Most readers will probably find the proof of Theorem 7.14 rather clever but very mysterious and unmotivated. The conceptual difficulty is that we really need to understand how polynomials in one variable "act" on vectors in terms of the matrix A. This can be done and yields a more "natural" proof. Actually, the reasoning is simpler and more general if we free ourselves from matrices and instead consider a finite-dimensional vector space E and some given linear map $f\colon E \to E$. Given any polynomial $p(X) = a_0X^n + a_1X^{n-1} + \cdots + a_n$ with coefficients in the field K, we define the *linear map* $p(f)\colon E \to E$ by

$$p(f) = a_0f^n + a_1f^{n-1} + \cdots + a_n\mathrm{id},$$

where $f^k = f \circ \cdots \circ f$, the k-fold composition of f with itself. Note that

$$p(f)(u) = a_0f^n(u) + a_1f^{n-1}(u) + \cdots + a_nu,$$

for every vector $u \in E$. Then we define a new kind of scalar multiplication $\cdot\colon K[X] \times E \to E$ by polynomials as follows: for every polynomial $p(X) \in K[X]$, for every $u \in E$,

$$p(X) \cdot u = p(f)(u).$$

It is easy to verify that this is a "good action," which means that

$$\begin{aligned}
p \cdot (u + v) &= p \cdot u + p \cdot v \\
(p + q) \cdot u &= p \cdot u + q \cdot u \\
(pq) \cdot u &= p \cdot (q \cdot u) \\
1 \cdot u &= u,
\end{aligned}$$

for all $p, q \in K[X]$ and all $u, v \in E$. With this new scalar multiplication, E is a $K[X]$-module.

If $p = \lambda$ is just a scalar in K (a polynomial of degree 0), then

$$\lambda \cdot u = (\lambda\mathrm{id})(u) = \lambda u,$$

which means that K acts on E by scalar multiplication as before. If $p(X) = X$ (the monomial X), then

$$X \cdot u = f(u).$$

Now if we pick a basis $(e_1, \ldots, e_n)$ of E, if a polynomial $p(X) \in K[X]$ has the property that

$$p(X) \cdot e_i = 0, \quad i = 1, \ldots, n,$$

then this means that $p(f)(e_i) = 0$ for $i = 1, \ldots, n$, which means that the linear map $p(f)$ vanishes on E. We can also check, as we did in Section 7.2, that if A and B are two $n \times n$ matrices and if $(u_1, \ldots, u_n)$ are any n vectors, then

$$A \cdot \left(B \cdot \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \right) = (AB) \cdot \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}.$$