remains true when the matrix A replaces the variable λ

$$P(A, A) = \sum_{k=0}^{N} p_k \, A^k = 0. \tag{1.258}$$

To see why, we use the formula (1.196)

$$\delta_{k\ell} \det A = \sum_{i=1}^{N} A_{ik} C_{i\ell} \tag{1.259}$$

to write the determinant $|A - \lambda I| = P(\lambda, A)$ as the product of the matrix $A - \lambda I$ and the transpose of its matrix of cofactors

$$(A - \lambda I)\, C(\lambda, A)^{\mathsf{T}} = |A - \lambda I|\, I = P(\lambda, A)\, I. \tag{1.260}$$

The transpose of the matrix of cofactors of the matrix $A - \lambda I$ is a polynomial in λ with matrix coefficients

$$C(\lambda, A)^{\mathsf{T}} = C_0 + C_1 \lambda + \cdots + C_{N-1} \lambda^{N-1}. \tag{1.261}$$

The left-hand side of equation (1.260) is then

$$\begin{aligned}(A - \lambda I) C(\lambda, A)^{\mathsf{T}} = {} & A C_0 + (A C_1 - C_0)\lambda + (A C_2 - C_1)\lambda^2 + \ldots \\ & + (A C_{N-1} - C_{N-2})\lambda^{N-1} - C_{N-1}\lambda^N.\end{aligned} \tag{1.262}$$

Equating equal powers of λ on both sides of (1.260), we have using (1.257) and (1.262)

$$\begin{aligned} A C_0 &= p_0 I \\ A C_1 - C_0 &= p_1 I \\ A C_2 - C_1 &= p_2 I \\ \ldots &= \ldots \\ A C_{N-1} - C_{N-2} &= p_{N-1} I \\ -C_{N-1} &= p_N I. \end{aligned} \tag{1.263}$$

We now multiply from the left the first of these equations by I, the second by A, the third by A^2, $\ldots$, and the last by A^N and then add the resulting equations. All the terms on the left-hand sides cancel, while the sum of those on the right gives $P(A, A)$. Thus a square matrix A obeys its characteristic equation $0 = P(A, A)$ or

$$0 = \sum_{k=0}^{N} p_k \, A^k = |A|\, I + p_1 A + \cdots + (-1)^{N-1} (\mathrm{Tr} A)\, A^{N-1} + (-1)^N \, A^N \tag{1.264}$$