a result known as the **Cayley-Hamilton theorem** (Arthur Cayley, 1821–1895, and William Hamilton, 1805–1865). This derivation is due to Israel Gelfand (1913–2009) (Gelfand, 1961, pp. 89–90).

Because every $N \times N$ matrix A obeys its characteristic equation, its Nth power A^N can be expressed as a linear combination of its lesser powers

$$A^N = (-1)^{N-1}\big(|A|\, I + p_1 A + p_2 A^2 + \cdots + (-1)^{N-1}(\mathrm{Tr}A)\, A^{N-1}\big). \quad (1.265)$$

For instance, the square A^2 of every 2×2 matrix is given by

$$A^2 = -|A| I + (\mathrm{Tr}A) A. \quad (1.266)$$

Example 1.35 (Spin-one-half rotation matrix) If $\boldsymbol{\theta}$ is a real 3-vector and $\boldsymbol{\sigma}$ is the 3-vector of Pauli matrices (1.32), then the square of the traceless 2×2 matrix $A = \boldsymbol{\theta} \cdot \boldsymbol{\sigma}$ is

$$(\boldsymbol{\theta} \cdot \boldsymbol{\sigma})^2 = -\, |\boldsymbol{\theta} \cdot \boldsymbol{\sigma}|\, I = - \begin{vmatrix} \theta_3 & \theta_1 - i\theta_2 \\ \theta_1 + i\theta_2 & -\theta_3 \end{vmatrix} I = \theta^2\, I \quad (1.267)$$

in which $\theta^2 = \boldsymbol{\theta} \cdot \boldsymbol{\theta}$. One may use this identity to show (exercise (1.28)) that

$$\exp\left(-i\boldsymbol{\theta} \cdot \boldsymbol{\sigma}/2\right) = \cos(\theta/2)\, I - i\hat{\boldsymbol{\theta}} \cdot \boldsymbol{\sigma}\, \sin(\theta/2) \quad (1.268)$$

in which $\hat{\boldsymbol{\theta}}$ is a unit 3-vector. For a spin-one-half object, this matrix represents a right-handed rotation of θ radians about the axis $\hat{\boldsymbol{\theta}}$. □

1.27 Functions of Matrices

What sense can we make of a function f of an $N \times N$ matrix A? and how would we compute it? One way is to use the characteristic equation (1.265) to express every power of A in terms of I, A, $\ldots$, A^{N-1} and the coefficients $p_0 = |A|, p_1, p_2, \ldots, p_{N-2}$, and $p_{N-1} = (-1)^{N-1}\mathrm{Tr}A$. Then if $f(x)$ is a polynomial or a function with a convergent power series

$$f(x) = \sum_{k=0}^{\infty} c_k\, x^k \quad (1.269)$$

in principle we may express $f(A)$ in terms of N functions $f_k(\boldsymbol{p})$ of the coefficients $\boldsymbol{p} \equiv (p_0, \ldots, p_{N-1})$ as

$$f(A) = \sum_{k=0}^{N-1} f_k(\boldsymbol{p})\, A^k. \quad (1.270)$$

The identity (1.268) for $\exp\left(-i\boldsymbol{\theta} \cdot \boldsymbol{\sigma}/2\right)$ is an $N = 2$ example of this technique which can become challenging when $N > 3$.