\{a_1, \cdots, a_r\} = A$. Since $h(t) := (t - a_r) \cdots (t - a_1)$ vanishes identically on A, we have $h(t) \in D[t]f(t)$. That is, $\deg h \geq \deg f$. But $f(t) = g_r(t)h(t)$. Therefore $f(t) = h(t)$. Finally, cyclic permutations are possible because a factorization of a polynomial in $F[t]$ into two factors in $D[t]$ is necessarily commutative; that is, if $\alpha \in F[t]$, $\alpha = \beta_1(t)\beta_2(t)$, $\beta_i \in D[t]$, then $\beta_1(t)\beta_2(t) = \beta_2(t)\beta_1(t)$).

Corollary :
With the same notations as above, if $f(t) = t^n + d_1 t^{n-1} + \cdots + d_n \in F[t]$, then d_1 is a sum of the elements of A and $(-1)^n d_n$ is the product of the elements of A.

Remarks :
From the above theorem we note that there are infinitely many factorizations. This is because, from the above theorem, we saw that a_1 was arbitrary. We know that A is infinite unless $A = \{a_1\}, a_1 \in F$. The next theorem deals further on the above theme for polynomial equations; it shows that polynomials with at least 2 conjugate zeroes has infinitely many.

Theorem (Gordon-Motzkin) :
Let D be a division ring and $f(t) = \sum_{i=0}^{n} a_i t^i \in D[t]$. Let A be a conjugacy class in D. Assume that f has atleast two zeroes in A. Then f has infinitely many zeroes in A. In particular, for $f = 0$, this means that $|A| \geq 2 \Rightarrow |A|$ is infinite.
Proof :
Fix any $a \in A$. If some dad^{-1} is a zero of f, then $\sum a_i da^i = 0$. So, we must look for $d \in D^*$ such that $\sum a_i da^i = 0$. Define $\Phi : D \to D$; such that $\Phi(d) = \sum a_i da^i$. Then $\Phi(dz) = \Phi(d)z \forall z \in C_D(a)$. Therefore the centralizer $C_D(a)$ of a acts on $\ker \Phi$ on the right. We have a map $\theta : D^* \cap \ker \Phi \to$ zeroes of f in A. $\theta(d) = dad^{-1}$. Note $\theta(d) = \theta(d')$ if and only if $d \in d'.C_D(a)$. Thus the set of zeroes of f in A is in bijection with $\ker \Phi \backslash \{0\}/C_D(a)$, the projective space of the right $C_D(a)$ vector space $\ker \Phi$. We are given that $\ker \Phi$ has dim ≥ 2 over $C_D(a)$. Thus, the corresponding projective space $\ker \Phi \backslash \{0\}/C_D(a)$ is infinite, since $C_D(a)$ is infinite (because D is not commutative since $|A| \geq 2$).

Corollary :
If $f(t) \in D[t]$ has degree n and Γ be the set of roots in D, then either $|\Gamma| \leq n$ or Γ is infinite.
Proof :
Suppose $|\Gamma| > n$ and let $a_1, \cdots, a_{n+1} \in \Gamma$ be distinct. Since the zero of f lie in at most n conjugacy classes, atleast two of the a_i's are conjugate. By the above theorem, the corresponding conjugacy class intersects Γ in an infinite set.

Remarks :
Over fields, we know that given n distinct points $c_1, \cdots, c_n$, there exists a unique polynomial of degree $\leq n$ vanishing at $c_1, \cdots, c_n$; namely, the polynomial $(t - c_1) \cdots (t - c_n)$. The analogue for division rings is the following theorem.

Theorem(**Bray-Whaples**) :
Let D be a division ring and let $c_1, \cdots, c_n$ be pairwise nonconjugate elements of D. Then there exists a unique polynomial $f(t) \in D[t]$ such that $f(c_i) = 0 \forall i$. Further, this polynomial necessarily satisfies :
(a) $c_1, \cdots, c_n$ are its only zeroes,
(b) if $h(t) \in D[t]$ vanishes at all the c_i 's, then $h(t) \in D[t]f(t)$.
Proof :
The uniqueness is clear. This is because the difference of two such polynomials would be a polynomial vanishing at n points in n distinct conjugacy classes while having degree $\leq n - 1$, which is a contradiction. To see the existence, we proceed by induction on n.
For $n = 1$, $f(t) = t - c_1$ clearly. For $n = 2$, choose d_2 so that $(t - d_2)(t - c_1)$ vanishes at c_2 (this clearly means $d_2 = (c_2 - c_1)c_2(c_2 - c_1)^{-1}$). Proceeding in this way, we can get clearly $f(t)$in the form $(t - d_n) \cdots (t - d_2)(t - c_1)$ where d_i is the conjugate of c_i. This proves the existence of f also.
To prove (b), we divide h by f and write $h(t) = q(t)f(t) + r(t)$. Since $h(c_i) = 0$, and $(q(t)f(t))(c_i) = 0$, $r(c_i) = 0$. And $\deg r < n$ and r vanishes at points from n distinct conjugacy classes. Thus $r(t) \equiv 0$.
To prove (a), we again proceed by induction on n. For $n = 1$, it is clear. Assume the result for all $m < n$. Write $f_n(t) = (t - d)f_m(t)$. where $f_n(t)$ is the unique polynomial vanishing at $c_1, \cdots, c_n$ and $f_m(t)$ is the one vanishing at $c_1, \cdots, c_{n-1}$. Observe that d is the conjugate of c_n. Suppose $f_n(c) = 0$ for some $c \in D$. If $f_m(c) = 0$, then c must be one among $c_1, \cdots, c_{n-1}$ by induc-

tion hypothesis. Therefore, assume $f_m(c) \neq 0$. Then c is a conjugate of d and therefore of c_n.
We claim that $d = c_n$. Suppose not. Consider the polynomial $g(t) = (t-e)(t-c_n)$ where e is chosen so that $g(c) = 0$ where e is chosen so that $g(c) = 0$. Indeed $e = (c-c_n)e(c-c_n)^{-1}$. Divide f_n by g and write $f_n(t) = q(t)g(t) + r(t)$, where $\deg r < \deg g = 2$. Since f_n vanishes at c and c_n and since $(q(t)g(t))$ vanishes at c and c_n, therefore $r(t)$ vanishes at c and c_n. But $\deg r \leq 1$. Thus, the only possibility is $r \equiv 0$ (as $c \neq c_n$). Therefore $f_n(t) = q(t)g(t)$. Therefore $\deg q \leq n-2$. But g does not vanish at $c_1, \cdots, c_{n-1}$, as any zero of g is conjugate of c_n. Now q vanishes at conjugates of $c_1, \cdots, c_{n-1}$, which is a contradiction since $\deg q \leq n-2$. Thus $c = c_n$.

Analogous to fields being algebraically closed, there is a notion of a division ring being *right-algebraically-closed.* D is defined to be so if every nonconstant polynomial f in one variable over D has a right root in D. A theorem of Baer says that the only noncommutative division ring which is right algebraically closed is necessarily the ring of quaternions over a real closed field. Recall that a field is said to be real closed if (like in $\mathbb{R}$) -1 is not a sum of squares in it.

Lemma :
Let D be any division ring with center F, and A be a conjugacy class of D which has a quadratic minimum polynomial $\lambda(t)$. If $f(t) \in D[t]$ has two roots in A, then $f(t) \in D[t]\lambda(t)$ and $f(A) = 0$.

Proposition(**Niven**) :
Take $\mathbb{R}$ to be a real closed field, and $D = \mathbb{R} \oplus \mathbb{R}i \oplus \mathbb{R}j \oplus \mathbb{R}k$ be the ring of quaternions over $\mathbb{R}$. For $0 \neq f(t) \in D[t]$, the following are equivalent:
(i)*$f(t)$ has infinitely many roots in D.*
(ii) *There exist $a, b \in \mathbb{R}$ with $b \neq 0$ such that $f(a+ib) = 0 = f(a-ib)$.*
(iii) *f has a right factor $\lambda(t)$ which is an irreducible quadratic in $\mathbb{R}[t]$.*
If these three equivalent conditions hold for f then f vanishes on the conjugacy class of $a+bi$.

Proof :
Assume (i) holds. Then $f(t)$ has two roots in certain conjugacy class A. The minimum polynomial $\lambda(t)$ of A over $\mathbb{R}$ is an irreducible quadratic over $\mathbb{R}$. By the above lemma, we have $f(t) \in D[t]\lambda(t)$. Thus, being quadratic its two roots are of the form $a+bi$ and $a-bi$. This proves (ii) holds.
Now, assume (iii). Let c be the root of $\lambda(t)$ in $\mathbb{R}(i)$. c has infinitely many

conjugates in D. All these are roots of $\lambda(t)$ and hence of $f(t)$. This proves (iii) from (i).
Now (ii) $\Rightarrow$ (iii) is easily deduced from the remainder theorem.

Here is a beautiful result on polynomials over division rings over real closed fields.

Proposition :
Let $D, \mathbf{R}$ be as above, and let $f(t) = \sum_{i=0}^{n} a_i t^i$, where $a_0 \in D \backslash \mathbf{R}$, and $a_1, \cdots, a_n \in \mathbf{R}$. Then f has at most n roots in D.
Proof :
Let α be any root of f. Then α commute with $\sum a_i \alpha^i = -a_0$. Therefore $\alpha \in C_D(\mathbb{R}(\alpha_0))$. But in this case $\mathbb{R}(\alpha_0)$ is a maximal subfield of D and therefore $C_D(\mathbb{R}(\alpha_0)) = \mathbb{R}(\alpha_0)$. Thus $\alpha \in \mathbb{R}(\alpha_0)$. Thus every root is in the field $\mathbb{R}(\alpha_0)$. Since f has at most n roots in a field, the proposition is proved.

Corollary (**Niven**) :
For $a \in D \setminus \mathbb{R}$, the equation $t^n = a$ has exactly n solutions in D and all of them lie in $\mathbb{R}(\alpha_0)$.

Vandermonde matrices

In the 18th century, the mathematician Vandermonde isolated the theory of determinants as a subject for independent study. The following type of matrix is usually known as a Vandermonde matrix :

$$V(a_1,\cdots,a_n) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ \vdots & \vdots & & \vdots \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{pmatrix}$$

where $a_1,\cdots,a_n$ are arbitrary complex numbers. It can be proved by induction that the determinant of this matrix is $\prod_{i>j}(a_i - a_j)$. The Vandermonde matrices evidently arise while solving polynomial equations. Indeed, if $f(x) = c_0 + c_1x + \cdots + c_{n-1}x^{n-1}$ is a polynomial whose values at the n points $a_1,\cdots,a_n$ are $b_1\cdots,b_n$ respectively, then we have the matrix equation

$$\begin{pmatrix} c_0 & c_1 & \cdots & c_{n-1} \end{pmatrix} V(a_1,\cdots,a_n) = \begin{pmatrix} b_1 & b_2 & \cdots & b_n \end{pmatrix}.$$

Therefore, once the points a_i are distinct, the Vandermonde matrix can be inverted and the polynomial can be obtained uniquely.

If we work with a division ring D, one has the notion of right (or left) D-vector spaces of finite dimension (which is a well-defined notion), and any D-vector space linear transformation can be represented by a matrix whose entries are from D. The composition of transformations leads to the definition of matrix multiplication. Thus, it makes sense to say that a matrix is nonsingular if it has an inverse. Over fields (respectively, general commutative rings with unity), this is also equivalent to the determinant being nonzero (respectively, a unit). The problem now is how to define the determinant in our noncommutative situation.

The Vandermonde matrix $V(a_1,\cdots,a_n)$ over a field is evidently nonsingular if, and only if, the a_i are distinct. The first thing we notice that this is *false* over division ring. For example, over H, the division ring of real quaternions, the Vandermonde matrix $V(i,j,k) = \begin{pmatrix} 1 & 1 & 1 \\ i & j & k \\ -1 & -1 & -1 \end{pmatrix}$ is clearly singular because the rows are dependent. However, notice that all the 3 elements i,j,k are conjugate.

Thus, if a notion of determinant can be defined over D it would have to be subtle. Dieudonne defined a notion of determinant which is a map from the set of all invertible matrices (of all sizes) over D to the abelian group $D^*/[D^*, D^*]$. We do not go into the definition of this subtle notion here but rather discuss a result of T.Y.Lam which gives a natural sufficient condition for the invertibility of the Vandermonde matrix whose proof involves the theory of equations as we discussed above.

Before proceeding, we first recall matrix mutiplication over division rings and discuss singularity of 3×3 Vandermonde matrices so as to motivate the general result to be proved. Let us start by recalling how matrix multiplication is defined when the entries are from a division ring. If A, B are $n \times n$ matrices with entries from a division ring D, define AB to be the matrix whose (i, j)-th entry is $\sum_{k=1}^n a_{ik} b_{kj}$. We must take care to keep the order of multiplication of the entries. This definition can be justified as follows. The matrix represents the D-module endomorphism of the n-dimensional right D-vector space with the ordered basis the columns $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$. The transformation represented by A is nothing but the map :

$$\begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11}d_1 + \cdots + a_{1n}d_n \\ a_{21}d_1 + \cdots + a_{2n}d_n \\ \vdots \cdots\cdots\cdots \vdots \\ a_{n1}d_1 + \cdots + a_{nn}d_n \end{pmatrix}.$$

Then, note that AB represents the transformation $A \circ B$. In this set-up, doing elementary row operations on a matrix involve multiplication of rows by scalars from the left. Likewise, doing elementary column operations on a matrix involve multiplication of columns by scalars from the right. It is easy to see that the left D-vector space generated by rows of a matrix and the right D-vector space of its columns have the same dimension; this common dimension is called the rank of the matrix. It may be that the right D-vector space generated by the rows may not have the same dimension.
Look at a 3×3 Vandermonde matrix (with a, b, c distinct)

$$V(a, b, c) = \begin{pmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{pmatrix}.$$

Then, doing two row operations, we have

$$\begin{pmatrix} 1 & 0 & 0 \\ -a & 1 & 0 \\ 0 & -a & 1 \end{pmatrix} V(a,b,c) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & b-a & c-a \\ 0 & (b-a)b & (c-a)c \end{pmatrix}.$$

Note aleady that 'taking $b-a$ and $c-a$ common' would involve *premultiplying the column by a scalar* which is not allowed by an elementary transformation. That is the reason, the noncommutative Vandermonde determinant is more complicated. Premultiplying the above matrix by $\begin{pmatrix} 1 & 0 & 0 \\ 0 & (c-a)^{-1} & 0 \\ 0 & 0 & (c-a)^{-1} \end{pmatrix}$, we get the matrix

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & (c-a)^{-1}(b-a) & 1 \\ 0 & (c-a)^{-1}(b-a)b & c \end{pmatrix}.$$

Yet another row operation (premultiplying by $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -c & 1 \end{pmatrix}$) leads finally to the matrix

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & (c-a)^{-1}(b-a) & 1 \\ 0 & (c-a)^{-1}(b-a)b - c(c-a)^{-1}(b-a) & 0 \end{pmatrix}.$$

Therefore, we note that this last matrix is singular (that is, the columns are D-linearly dependent) if, and only if,

$$(c-a)^{-1}(b-a)b - c(c-a)^{-1}(b-a) = 0;$$

that is, if and only if

$$(b-a)b(b-a)^{-1} = (c-a)c(c-a)^{-1}.$$

In fact, this computation is what leads to a definition of the Dieudonne determinant which we have not gone into. Now, we can state the main final result.

Theorem (T.Y.Lam)
Let $\Delta = \{a_1, ...a_n\}$ be a subset of a division ring D. If no three of the a_i's

lie in a single conjugacy class, then the Vandermonde matrix $V_n(a_1, ...a_n)$ is invertible.

Remarks and definitions :
Before starting the proof, we remark that this theorem gives a sufficiency criterion for the Vandermonde matrix to be invertible. We also saw that when the criterion is not satisfied, the Vandermonde *may* not be invertible. However, there are examples of division algebras where this criterion is not satisfied but the Vandermonde matrix is invertible. Thus, this theorem is the best one could hope for.
A subset $\Delta \subseteq D$ is said to be algebraic if there exists a nonzero polynomial $f(t) \in D[t]$ which is zero on Δ. The set of polynomials vanishing on Δ forms a left ideal in $D[t]$. It is principal and the monic generator is called the minimum polynomial of Δ; it is denoted by f_Δ and the degree of f_Δ will be the *rank* of Δ.
An element $d \in D$ is said to be *P-dependent* (or polynomial-dependent) on Δ if every polynomial in $D[t]$ vanishing on Δ also vanishes on d. Further an algebraic set Δ is *P-independent* if no element $b \in \Delta$ is P-dependent on $\Delta \backslash \{b\}$.
A subset $B \subseteq D$, is said to be a *P-basis* if B is P-independent, and every $d \in D$ is P-dependent on B. With these notations, we can prove the theorem now.

Proof of theorem :
To prove that $V_n(a_1, ...a_n)$ is invertible, we have to show that rank of $V_n(a_1, ...a_n)$ is n. We give the proof in steps as follows.

Step 1 :
We claim that rank $V_n(a_1, ...a_n) =$ rank Δ.
Proof :
Let r and c denote, respectively, the row rank and the column rank of $V_n(a_1, ...a_n)$. We know already that $c = r$ but the proof here proceeds literally by showing that $r \leq \delta \leq c$ where $\delta =$ rank Δ.
Note that a polynomial $g(t) = \sum_{i=1}^{n-1} b_i t^i$ vanishes on Δ if and only if

$$(b_o, ...b_{n-1})V_n(a_1, ...a_n) = 0.$$

To show that $\delta \leq c$ it suffices to find a nonzero polynomial $g(t)$ of degree $\leq c$ such that $g(\Delta) = 0$.
Among the columns $C_1, ..., C_n$ of $V_n(a_1, ...a_n)$, there are c of them which

form a basis of the column space. Assume that these are the first c columns. Let $g(t) = \sum b_i t^i$ be the minimum polynomial of the set $\{a_1, ...a_c\}$. Then $\deg g \leq c \leq n$ and $(b_o, ...b_{n-1}).C_j = 0$ for $1 \leq j \leq c$. Since C_j is a right linear combination of $C_1, \cdots, C_c$, we have $(b_0, ...b_{n-1}).C_j = 0$ and so $(b_o, ...b_{n-1})V_n(a_1, ...a_n) = 0$. That is, $g(\Delta) = 0$.
Next we show that $r \leq \delta$. Let $f = f_\Delta(t)$ be the minimal polynomial of Δ. It suffices to show that each row R_i of $V_n(a_1, ...a_n)$ is a left linear combination of the first δ rows $R_1, ...R_\delta$. By the (left) division algorithm, we can write $t^i = q(t)f(t) + (d_0 + d_1 t + ... + d_{\delta-1}t^{\delta-1})$ since $\delta =$ rank Δ is the degree of f. Evaluating at a_j, we have $a_j^i = d_0 + d_1 a_j + ... + d_{\delta-1}a_j^{\delta-1}$. Thus $R_i = d_0 R_1 + d_1 R_2 + ... + d_{\delta-1}R_\delta$ which means that rank $V_n(a_1, ...a_n) =$ rank Δ.

Step 2 :
Let B, B' be algebraic subsets of D each of which is P-independent such that no element of B is conjugate to an element of B'. Then $B \cup B'$ is also P-independent.
If not, let $c \in B$ is P-dependent on $\Omega := B_o \cup B'$, where $B_o = B \backslash \{c\}$. Let $\mathbf{C}$ be the conjugacy class of D determined by an element $c \in D$. If c is P-dependent on a algebraic set Ω, then c is P-dependent on $\Omega \cap \mathbf{C}$. For, assume that h is the minimum polynomial of $\Omega \cap \mathbf{C}$.Then for any $d \in \Omega \backslash \mathbf{C}$, we have $h(d) \neq 0$. Let $g(t)$ be the minimum polynomial of the set $\Gamma := \{h(d)dh(d)^{-1} : d \in \Omega \backslash \mathbf{C}\}$ and let $f(t) = g(t)h(t)$. Thus f vanishes on $\Delta \backslash \mathbf{C}$ and on $\Delta \cap \mathbf{C}$. Thus f vanishes on Δ and on c. If $h(c) \neq 0$, then $g(t)$ must vanish on $h(c)ch(c)^{-1} \in \mathbf{C}$.Since g is the minimum polynomial of Γ, each root of g is conjugate to some $d \in \Omega \backslash \mathbf{C}$. This contradiction implies that $h(c) = 0$. By hypothesis, $\Omega \cap \mathbf{C}$ is disjoint from B', so $\Omega \cap \mathbf{C} \subseteq B_o$. Thus c is P-dependent on B_o, contradicting the P-independence of B.

Step 3 :
Let Δ, Δ' be algebraic sets in D such that no element of Δ is conjugate to an element of Δ'. Let B, B' be P-bases for Δ and Δ'. Then $rank(\Delta \cup \Delta') = rank\Delta + rank\Delta'$.
If a polynomial vanishes on $B \cup B'$, then it vanishes on $\Delta \cup \Delta'$. Therefore $\Delta \cup \Delta'$ is algebraic and every element of it is P-dependent on $B \cup B'$. By step 2, we have $B \cup B'$ is P-dependent, from which it follows that $B \cup B'$ is a P-basis for $\Delta \cup \Delta'$. Thus $rank(\Delta \cup \Delta') = rank\Delta + rank\Delta'$.

Step 4 :

Let Δ be the algebraic set in D. Then $rank\Delta = \sum rank(\Delta \cap \mathbf{C})$, where C ranges over the finitely many conjugacy classes which intersects Δ. Further if $|\Delta \cap \mathbf{C}| \leq 2$ for each $\mathbf{C}$ then $|\Delta| < \infty$ and Δ is P-independent.
The first statement is proved from step 3 and induction. If there exists only one conjugacy class which intersects Δ, then the statement immediately follows. If there are two conjugacy classes which intersect Δ, then the statement follows from step 3. Let the statement be true for $n-1$ conjugacy classes. Suppose there exists n conjugacy classes which intersects Δ. Let the conjugacy classes be $\mathbf{C}_i$ for $1 \leq i \leq n$ and let $\Delta_i = \Delta \cap \mathbf{C}_i$.Let $\Delta' = \cup_{i=1}^{n-1}\Delta_i$. Again, by step 3 we have $rank(\Delta_{n-1} \cup \Delta') = rank\Delta_{n-1} + rank\Delta'$.
Thus we have $rank\Delta = \sum_{i-1}^{n} rank(\Delta_i) = \sum_{i-1}^{n} rank(\Delta \cap \mathbf{C}_i)$.
The second statement of step 4 follows from the first statement and the fact that no doubleton set is P-independent.

Let us see how the proof follows from these steps. Now since no three a_i's are in the same conjugacy class, the hypothesis of step 4 is fulfilled and we have $\Delta = \{a_1, ...a_n\}$ is P-independent. Inductively we may assume that $rank(\Delta\backslash\{a_1\}) = n-1$. Let f be the minimum polynomial of Δ and g be the minimum polynomial of $\Delta\backslash\{a_1\}$. Then f is a left multiple of g but $f \neq g$. Hence $\deg f \geq 1 + \deg g = n$.
On the other hand, $\deg f \leq |\Delta| = n$, hence rank $\Delta = \deg f = n$.
Since rank $\Delta = n$, it follows from step 1 that $V_n(a_1, \cdots, a_n)$ is invertible. This completes the proof.