Example 2 Let T be the linear operator on $\mathbb{R}^2$ defined by $T(a,b) = (2a+5b, 6a+b)$ and β be the standard ordered basis for $\mathbb{R}^2$. Then $[T]_\beta = \begin{pmatrix} 2 & 5 \\ 6 & 1 \end{pmatrix}$ and hence the characteristic polynomial of T is $f(t) = \det\begin{pmatrix} 2-t & 5 \\ 6 & 1-t \end{pmatrix} = (t-7)(t+4)$

Thus the minimal polynomial of T is also $(t-7)(t+4)$.

Example 3 Let D be the linear operator on $P_2(\mathbb{R})$ defined by $D(g(x)) = g'(x)$, the derivative of $g(x)$. We compute the minimal polynomial of D. Let β be the standard ordered basis for $P_2(\mathbb{R})$. Then $[D]_\beta = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$ and it follows that the characteristic polynomial of D is $-t^3$. So, by corollary 2.15.7, the minimal polynomial of D is t, t^2 or t^3. Since $D^2(x^2) = 2 \neq \theta$, it follows that $D^2 \neq T_\theta$; hence the minimal polynomial of D must be t^3.

Theorem 2.15.8 Let T be a linear operator on an n dimensional vector space V such that V is a T-cyclic subspace of itself. Then the characteristic polynomial $f(t)$ and the minimal polynomial $p(t)$ have the same degree, and hence $f(t) = (-1)^n p(t)$

Proof: Since V is a T-cyclic space, there exists an $x \in V$ such $\beta = \{x, T(x), \dots, T^{n-1}(x)\}$ is a basis for V (Theorem 2.14.4, page-121). Let $g(t) = a_0 + a_1 t + \dots + a_k t^k$ be a polynomial of degree $k < n$. Then $a_k \neq 0$ and we have

$g(T)(x) = a_0 x + a_1 T(x) + \cdots + a_k T^k(x)$ and so $g(T)(x)$ is a linear combination of the vectors of β, having at least one non-zero coefficient, namely, a_k. Since β is linearly independent, it follows that $g(T)(x) \neq \theta$; hence $g(T) \neq T_\theta$. Therefore the minimal polynomial of T has degree n, which is also the degree of the characteristic polynomial of T.

<u>Theorem 2.15.9</u> Let T be a linear operator on a finite dimensional vector space V. Then T is diagonalizable if and only if the minimal polynomial of T is of the form $p(t) = (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_k)$ where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of T.

<u>Proof</u>: Suppose that T is diagonalizable. Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of T, and define

$$p(t) = (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_k).$$

By Theorem 2.15.6, $p(t)$ divides the minimal polynomial of T. Let $\beta = \{v_1, v_2, \ldots, v_n\}$ be a basis for V consisting of eigenvectors of T, and consider any $v_i \in \beta$. Then $(T-\lambda_j I)(v_i) = \theta$ for some eigenvalue λ_j. Since $(t-\lambda_j)$ divides $p(t)$, there is a polynomial $q_j(t)$ such that $p(t) = q_j(t)(t-\lambda_j)$. Hence

$$p(T)(v_i) = q_j(T)(T-\lambda_j I)(v_i) = \theta.$$

It follows that $p(T) = T_0$, since $p(T)$ takes each vector in a basis for V into θ. Therefore $p(t)$ is the minimal polynomial of T.

Conversely, suppose that there are distinct scalars $\lambda_1, \lambda_2, \ldots, \lambda_k$ such that the minimal polynomial $p(t)$ of T factors as

$p(t) = (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_k)$.

By Theorem 2.15.6, the λ_i's are eigenvalues of T. We apply mathematical induction on ~~dim~~ $n = \dim(V)$. Clearly T is diagonalizable for $n=1$. Now suppose that T is diagonalizable whenever $\dim(V) < n$ for some $n > 1$ and let $n = \dim(V)$ and $W = R(T-\lambda_k I)$. Obviously ~~$V \neq W$~~ $W \neq V$, because λ_k is an eigenvalue of T. If $W = \{\theta\}$, then $T = \lambda_k I$, which is clearly diagonalizable. So, suppose $0 < \dim(W) < n$. Then W is T-invariant, and for any $x \in W$,

$$(T-\lambda_1 I)(T-\lambda_2 I)\cdots(T-\lambda_k I)(x) = \theta.$$

It follows that the minimal polynomial of T_W divides the polynomial $(t-\lambda_1)(t-\lambda_2)\cdots \cancel{(t-\lambda_k)}\,(t-\lambda_{k-1})$. Hence by induction hypothesis T_W is diagonalizable. Furthermore, λ_k is not an eigenvalue of T_W by Theorem 2.15.6. Therefore $W \cap N(T-\lambda_k I) = \{\theta\}$. Now let $\beta_1 = \{v_1, v_2, \ldots, v_m\}$ be a basis for W consisting of the eigenvectors of T_W (and hence of T), and let $\beta_2 = \{w_1, w_2, \ldots, w_p\}$ be a basis for $N(T-\lambda_k I)$, the eigenspace of T corresponding to λ_k. Then β_1 and β_2 are disjoint by the previous comment. Moreover,

$\beta = \beta_1 \cup \beta_2$ is linearly independent. Consider scalars $a_1, a_2, \ldots, a_m$ and $b_1, b_2, \ldots b_p$ such that

$$a_1 v_1 + a_2 v_2 + \cdots + a_m v_m + b_1 w_1 + b_2 w_2 + \cdots + b_p w_p = \theta$$

Let $x = \sum_{i=1}^{m} a_i v_i$ and $y = \sum_{i=1}^{p} b_i w_i$

Then $x \in W$, $y \in N(T - \lambda_k I)$ and $x + y = \theta$. It follows that $x = -y \in W \cap N(T - \lambda_k I)$ and therefore $x = \theta$. Since β_1 is linearly independent, $a_1 = a_2 = \cdots = a_m = 0$. Similarly, $b_1 = b_2 = \cdots = b_p = 0$ and we conclude that β is a linearly independent subset of V consisting of n ~~vectors~~ eigenvectors. It follows that β is a basis for V consisting of eigenvectors of T, and consequently, T is diagonalizable.

Example 4 We determine all matrices $A \in M_{2\times 2}(\mathbb{R})$ for which $A^2 - 3A + 2I_2 = O$. Let $g(t) = t^2 - 3t + 2 = (t-1)(t-2)$. Since $g(A) = O$, the minimal polynomial $p(t)$ of A divides $g(t)$. Hence the only possible candidates for $p(t)$ are $t-1$, $t-2$, and $(t-1)(t-2)$. If $p(t) = t-1$ or $p(t) = t-2$ then $A = I$ or $A = 2I$ respectively. If $p(t) = (t-1)(t-2)$, then A is diagonalizable with eigenvalues 1 and 2, and hence A is similar to $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$.