Polynomials with non commutative coefficients: If R is any ring, not necessarily commutative, define the polynomial ring R[t] as usual, but where powers of t commute with all coefficients in R, although the coefficients may not commute among themselves.

Hence $f(t) = \sum a_i t^i = \sum t^i a_i$, but if we set t = c, where c is in R, it makes a difference whether we set t = c in the first or the second of these expressions. We call $f_r(c) = \sum a_i c^i$ the right value of f at c, and $f_l(c) =$
$\sum c^i a_i$, the left value of f at c.

Remainder theorem: If f(t) is a polynomial in R[t], then we can write $f(t) = (t-c)q(t) + f_l(c) = p(t)(t-c) + f_r(c)$, i.e. we can divide f(t) by (t-c) from the left, with remainder the left value of f at c, and similarly from the right. The quotients and remainders are unique if we require the remainder belong to R.
proof: We do it for left evaluations and left division. This is the binomial theorem, i.e. replace t in f(t), by (t-c)+c and expand. We get in each term $t^i a_i$, terms in which all but the last have a factor of (t-c), i.e.
$t^i a_i = [(t-c)+c]^i a_i = [(t-c)q(t) + c^i] a_i$. Thus $f(t) = \sum t^i a_i = (t-c)Q(t) + \sum c^i a_i$, and we see the remainder is indeed the left evaluation of f at c.
This proves existence. For uniqueness, assume $f(t) = (t-c)q(t)+r = (t-c)(p(t)+s$, where r,s belong to R. Then $(t-c)[q(t)-p(t)] = s-r$. Thus the left hand side also belongs to R. But multiplication by (t-c) raises the degree by one, so the left ahnd side has degree ≥ 1, unless $[q(t)-p(t)] = 0$. then also r-s = 0. Hence both quotient and remainder are unique. **QED.**

Corollary: If f(t) is any polynomial in R[t], f is left divisible by (t-c) if and only if $f_l(c) = 0$. Similarly for right divisibility.
proof: The expression we gave shows that $f(t) = (t-c)q(t) + f_l(c)$, Hence if $f_l(c) = 0$, then f is left divisible by (t-c). Conversely, if f is left divisaible by (t-c), uniqueness shows the remainder, which is zero, must equal $f_l(c)$, so $f_l(c)$ 0. **QED.**

Next to apply these results about divisibility of polynomials, to products of matrices, we prove that matrices with polynomial entries are equivalently polynomials with matrix coefficients.

Lemma: If k is a field, the non commutative ring $Mat_n(k[t])$ of n by n matrices with entries from k[t], is isomorphic to $Mat_n(k)[t]$, the ring of polynomials with coefficients in the non commutative ring $Mat_n(k)$.
proof: Just as with commutative rings, a ring map R[t]-->S is obtained from a ring map R--->S plus a choice of element in S to send t to, only this time, since t commutes with R in R[t], we must choose as image of t, an element that commutes with the image of R in S. So we map $Mat_n(k)$ into $Mat_n(k[t])$ by viewing scalar matrices as polynomial matrices, and then send t to

the matrix t.I, which is in the center of $Mat_n(k[t])$, i.e. it commutes with everything. It is an exercise to check this ring map is injective and surjective. **QED.**

It follows that if we have two matrices of polynomials and we multiply them as matrices, we get the same result by viewing them as polynomials with matrix entries, and multiplying them as polynomials.

Corollary: Cayley Hamilton. A square matrix A over a commutative ring R, is a root of its characteristic polynomial $ch_A(t)$.
proof: By Cramer's rule, we have $(tI-A).adj(tI-A) = ch_A(t).I$, as products of matrices. Then it holds also as products of polynomials. Setting t = A gives zero on the left, hence also on the right side. I.e. if $ch_A(t) = \sum t^i c_i$, where the c_i belong to R, then $ch_A(t).I = (\sum t^i c_i).I = \sum t^i(c_i.I)$. Thus setting t = A gives $0 = \sum A^i(c_i.I) = \sum A^i(c_i) = \sum c_i A^i = ch_A(A)$. **QED.**

If in the lemma above, we think of the matrix on the left acting individually on each column vector of the matrix on the right, we can also consider matrices of polynomials acting on column vectors of polynomials, as multiplication from the left of polynomials with matrix coefficients, times polynomials with column vector coefficients. I.e. the lemma also holds, with the same proof, for polynomials with coefficients in any ring R with identity, acting from the left on polynomials with coefficients in any (unitary) left module over R.

So let $k^n[t]$ denote polynomials with coefficients which are column vectors from k^n. This is not a ring, in particular the coefficents do not have an element 1, so this object does not contain t. But the coefficients do contain the basic vectors e_i, and we can multiply these by polynomials over k and add up. In particular this object is a k[t] module, and is isomorphic as such to the free k[t] module $(k[t])^n$.

I.e. if E_i are the standard free k[t] basis vectors in $(k[t])^n$, just send E_i to e_i, and $\sum f_i E_i$ to $\sum f_i e_i$ where f_i are polynomials in k[t]. The expression $\sum f_i e_i$ can be re - expanded as a polynomial in t with vector coefficients by expanding each term as $fe_i = (a_0+a_1t+...+t^n)e_i = (a_0e_i + t\, a_1e_i +...+ t^ne_i)$, and then combining coefficients of like powers of t, from various terms, to get coefficient vectors.

Exercise: Show this gives a k[t] module isomorphism $(k[t])^n \to k^n[t]$.
As we have remarked above, the previous lemma, shows multiplication of matrices corresponds to multiplication of polynomials, i.e. the isomorphisms above, give isomorphisms of multiplication diagrams with matrix multiplication $Mat_n(k[t]) \times (k[t])^n \to (k[t])^n$, corresponding to polynomial multiplication $Mat_n(k)[t] \times k^n[t] \to k^n[t]$.

Now we can prove the main presentation theorem.

Theorem: Given any n by n matrix A over a field k, defining a k[t] module structure on k^n, the

k[t] module map $(k[t])^n$--->k^n, sending $\sum f_i(t)E_i$ to $\sum f_i(A)e_i$, is surjective, and its kernel is a free k[t] module, freely generated by the columns of [tI-A], the characteristic matrix of A. I.e. this sequence is exact: 0--->$(k[t])^n$--->$(k[t])^n$--->k^n--->0, as k[t] - modules, where the left map is multiplication by [tI-A].

proof: We know the last map is surjective.

Recall the right map takes $\sum f_i(t)E_i$ to $\sum f_i(A)e_i$, which is exactly the result of viewing $\sum f_i(t)E_i$ as a polynomial $\sum f_i(t)e_i$ with coefficient vectors in k^n, and then setting t = A. So if we view these as maps of polynomials $k^n[t]$--->$k^n[t]$--->k^n--->0, the right map $k^n[t]$--->k^n, is left evaluation of a polynomial f(t) with vector coefficients, at t = A. By the factor theorem above, this is zero if and only if f(t) is left divisible by (t-A), i.e. if and only if f(t) is in the image of the left hand map $k^n[t]$--->$k^n[t]$.

Since multiplication by a monic polynomial never sends a non zero polynomial to zero, the left map is injective. Hence the sequence 0--->$(k[t])^n$--->$(k[t])^n$--->k^n--->0 is exact, and (tI-A) is indeed a presentation matrix for the module (k^n,A). **QED.**