from the direct sum decomposition is reversible. Suppose $b_1(x), b_2(x), \dots, b_t(x)$ are monic polynomials in $F[x]$ of degree at least one such that $b_i(x) \mid b_{i+1}(x)$ for all i and suppose for some basis $\mathcal{E}$ of V, that the matrix of T with respect to the basis $\mathcal{E}$ is the direct sum of the companion matrices of the $b_i(x)$. Then V must be a direct sum of T-stable subspaces D_i, one for each $b_i(x)$ in such a way that the matrix of T on each D_i is the companion matrix of $b_i(x)$. Let $\mathcal{E}_i$ be the corresponding (ordered) basis of D_i (so $\mathcal{E}$ is the union of the $\mathcal{E}_i$) and let e_i be the first basis element in $\mathcal{E}_i$. Then it is easy to see that D_i is a cyclic $F[x]$-module with generator e_i and that the annihilator of D_i is $b_i(x)$. Thus the torsion $F[x]$-module V decomposes into a direct sum of cyclic $F[x]$-modules in two ways, both of which satisfy the conditions of Theorem 5, i.e., both of which give lists of invariant factors. Since the invariant factors are unique by Theorem 9, $a_i(x)$ and $b_i(x)$ must differ by a unit factor in $F[x]$ and since the polynomials are monic by assumption, we must have $a_i(x) = b_i(x)$ for all i. This proves the following result:

Theorem 14. *(Rational Canonical Form for Linear Transformations)* Let V be a finite dimensional vector space over the field F and let T be a linear transformation of V.

(1) There is a basis for V with respect to which the matrix for T is in rational canonical form, i.e., is a block diagonal matrix whose diagonal blocks are the companion matrices for monic polynomials $a_1(x), a_2(x), \dots, a_m(x)$ of degree at least one with $a_1(x) \mid a_2(x) \mid \cdots \mid a_m(x)$.

(2) The rational canonical form for T is unique.

The use of the word *rational* is to indicate that this canonical form is calculated entirely within the field F and exists for any linear transformation T. This is not the case for the Jordan canonical form (considered later), which only exists if the field F contains the eigenvalues for T (cf. also the remarks following Corollary 18).

The following result translates the notion of similar linear transformations (i.e., the same linear transformation up to a change of basis) into the language of modules and relates this notion to rational canonical forms.

Theorem 15. Let S and T be linear transformations of V. Then the following are equivalent:

(1) S and T are similar linear transformations

(2) the $F[x]$-modules obtained from V via S and via T are isomorphic $F[x]$-modules

(3) S and T have the same rational canonical form.

Proof: [(1) implies (2)] Assume there is a nonsingular linear transformation U such that $S = UTU^{-1}$. The vector space isomorphism $U : V \to V$ is also an $F[x]$-module homomorphism, where x acts on the first V via T and on the second via S, since for example $U(xv) = U(Tv) = UT(v) = SU(v) = x(Uv)$. Hence this is an $F[x]$-module isomorphism of the two modules in (2).

[(2) implies (3)] Assume (2) holds and denote by V_1 the vector space V made into an $F[x]$-module via S and denote by V_2 the space V made into an $F[x]$-module via T. Since $V_1 \cong V_2$ as $F[x]$-modules they have the same list of invariant factors. Thus S and T have a common rational canonical form.

[(3) implies (1)] Assume (3) holds. Since S and T have the same matrix representation with respect to some choice of (possibly different) bases of V by assumption, they are, up to a change of basis, the same linear transformation of V, hence are similar.

Let A be any $n \times n$ matrix with entries from F. Let V be an n-dimensional vector space over F. Recall we can then *define* a linear transformation T on V by choosing a basis for V and setting $T(v) = Av$ where v on the right hand side means the $n \times 1$ column vector of coordinates of v with respect to our chosen basis (this is just the usual identification of linear transformations with matrices). Then (of course) the matrix for this T with respect to this basis is the given matrix A. Put another way, any $n \times n$ matrix A with entries from the field F arises as the matrix for some linear transformation T of an n-dimensional vector space.

This dictionary between linear transformations of vector spaces and matrices allows us to state our previous two results in the language of matrices:

Theorem 16. *(Rational Canonical Form for Matrices)* Let A be an $n \times n$ matrix over the field F.

(1) The matrix A is similar to a matrix in rational canonical form, i.e., there is an invertible $n \times n$ matrix P over F such that $P^{-1}AP$ is a block diagonal matrix whose diagonal blocks are the companion matrices for monic polynomials $a_1(x), a_2(x), \dots, a_m(x)$ of degree at least one with $a_1(x) \mid a_2(x) \mid \cdots \mid a_m(x)$.

(2) The rational canonical form for A is unique.

Definition. The *invariant factors* of an $n \times n$ matrix over a field F are the invariant factors of its rational canonical form.

Theorem 17. Let A and B be $n \times n$ matrices over the field F. Then A and B are similar if and only if A and B have the same rational canonical form.

If A is a matrix with entries from a field F and F is a subfield of a larger field K then we may also consider A as a matrix over K. The next result shows that the rational canonical form for A and questions of similarity do not depend on which field contains the entries of A.

Corollary 18. Let A and B be two $n \times n$ matrices over a field F and suppose F is a subfield of the field K.

(1) The rational canonical form of A is the same whether it is computed over K or over F. The minimal and characteristic polynomials and the invariant factors of A are the same whether A is considered as a matrix over F or as a matrix over K.

(2) The matrices A and B are similar over K if and only if they are similar over F, i.e., there exists an invertible $n \times n$ matrix P with entries from K such that $B = P^{-1}AP$ if and only if there exists an (in general different) invertible $n \times n$ matrix Q with entries from F such that $B = Q^{-1}AQ$.

Proof: (1) Let M be the rational canonical form of A when computed over the smaller field F. Since M satisfies the conditions in the definition of the rational canonical form over K, the uniqueness of the rational canonical form implies that M is also the rational canonical form of A over K. Hence the invariant factors of A are the same whether A is viewed over F or over K. In particular, since the minimal polynomial is the largest invariant factor of A it also does not depend on the field over which A is viewed. It is clear from the determinant definition of the characteristic polynomial of A that this polynomial depends only on the entries of A (we shall see shortly that the characteristic polynomial is the product of all the invariant factors for A, which will give an alternate proof of this result).

(2) If A and B are similar over the smaller field F they are clearly similar over K. Conversely, if A and B are similar over K, they have the same rational canonical form over K. By (1) they have the same rational canonical form over F, hence are similar over F by Theorem 17.

This corollary asserts in particular that the rational canonical form for an $n \times n$ matrix A is an $n \times n$ matrix with entries in the smallest field containing the entries of A. Further, this canonical form is the same matrix even if we allow conjugation of A by nonsingular matrices whose entries come from larger fields. This explains the terminology of *rational* canonical form.

The next proposition gives the connection between the characteristic polynomial of a matrix (or of a linear transformation) and its invariant factors and is quite useful for determining these invariant factors (particularly for matrices of small size).

Lemma 19. Let $a(x) \in F[x]$ be any monic polynomial.

(1) The characteristic polynomial of the companion matrix of $a(x)$ is $a(x)$.

(2) If M is the block diagonal matrix

$$M = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_k \end{pmatrix},$$

given by the direct sum of matrices $A_1, A_2, \dots, A_k$ then the characteristic polynomial of M is the product of the characteristic polynomials of $A_1, A_2, \dots, A_k$.

Proof: These are both straightforward exercises.

Proposition 20. Let A be an $n \times n$ matrix over the field F.

(1) The characteristic polynomial of A is the product of all the invariant factors of A.

(2) *(The Cayley–Hamilton Theorem)* The minimal polynomial of A divides the characteristic polynomial of A.

(3) The characteristic polynomial of A divides some power of the minimal polynomial of A. In particular these polynomials have the same roots, not counting multiplicities.

The same statements are true if the matrix A is replaced by a linear transformation T of an n-dimensional vector space over F.