Generalized Spectral Resolution

& some of its applications

Nicholas Wheeler, Reed College Physics Department

27 April 2009

Introduction. Familiarly, if the $n\times n$ matrix $\mathbb{H}$ is complex hermitian (or, more particularly, real symmetric or imaginary antisymmetric) then the eigenvalues $\{\lambda_1,\lambda_2,\dots,\lambda_n\}$ are necessarily real, and amongst the eigenvectors $\{\mathbf{h}_1,\mathbf{h}_2,\dots,\mathbf{h}_n\}$ those associated with distinct eigenvalues are necessarily orthogonal:

$$\mathbf{h}_i\perp\mathbf{h}_j\quad\text{if}\quad\lambda_i\neq\lambda_j$$

If the spectrum is non-degenerate the eigenvectors (which we might—but won't—assume to have been normalized) provide an orthogonal basis in $\mathcal{V}_n$. More generally, we write $\{(\lambda_1,\delta_1),(\lambda_2,\delta_2),\dots,(\lambda_\nu,\delta_\nu)\}$ where

$$\delta_\alpha\equiv\text{degeneracy of }\lambda_\alpha\quad:\quad\sum_\alpha\delta_\alpha=n$$

and the λ_α are distinct: each such λ_α identifies an "eigenspace" $\mathcal{V}_{(\alpha)}$—a δ_α-dimensional subspace of $\mathcal{V}_n$. Every element of $\mathcal{V}_{(\alpha)}$ is orthogonal to every element of $\mathcal{V}_{(\beta)}$ $(\alpha\neq\beta)$. The set of eigenvectors is similarly partitioned

$$\{(\mathbf{h}_1,\dots,\mathbf{h}_{\delta_1}),(\mathbf{h}_{\delta_1+1},\dots,\mathbf{h}_{\delta_1+\delta_2}),\dots,(\mathbf{h}_{n-\delta_\nu},\dots,\mathbf{h}_n)\}$$

where the vectors $(\mathbf{h}_1,\dots,\mathbf{h}_{\delta_1})$ span $\mathcal{V}_{(1)}$ and can be assumed to have been orthogonalized ("by hand"), *etc*.

With this apparatus in hand, we construct matrices[1]

$$\mathbb{P}_i\equiv\left[\frac{\mathbf{h}\cdot\mathbf{h}^t}{(\mathbf{h},\mathbf{h})}\right]_i=\left[(\mathbf{h},\mathbf{h})^{-1}\begin{pmatrix}h_1\bar h_1 & h_1\bar h_2 & \cdots & h_1\bar h_n\\ h_2\bar h_1 & h_2\bar h_2 & \cdots & h_2\bar h_n\\ \vdots & \vdots & \ddots & \vdots\\ h_n\bar h_1 & h_n\bar h_2 & \cdots & h_n\bar h_n\end{pmatrix}\right]_i$$

[1] Here $(\mathbf{h},\mathbf{h})\equiv\mathbf{h}^t\cdot\mathbf{h}$ and ^{t} signifies conjugated transposition.

which project onto the "eigenrays":

$$\mathbb{P}_i\,\mathbf{h}_i = \mathbf{h}_i$$

From the circumstance that the $\{\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n\}$ comprise—automatically, else (in the case of degeneracy) by contrivance—an orthogonal basis in $\mathcal{V}_n$ it follows that the matrices $\{\mathbb{P}_1, \mathbb{P}_2, \dots, \mathbb{P}_n\}$ comprise a complete

$$\sum_i \mathbb{P}_i = \mathbb{I} \tag{1.1}$$

orthogonal set

$$\mathbb{P}_i\mathbb{P}_j = \mathbb{O} \quad : \quad i \neq j \tag{1.2}$$

of projection matrices

$$\mathbb{P}_i^2 = \mathbb{P}_i \quad : \quad \text{all } i \tag{1.3}$$

in terms of which we have the "spectral resolution of $\mathbb{H}$:

$$\mathbb{H} = \sum_i \lambda_i\,\mathbb{P}_i \tag{2}$$

In degenerate cases we can lump the projectors onto the same eigenspace, writing

$$\mathbb{H} = \sum_\alpha \lambda_\alpha \mathbb{P}_{(\alpha)} \quad \text{with} \quad \mathbb{P}_{(1)} \equiv \mathbb{P}_1 + \mathbb{P}_2 + \cdots + \mathbb{P}_{\delta_1}, \textit{ etc.}$$

where

$$\sum_\alpha \mathbb{P}_{(\alpha)} = \mathbb{I}$$
$$\mathbb{P}_{(\alpha)}\mathbb{P}_{(\beta)} = \mathbb{O} \quad : \quad \alpha \neq \beta$$
$$\mathbb{P}_{(\alpha)}^2 = \mathbb{P}_{(\alpha)} \quad : \quad \text{all } \alpha$$

From (1) and (2) we obtain

$$\mathbb{H}^k = \sum_i \lambda_i^k\,\mathbb{P}_i$$

which in the case $k = 0$ gives back the completeness condition (1.1). For all $f(x)$ that can be expressed as weighted sums of powers we have

$$f(\mathbb{H}) = \sum_i f(\lambda_i\,\mathbb{P}_i) = \sum_i f(\lambda_i)\,\mathbb{P}_i \tag{3}$$

Relaxation of the hermiticity assumption. Let $\mathbb{M}$ be *any* $n \times n$ matrix (no symmetry properties assumed). We proceed from the observation that, while $\mathbb{M}$ and its transpose $\mathbb{M}^\mathsf{T}$ have identical spectra,[2] they can be expected to have distinct eigenvectors. We are led thus to distinguish **right eigenvectors**—defined

[2] So long as $\mathbb{M}$ remains unspecialized we can say nothing about the any special properties of the eigenvalues.

$$\mathbb{M}\,\mathbf{a}_i = \lambda_i\,\mathbf{a}_i$$

—from **left eigenvectors**, defined

$$\mathbb{M}^{\mathsf{T}}\mathbf{b}_i = \lambda_i\,\mathbf{b}_i \quad\text{equivalently}\quad \mathbf{b}_i^{\mathsf{T}}\mathbb{M} = \lambda_i\,\mathbf{b}_i^{\mathsf{T}}$$

Immediately

$$\mathbf{b}_i^{\mathsf{T}}\mathbb{M}\,\mathbf{a}_j = \begin{cases}\lambda_i\cdot\mathbf{b}_i^{\mathsf{T}}\mathbf{a}_j & \text{on the one hand}\\ \lambda_j\cdot\mathbf{b}_i^{\mathsf{T}}\mathbf{a}_j & \text{on the other}\end{cases}$$

from which we conclude that the sets $\{\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n\}$ and $\{\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_n\}$ are **biorthogonal** in the following sense:

$$\mathbf{b}_i\perp\mathbf{a}_j \quad\text{if}\quad \lambda_i\neq\lambda_j$$

Here

$$\mathbf{b}_i\perp\mathbf{a}_j \quad\text{means that}\quad (\mathbf{b}_i,\mathbf{a}_j)\equiv\mathbf{b}_i^{\mathsf{T}}\mathbf{a}_j = 0$$

Note that in the preceding equations we encounter the *simple transpose, not* the conjugated transpose. And that the $\mathbf{a}/\mathbf{b}$ distinction disappears when $\mathbb{M}$ is real symmetric: "biorthogonality" reduces then to "simple orthogonality."

Use the material now in hand to define

$$\mathbb{P}_i \equiv \left[\frac{\mathbf{a}\cdot\mathbf{b}^{\mathsf{T}}}{(\mathbf{a},\mathbf{b})}\right]_i = \left[(\mathbf{a},\mathbf{b})^{-1}\begin{pmatrix}a_1b_1 & a_1b_2 & \cdots & a_1b_n\\ a_2b_1 & a_2b_2 & \cdots & a_2b_n\\ \vdots & \vdots & \ddots & \vdots\\ a_nb_1 & a_nb_2 & \cdots & a_nb_n\end{pmatrix}\right]_i$$

and observe that

$$\left.\begin{aligned}&\mathbb{P}_i\,\mathbf{a}_i = \mathbf{a}_i\\ &\mathbb{P}_i\,\mathbf{a}_j = \mathbf{0} \quad\text{if}\quad \lambda_i\neq\lambda_j \quad\text{(by biorthogonality)}\\ &\mathbf{b}_i^{\mathsf{T}}\mathbb{P}_i = \mathbf{b}_i^{\mathsf{T}}\\ &\mathbf{b}_i^{\mathsf{T}}\mathbb{P}_j = \mathbf{0}^{\mathsf{T}} \quad\text{if}\quad \lambda_i\neq\lambda_j \quad\text{(by biorthogonality)}\end{aligned}\right\} \tag{4}$$

Moreover,

$$\mathbb{P}_i\,\mathbb{P}_i = \mathbb{P}_i \tag{5.1}$$

$$\mathbb{P}_i\,\mathbb{P}_j = \mathbb{O} \quad\text{if}\quad \lambda_i\neq\lambda_j \quad\text{(by biorthogonality)} \tag{5.2}$$

and if the spectrum is non-degenerate it is assuredly the case that

$$\sum_i \mathbb{P}_i = \mathbb{I} \tag{5.3}$$

Finally, we have this universally valid generalization of the familiar spectral resolution formula (2):

$$\mathbb{M} = \sum_i \lambda_i\,\mathbb{P}_i \tag{5.4}$$

In the presence of spectral degeneracies we can by contrivance *arrange* for equations (5) to be valid.

First application: Hamiltonian generators of quantum gates. The controlled evolution of the state $|\Psi\rangle$ is accomplished by the action of "gates," represented by unitary matrices the designs of which is reflect the basic elements of Boolean logic. The action of such gates is quantum dynamical

$$|\Psi\rangle_0 \longrightarrow |\psi\rangle_t = \underbrace{e^{-(i/\hbar)\,\mathbb{H}\,t}}_{\text{becomes } \mathbb{U}_{\text{gate}} \text{ at time } t=1}|\Psi\rangle_0$$

In quantum mechanics we are most commonly given $\mathbb{H}$, and asked to construct $\mathbb{U}(t)$, but here we confront the inverse problem: we are given $\mathbb{U}_{\text{gate}}$ and asked to construct the generator $\mathbb{H}_{\text{gate}}$ of that matrix. Since the unitarity of $\mathbb{U}_{\text{gate}}$ implies the hermiticity of $\mathbb{H}_{\text{gate}}$ we can solve the problem by appeal simply to (2), don't in this instance need the generality of (5.4). I illustrate the procedure by looking to a specific example:

The most common instance of the important cNOT ("controlled NOT") gate is defined[3]

$$\mathbb{U}_{\text{cNOT}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Mathematica supplies $\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\} = \{-1, 1, 1, 1\}$ and the (unnormalized but orthogonal) eigenvectors

$$\mathbf{h}_1 = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}, \quad \mathbf{h}_2 = \begin{pmatrix} 0 \\ 0 \\ +1 \\ 1 \end{pmatrix}, \quad \mathbf{h}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{h}_4 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

from which we obtain the projection matrices

$$\mathbb{P}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & +\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & +\frac{1}{2} \end{pmatrix}, \quad \mathbb{P}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & +\frac{1}{2} & +\frac{1}{2} \\ 0 & 0 & +\frac{1}{2} & +\frac{1}{2} \end{pmatrix}$$

$$\mathbb{P}_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbb{P}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

So we have

$$\mathbb{U}_{\text{cNOT}} = (-1)\mathbb{P}_1 + (+1)\mathbb{Q} \tag{6}$$

---

[3] See N. David Mermin, *Quantum Computer Science: An Introduction* (2007), page 10.

with

$$\mathbb{Q} \equiv \mathbb{P}_2 + \mathbb{P}_3 + \mathbb{P}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Equation (6) can now be written

$$\mathbb{U}_{\text{cNOT}} = e^{i\pi}\,\mathbb{P}_1 + e^{i0}\,\mathbb{Q} = e^{i\pi\,\mathbb{P}_1} = e^{-(i/\hbar)\,\mathbb{H}_{\text{cNOT}}\,t}\Big|_{t=1}$$

with

$$\mathbb{H}_{\text{cNOT}} = -\pi\,\hbar\,\mathbb{P}_1$$

Second application: Continuously interpolated Markoff processes. Let the elements of

$$\mathsf{P} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix}$$

be probabilities, with $\sum_i p_i = 1$. Markoff processes have the structure

$$\mathsf{P}_{k-1} \longrightarrow \mathsf{P}_k = \mathbb{M}\,\mathsf{P}_{k-1} = \mathbb{M}^k\,\mathsf{P}_0$$

where the elements of $\mathbb{M} = \|m_{ij}\|$ are "transition probabilities" and

$$\sum \text{elements of } \mathsf{P}_k = 1 \quad : \quad \text{all k}$$

requires that the columns of $\mathbb{M}$ sum to unity: $\sum_i m_{ij} = 1$ (all j).

Look, for example, to the case

$$\mathbb{M} = \begin{pmatrix} 0.261 & 0.087 & 0.052 \\ 0.006 & 0.042 & 0.862 \\ 0.733 & 0.871 & 0.086 \end{pmatrix}$$

The eigenvalues are $\{\lambda_1, \lambda_2, \lambda_3\} = \{1.000, -0.790, 0.179\}$. It is characteristic of Markoff matrices that one of the eigenvalues is unity, and the others have absolute values that are less than one. The negative eigenvalue is admissible, but has a complex logarithm. To avoid the "complex probabilities" to which we would be led when we construct the matrix $\mathbb{M}^t$ that interpolates between the matrices $\mathbb{M}^k$ $(k = 0, 1, 2, \ldots)$ I therefore adopt this modified definition:

$$\mathbb{M} = \begin{pmatrix} 0.261 & 0.087 & 0.052 \\ 0.006 & 0.042 & 0.862 \\ 0.733 & 0.871 & 0.086 \end{pmatrix}\begin{pmatrix} 0.261 & 0.087 & 0.052 \\ 0.006 & 0.042 & 0.862 \\ 0.733 & 0.871 & 0.086 \end{pmatrix}$$

The spectrum of this $\mathbb{M}$ is the assuredly non-negative square of the previous spectrum (it reads $\{1.000, 0.624, 0.032\}$). We construct the right eigenvectors $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$, the left eigenvectors $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$, and from them assemble

$$\begin{aligned}
\mathbb{P}_1 &= \frac{\mathbf{a}_1 \cdot \mathbf{b}_1^{\mathsf{T}}}{(\mathbf{a}_1, \mathbf{b}_1)} = \begin{pmatrix} 0.085 & 0.085 & 0.085 \\ 0.434 & 0.434 & 0.434 \\ 0.481 & 0.481 & 0.481 \end{pmatrix} \\
\mathbb{P}_2 &= \frac{\mathbf{a}_2 \cdot \mathbf{b}_2^{\mathsf{T}}}{(\mathbf{a}_2, \mathbf{b}_2)} = \text{a matrix of undistinguished appearance} \\
\mathbb{P}_3 &= \frac{\mathbf{a}_3 \cdot \mathbf{b}_3^{\mathsf{T}}}{(\mathbf{a}_3, \mathbf{b}_3)} = \text{ditto}
\end{aligned}$$

which do in fact comprise a complete set of orthogonal projection matrices. We now have

$$\mathbb{M} = \mathbb{P}_1 + (0.624)\,\mathbb{P}_2 + (0.032)\,\mathbb{P}_3$$

giving

$$\begin{aligned}
\mathbb{M}^k &= \mathbb{P}_1 + (0.624)^k\,\mathbb{P}_2 + (0.032)^k\,\mathbb{P}_3 \qquad (7) \\
&\downarrow \\
&= \mathbb{P}_1 \quad \text{in the limit } k \uparrow \infty
\end{aligned}$$

The implication is that *all* initial probability vectors $\mathbf{P}_0$ proceed asymptotically to the state

$$\mathbf{P}_\infty = \frac{\mathbf{a}_1}{\text{sum of the elements of } \mathbf{a}_1} = \begin{pmatrix} 0.085 \\ 0.434 \\ 0.481 \end{pmatrix}$$

Notice that the elements of $\mathbf{P}_\infty$ are precisely the elements that we see repeated in $\mathbf{P}_\infty$ are—for readily understood reasons—precisely the elements that we see repeated in the columns of $\mathbb{P}_1$.

Returning now to (7), we have

$$\begin{aligned}
\mathbb{M} &= e^{\log 1.000}\,\mathbb{P}_1 + e^{\log 0.624}\,\mathbb{P}_2 + e^{\log 0.032}\,\mathbb{P}_3 \\
&= e^{-0.471\,\mathbb{P}_2 - 3.441\,\mathbb{P}_3}
\end{aligned}$$

giving the interpolating matrix

$$\mathbb{M}^t = e^{\mathbb{L}\,t} \quad \text{with} \quad \mathbb{L} = \log \lambda_2 \cdot \mathbb{P}_2 + \log \lambda_3 \cdot \mathbb{P}_3 \qquad (8)$$

In physical applications the elements of $\mathbb{M}$ are subject to a principle of detailed balancing: $m_{ij} = m_{ji}$. The analysis proceeds then not from (5.4) but from the more familiar equation (2). It is found in such cases that $\mathbb{L}$ is symmetric, and that the elements in its columns (rows) sum to zero. And that the calculation typically proceeds $\mathbb{L} \Longrightarrow \mathbb{M}$ rather than (as above) $\mathbb{M} \Longrightarrow \mathbb{L}$, with the structure of $\mathbb{L}$ read directly from (say) the "adjacency matrix" of a graph (as in Matt Jemielita's thesis (2009)).

Third application: Proof of an elegant identity. One frequently encounters arguments that hinge on the identity

$$\det \mathbb{M} = e^{\operatorname{tr}\log \mathbb{M}} \quad, i.e., \quad \log\det \mathbb{M} = \operatorname{tr}\log \mathbb{M}$$

—proofs of which usually pertain only to cases in which $\mathbb{M}$ is equivalent to a diagonal matrix: $\mathbb{M} = \mathbb{S}^{-1}\,\mathbb{D}\,\mathbb{S}$. We are in position now to construct a proof which is subject to no such limitation. For we have

$$\mathbb{M} = \sum_\alpha \lambda_\alpha \mathbb{P}_{(\alpha)} = \exp\Big\{\sum_\alpha \log\lambda_\alpha \mathbb{P}_{(\alpha)}\Big\}$$

where the λ_α are distinct and $\mathbb{P}_{(\alpha)}$ projects onto the δ_α-dimensional eigenspace $\mathcal{V}_{(\alpha)}$. Familiarly, $\det \mathbb{M} = (\lambda_1)^{\delta_1}(\lambda_2)^{\delta_2}\cdots(\lambda_\nu)^{\delta_\nu}$. But the trace of a projection matrix is the dimension of the space onto which it projects, so we have

$$\begin{aligned}\operatorname{tr}\Big\{\sum_\alpha \log\lambda_\alpha \mathbb{P}_{(\alpha)}\Big\} &= \sum_\alpha \delta_\alpha \log\lambda_\alpha \\ &= \log\Big\{(\lambda_1)^{\delta_1}(\lambda_2)^{\delta_2}\cdots(\lambda_\nu)^{\delta_\nu}\Big\} \qquad QED\end{aligned}$$

Concluding comments. In this short note my intent has been to make more conveniently available some of the material of which I made critical use in a *Mathematica* notebook ("New Markoff: Classical/Quantum Markoff Processes (22 April 2009)) written in conjunction with Matt Jemielita's thesis, which is concerned with classical/quantum random walks on graphs. The heart of the note resides in the "generalized spectral representation" (5.4) to which my title refers. I have no doubt that mathematicians would consider (5.4) to be a commonplace triviality, but think it fair to say that (5.4) and its powerful implications are unfamiliar to most physicists—though it was a few lines in a paper by a physicist that introduced me to this topic.[4] To say the same thing another way, one only seldom encounters references in the physics literature to "biorthogonality," though it underlies the entry of the "reciprocal lattice" into the solid state physics of periodic structures (crystals).[5] This has probably to do with the fact that the matrices encountered in physical applications are usually (anti)symmetric or rotational (in either the Euclidean or Lorentzian sense), (anti)hermitian or unitary—seldom asymmetric. Or rectangular, in which context something very like the present line of argument leads to the singular value decomposition (SVD).

---

[4] I allude to remarks on page 418 of Elliott W. Montroll's "Markoff chains, Wiener integrals and quantum theory," Comm. Pure & Appl. Math **5**, 415–453 (1952).

[5] I have explored aspects of this subject in "Recirocal systems of non-orthogonal quantum states" (1998).