15.2 Reduction to Upper Triangular Form

Unfortunately, not every linear map on a complex vector space can be diagonalized. The next best thing is to "triangularize," which means to find a basis over which the matrix has zero entries below the main diagonal. Fortunately, such a basis always exist.

We say that a square matrix A is an *upper triangular matrix* if it has the following shape,

$$\begin{pmatrix} a_{1\,1} & a_{1\,2} & a_{1\,3} & \dots & a_{1\,n-1} & a_{1\,n} \\ 0 & a_{2\,2} & a_{2\,3} & \dots & a_{2\,n-1} & a_{2\,n} \\ 0 & 0 & a_{3\,3} & \dots & a_{3\,n-1} & a_{3\,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & a_{n-1\,n-1} & a_{n-1\,n} \\ 0 & 0 & 0 & \dots & 0 & a_{n\,n} \end{pmatrix},$$

i.e., $a_{i\,j} = 0$ whenever $j < i$, $1 \leq i, j \leq n$.

Theorem 15.5. *Given any finite dimensional vector space over a field K, for any linear map $f\colon E \to E$, there is a basis $(u_1, \dots, u_n)$ with respect to which f is represented by an upper triangular matrix (in $\mathrm{M}_n(K)$) iff all the eigenvalues of f belong to K. Equivalently, for every $n \times n$ matrix $A \in \mathrm{M}_n(K)$, there is an invertible matrix P and an upper triangular matrix T (both in $\mathrm{M}_n(K)$) such that*

$$A = PTP^{-1}$$

iff all the eigenvalues of A belong to K.

Proof. If there is a basis $(u_1, \dots, u_n)$ with respect to which f is represented by an upper triangular matrix T in $\mathrm{M}_n(K)$, then since the eigenvalues of f are the diagonal entries of T, all the eigenvalues of f belong to K.

For the converse, we proceed by induction on the dimension n of E. For $n = 1$ the result is obvious. If $n > 1$, since by assumption f has all its eigenvalues in K, pick some eigenvalue $\lambda_1 \in K$ of f, and let u_1 be some corresponding (nonzero) eigenvector. We can find $n - 1$ vectors $(v_2, \dots, v_n)$ such that $(u_1, v_2, \dots, v_n)$ is a basis of E, and let F be the subspace of dimension $n - 1$ spanned by $(v_2, \dots, v_n)$. In the basis $(u_1, v_2 \dots, v_n)$, the matrix of f is of the form

$$U = \begin{pmatrix} \lambda_1 & a_{1\,2} & \dots & a_{1\,n} \\ 0 & a_{2\,2} & \dots & a_{2\,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n\,2} & \dots & a_{n\,n} \end{pmatrix},$$

since its first column contains the coordinates of $\lambda_1 u_1$ over the basis $(u_1, v_2, \dots, v_n)$. If we let $p\colon E \to F$ be the projection defined such that $p(u_1) = 0$ and $p(v_i) = v_i$ when $2 \leq i \leq n$, the linear map $g\colon F \to F$ defined as the restriction of $p \circ f$ to F is represented by the