SOME UNCOMMON MATRIX THEORY

Nicholas Wheeler
Reed College Physics Department
April 2012

Introduction. In recent work, having mainly to do with continuous classical/quantum random walks on graphs and the quantum theory of open systems, I have repeatedly had need to draw upon several aspects of elementary matrix theory that, while doubtless "well known" in some quarters, seem not to be treated in any of the standard texts. Discussions of the points in question are at present folded—often redundantly—into the texts of a large number of *Mathematica* notebooks, and have proven difficult to retrieve because the notebooks bear names that refer to diverse areas of application, not to incidental points of mathematical method. My primary objective here will be to construct a unified account of that material, to be released from the inefficient distraction of having to "reinvent" such material each time I encounter need of this or that idea or detail.

I have come to belated realization that the ideas in question—though developed independently of one another, to serve a variety of special needs—are in fact interrelated in a variety of interesting ways; I will seek to expose some of those interconnections.

The matrices of interest will in all cases be finite dimensional. Occasionally the reality of the matrix elements will (at least for application-based interpretive purposes) be essential, but more commonly we will assume the elements of $\mathbb{M}$ to be complex in the general case, real as a special case.

Our typical objective, broadly speaking, will be to show how matrices of some frequently-encounterred specified type can be constructed (additively, multiplicatively, by exponentiation...) from matrices of some more specialized type.

Standard spectral decomposition. Let $\mathbb{A}$ be a $d \times d$ hermitian matrix: $\mathbb{A} = \mathbb{A}^+$ where in the finite-dimensional theory $^+$ signifies nothing more abstract than simple conjugated transposition. We avail ourselves of Dirac notation, which in this simple context has the status merely of a handy triviality: let $\mathcal{V}_d$ signify the complex space of column vectors

$$|y) = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_d \end{pmatrix}$$

Application of $^+$ to $|x)$ produces

$$(x| \equiv [|x)]^+ = (\, x_1^*, x_2^*, \ldots, x_d^* \,)$$

—a row vector that lives in the dual space $\mathcal{V}_d^\star$. The inner product

$$(x|y) = x_1^* y_1 + x_2^* y_2 + \cdots + x_d^* y_d$$

is a complex number-valued object that lives with one foot in $\mathcal{V}_d$ and the other foot in $\mathcal{V}_d^\star$. Clearly, $(x|x)$ is real, non-negative, and vanishes iff $|x) = |0)$.

Look to complex numbers of the construction $(x|\mathbb{M}|y)$. Clearly,

$$[(x|\mathbb{M}|y)]^+ = (y|\mathbb{M}^+|x) = \text{complex conjugate of } (x|\mathbb{M}|y)$$

so from the hermiticity assumption $\mathbb{A} = \mathbb{A}^+$ we obtain

$$(y|\mathbb{A}|x) = \text{complex conjugate of } (x|\mathbb{A}|y)$$

which implies the reality of $(x|\mathbb{A}|x)$. Supposing a to be an eigenvalue of $\mathbb{A}$ and $|a)$ the associated eigenvector

$$\mathbb{A}|a) = a|a)$$

we have

$$a = \frac{(a|\mathbb{A}|a)}{(a|a)} \quad \Longrightarrow \quad \text{eigenvalues of hermitian matrices are real}$$

Suppose, moreover, that $\mathbb{A}|a_1) = a_1|a_1)$ and $\mathbb{A}|a_2) = a_2|a_2)$ with $a_1 \neq a_2$. Then

$$\begin{aligned} &(a_2|\mathbb{A}|a_1) = a_1(a_2|a_1) \\ &(a_1|\mathbb{A}|a_2) = a_2(a_1|a_2) \quad \Longrightarrow \quad (a_2|\mathbb{A}|a_1) = a_1^*(a_2|a_1) \end{aligned}$$

which by $a_1^* = a_1 \neq a_2$ implies

$$(a_2|a_1) = 0 \quad : \quad \begin{cases} \text{eigenvectors associated with distinct} \\ \text{eigenvalues of a hermitian matrix } \mathbb{A} \\ \text{are orthogonal} \end{cases}$$

Suppose for the moment that the spectrum $\{a_1, a_2, \ldots, a_d\}$ of $\mathbb{A}$ is non-degenerate. The associated eigenvectors $\{|a_1), |a_2), \ldots, |a_d)\}$, which we may assume to have been normalized, comprise then an orthonormal basis in $\mathcal{V}_d$. The general element $|x) \in \mathcal{V}_d$ can be developed

$$|x) = \xi^1|a_1) + \xi^2|a_2) + \cdots \xi^d|a_d)$$

The matrices

$$\mathbb{P}_k = |a_k)(a_k|$$

—which are clearly projective and orthogonal

$$\mathbb{P}_j\mathbb{P}_k = \delta_{jk}\mathbb{P}_k$$

serve to project out the respective components of vectors $|x)$

$$\mathbb{P}_k|x) = \xi^k|a_k)$$

and are complete in the sense that for all $|x) \in \mathcal{V}_d$

$$\sum_{k=1}^d \mathbb{P}_k|x) = |x) \quad ; \textit{i.e.}, \quad \sum_{k=1}^d \mathbb{P}_k = \mathbb{I}$$

We arrive thus at the spectral decomposition

$$\mathbb{A} = \sum_{k=1}^d a_k\mathbb{P}_k = \sum_{k=1}^d |a_k)\,a_k\,(a_k|$$

of the hermitian matrix $\mathbb{A}$.[1] More generally,

$$\mathbb{A}^n = \sum_{k=1}^d a_k^n\,\mathbb{P}_k = \sum_{k=1}^d |a_k)\,a_k^n\,(a_k|$$

and, for all $f(\bullet)$ that can be developed as formal power series,

$$f(\mathbb{A}) = \sum_{k=1}^d f(a_k)\,\mathbb{P}_k = \sum_{k=1}^d |a_k)\,f(a_k)\,(a_k|$$

---

[1] Degenerate spectra are described $\{(a_1, \mu_1), (a_2, \mu_2), \ldots, (a_\nu, \mu_\nu)\}$, where the a_k are distinct, μ_k is the degeneracy of a_k and $\sum_k \mu_k = d$. In such cases we have

$$\mathbb{A} = \sum_{k=1}^\nu a_k\mathbb{P}_k$$

where $\mathbb{P}_k^2 = \mathbb{P}_k$ projects onto a μ_k-dimensional subspace of $\mathcal{V}_d$: $\mathrm{tr}\,\mathbb{P}_k = \mu_k$. Those subspaces are mutually orthogonal. One can proceed arbitrarily to erect orthonormal bases on each of those subspaces; *i.e.*, to construcrt subresolutions $\mathbb{P}_k = \sum_{j=1}^{\mu_k} \mathbb{P}_{jk}$. The short of it: the issues posed by spectral degeneracy are, for present purposes, uninteresting, and will henceforth be ignored.

In particular, we have this spectral decomposition of the unitary matrix "generated" by the antihermitian matrix $i\mathbb{A}$:

$$\mathbb{U} \equiv e^{i\mathbb{A}} = \sum_{k=1}^{d} |a_k)\, e^{i\,a_k}\, (a_k|$$

I have belabored this familiar material in order to facilitate discussion of some closely related material which, because only rarely called upon in physical applications, is much less familiar.

Generalized spectral decomposition. Abandoning now our former hermiticity assumption, we assume $\mathbb{M}$ to be an arbitrary $d\times d$ complex matrix. We confront now a *pair* of "eigenproblems" which, because the eigenvalues of $\mathbb{M}^+$ are (by an easy argument) complex conjugates of the eigenvalues of $\mathbb{M}$, can be formulated

$$\begin{aligned} &\mathbb{M}\ \ |r_k) = m_k|r_k) \\ &\mathbb{M}^+|\ell_k) = m_k^*|\ell_k) \quad \iff \quad (\ell_k|\mathbb{M} = (\ell_k|m_k \end{aligned}$$

So we have, in general, to distinguish between *right eigenvectors* $\{|r_k)\}$ and *left eigenvectors* $\{(\ell_k|\}$. Though we are generally in position to say nothing about inner products of the forms $(r_j|r_k)$ or $(\ell_j|\ell_k)$, it follows from

$$(\ell_j|\mathbb{M}|r_k) = m_j(\ell_j|r_k) = m_k(\ell_j|r_k)$$

that

$$(\ell_j|r_k) = 0 \quad \text{if} \quad m_j \neq m_k$$

In the absence of spectral degeneracy

$$(\ell_j| \text{ is } \perp \text{ every } |r_k) \text{ except its "mate" } |r_j)\text{, and } \textit{vice versa}$$

The bases $\{|r_k)\}$ and $\{(\ell_k|\}$ are, in this sense, "biorthogonal."[2]

---

[2] This is a concept encountered in solid state physics and crystalography. Let $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ be an arbitrary set of linearly independent unit vectors in 3-space, and define

$$\mathbf{A} \equiv \mathbf{b}\times\mathbf{c}, \quad \mathbf{B} \equiv \mathbf{c}\times\mathbf{a}, \quad \mathbf{C} \equiv \mathbf{a}\times\mathbf{b}$$

Then

$$\begin{aligned} \mathbf{A} &\perp \mathbf{b} \,\&\, \mathbf{c} \\ \mathbf{B} &\perp \mathbf{c} \,\&\, \mathbf{a} \\ \mathbf{C} &\perp \mathbf{a} \,\&\, \mathbf{b} \end{aligned}$$

The elements $\{\mathbf{A}, \mathbf{B}, \mathbf{C}\}_{\text{normalized}}$ of the "reciprocal basis" are used to construct the "reciprocal lattice." For related material, see P. M. Morse & H. Feshbach, *Methods of Theoretical Physics* (1953), pages 884 & 931.

Introduce matrices

$$\mathbb{P}_k = \frac{|r_k)(\ell_k|}{(\ell_k|r_k)}$$

and notice that the definition gives back $\mathbb{P}_k = |r_k)(r_k|$ when the hermiticity assumption $\mathbb{M}^+ = \mathbb{M}$ is reinstated. Those matrices are projective

$$\mathbb{P}_k^2 = \frac{|r_k)(\ell_k|r_k)(\ell_k|}{(\ell_k|r_k)^2} = \mathbb{P}_k$$

and (by biorthogonality) orthogonal:

$$\mathbb{P}_j\mathbb{P}_k = \frac{|r_j)(\ell_j|r_k)(\ell_k|}{(\ell_j|r_j)(\ell_k|r_k)} = \mathbb{O} \quad : \quad j \neq k$$

To establish that they are also complete we note that if the orthonormal frames $\{|r_j)\}$ and $\{|\ell_k)\}$ were arbitrary the r-coordinates $(r_j|x)$ and the ℓ-coordinates $(\ell_k|x)$ of an arbitrary vector $|x)$ would stand in a relationship that reduces to simple proportionality in the presence of biorthogonality:

$$\begin{aligned}(\ell_j|x) &= \sum_k (\ell_j|r_k)(r_k|x)\\ &\downarrow\\ &= (\ell_j|r_j)(r_j|x) \quad \text{by biorthogonality}\end{aligned}$$

of which, by the way, $(r_j|\ell_j) = (\ell_j|r_j)^{-1}$ is a corollary. We therefore have

$$\begin{aligned}\Big\{\sum_j \mathbb{P}_j\Big\}|x) &= \sum_j |r_j)\frac{(\ell_j|x)}{(\ell_j|r_j)}\\ &= \sum_j |r_j)(r_j|x) = |x) \quad : \quad \text{all } |x)\\ &\Downarrow\\ \sum_j \mathbb{P}_j &= \mathbb{I}\end{aligned}$$

Similarly,

$$\begin{aligned}\mathbb{M} = \sum_j m_j|r_j)(r_j| &= \sum_j\sum_k m_j|r_j)(r_j|\ell_k)(\ell_k|\\ &= \sum_j m_j|r_j)(r_j|\ell_j)(\ell_j|\\ &= \sum_j m_j|r_j)(\ell_j|r_j)^{-1}(\ell_j|\\ &\Downarrow\\ \mathbb{M} &= \sum_j m_j\mathbb{P}_j\end{aligned}$$

From this "generalized spectral decomposition" we recover the "standard" spectral decomposition when, as a special circumstance, $\mathbb{M}$ is hermitian.

We are in position now to write

$$f(\mathbb{M}) = \sum_j f(m_j)\,\mathbb{P}_j$$

and, in particular, to speak of the "logarithm" of any non-singular matrix $\mathbb{M}$:

$$\mathbb{M} = e^{\mathbb{L}} \quad \text{with} \quad \mathbb{L} = \sum_j \log(m_j)\,\mathbb{P}_j$$

Application to proof of an elegant identity. Suppose

$$\mathbb{A} = e^{\mathbb{B}}$$

The (nameless!) elegant identity in question—which I first encountered in a paper by Schwinger, and of which I have made essential use many times in the course of my career, in a great variety of contexts—asserts that

$$\det\mathbb{A} = e^{\operatorname{tr}\mathbb{B}}$$

A simple proof is available when the square matrix $\mathbb{A}$ is non-singular and can be diagonalized by similarity transformation, for

$$\mathbb{A} = \mathbb{S}^{-1}\begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_n \end{pmatrix}\mathbb{S} \Longrightarrow \mathbb{B} = \mathbb{S}^{-1}\begin{pmatrix} \log a_1 & 0 & \dots & 0 \\ 0 & \log a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \log a_n \end{pmatrix}\mathbb{S}$$

and we have

$$\begin{aligned} \det\mathbb{A} &= \prod_k a_k \\ e^{\operatorname{tr}\mathbb{B}} &= \exp\Big\{\sum_k \log a_k\Big\} = \exp\Big\{\log\Big[\prod_k a_k\Big]\Big\} = \det\mathbb{A} \end{aligned}$$

Non-singularity was assumed to avoid reference to "log 0." Matrices encountered in physical contexts often conform to the conditions assumed in that argument.[3]

[3] Think of rotation matrices, which by $\mathbb{R}^{\mathsf{T}}\mathbb{R} = \mathbb{I}$ are proper/improper according as $\det\mathbb{R} = \pm 1$. Matrices of the form $\mathbb{R} = e^{\mathbb{A}}$ are rotational if and only if $\mathbb{A}$ is antisymmetric, therefore traceless. But $\operatorname{tr}\mathbb{A} = 0 \;\Rightarrow\; \det\mathbb{R} = 1$, from which we conclude that rotation matrices do not possess (real) antisymmetric "generators." Similarly, the unitarity of $\mathbb{U}$ implies $\det(\mathbb{U}^{+}\mathbb{U}) = |\det\mathbb{U}|^2 = 1$. But $\mathbb{U} = e^{\mathbb{G}}$ is unitary if and only if $\mathbb{G}$ is antihermitian, and the diagonal elements of antihermitian matrices are necessarily imaginary. So we have the sharpened statement $\det\mathbb{U} = e^{i\phi}$ with $i\phi = \operatorname{tr}\mathbb{G}$. Similar remarks arise when Lorentz matrices—defined $\mathbb{L}^{\mathsf{T}}\mathbb{G}\mathbb{L} = \mathbb{G}$ or equivalently $\mathbb{L}^{-1} = \mathbb{G}^{-1}\mathbb{L}^{\mathsf{T}}\mathbb{G}$, where $\mathbb{G}$ is the Lorentz metric—are written in "polar form" $\mathbb{L} = e^{\mathbb{A}}$.

But not every square matrix $\mathbb{M}$—more particularly, not every $\mathbb{M}$ encountered in the course of physical argument—admits of diagonalization by similarity transformation.[4] We are, however, in position now to observe that

$$\det\mathbb{M} = \prod_{k=1}^{d} m_k = \prod_{j=1}^{\nu}(m_j)^{\mu_j}$$

and to construct

$$\mathbb{L} = \log\mathbb{M} = \sum_{j=1}^{\nu}\log(m_j)\,\mathbb{P}_j$$

But as previously remarked, $\mathrm{tr}\,\mathbb{P}_j = \mu_j$, so

$$\mathrm{tr}\,\mathbb{L} = \sum_{j=1}^{\nu}\mu_j\log(m_j)$$

$$\Downarrow$$

$$\exp(\mathrm{tr}\,\mathbb{L}) = \prod_{j=1}^{\nu}(m_j)^{\mu_j}$$

and our "elegant identity" is established now—with remarkable simplicity—in its full generality.

By way of application, look to the improper rotation matrix

$$\mathbb{R} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\cdot\exp\left\{\begin{pmatrix} 0 & \phi \\ -\phi & 0 \end{pmatrix}\right\} = \begin{pmatrix} \cos\phi & \sin\phi \\ \sin\phi & -\cos\phi \end{pmatrix}$$

We have argued[3] that $\mathbb{R}$—since improper—cannot be developed $\mathbb{R} = e^{\mathbb{A}}$ with $\mathbb{A}$ real and antisymmetric. But $\mathbb{R}$ is non-singular (its eigenvalues are $\omega_1 = -1$ and $\omega_2 = +1$) so must admit of generalized spectral decomposition. By quick computation we obtain

$$\mathbb{R} = \omega_1\mathbb{P}_1 + \omega_2\mathbb{P}_2$$

with

$$\mathbb{P}_1 = \begin{pmatrix} \sin^2\tfrac{1}{2}\phi & -\cos\tfrac{1}{2}\phi\sin\tfrac{1}{2}\phi \\ -\cos\tfrac{1}{2}\phi\sin\tfrac{1}{2}\phi & \cos^2\tfrac{1}{2}\phi \end{pmatrix}$$

$$\mathbb{P}_2 = \begin{pmatrix} \cos^2\tfrac{1}{2}\phi & +\cos\tfrac{1}{2}\phi\sin\tfrac{1}{2}\phi \\ +\cos\tfrac{1}{2}\phi\sin\tfrac{1}{2}\phi & \sin^2\tfrac{1}{2}\phi \end{pmatrix}$$

---

[4] It is, in the present context, of no help to recall that *every* matrix—whether square *or rectangular*!—can by singular value decomposition (SVD) be displayed

$$\mathbb{M} = \mathbb{S}^{-1}\mathbb{D}\,\mathbb{T}\ \text{ with } \mathbb{D} \text{ diagonal}$$

since in the general case $\mathbb{S} \neq \mathbb{T}$, which subverts the logic of the simple proof.

whence

$$\log \mathbb{R} = \log(\omega_1)\,\mathbb{P}_1 + \log(\omega_2)\,\mathbb{P}_2 = i\pi\,\mathbb{P}_1$$

With *Mathematica*'s assistance we verify that indeed

$$\begin{pmatrix} \cos\phi & \sin\phi \\ \sin\phi & -\cos\phi \end{pmatrix} = \exp\left\{ i\pi \begin{pmatrix} \sin^2 \frac{1}{2}\phi & -\cos\frac{1}{2}\phi \sin\frac{1}{2}\phi \\ -\cos\frac{1}{2}\phi\sin\frac{1}{2}\phi & \cos^2\frac{1}{2}\phi \end{pmatrix} \right\}$$

The generator is now not real antisymmetric but *imaginary symmetric*, so the real rotation matrix $\mathbb{R}$ is actually *unitary*. Finally, we have

$$\det \mathbb{R} = e^{i\pi\,\mathrm{tr}\,\mathbb{P}_1} = e^{i\pi} = -1$$

Continuously interpolated classical/quantum Markov processes. Sequences

$$\{f_0, f_1, f_2, \ldots, f_n, \ldots\}$$

—and, more particularly, sequences generated by iterative processes

$$x \to f(x) \to f(f(x)) \to \cdots \to f(f(\cdots f(x)\cdots)) \to \cdots$$

—of inexhaustible variety are encountered in pure/applied mathematics. In a large subclass of those cases it is meaningful to ask "What meaning can be assigned to the objects that occupy the interstices?" Think, for example, of the meaning assigned by the gamma function to the symbol $x!$ when x is non-integral, or of the meaning assigned by the beta function to $\binom{x}{y}$ when x and y are non-integral, or of the meanings assigned by the fractional calculus to derivatives and integrals of non-integral order. I will be concerned here with certain linear algebraic iterative processes

$$\mathsf{x} \to \mathbb{F}\,\mathsf{x} \to \mathbb{F}^2\,\mathsf{x} \to \cdots \to \mathbb{F}^n\,\mathsf{x} \to \cdots$$

and with the question "What meaning can be assigned to $\mathbb{F}^n\,\mathsf{x}$ when n is non-integral?" The short answer: we are in position now to construct

$$\mathbb{F} = e^{\log \mathbb{F}}$$

and therefore to write

$$\mathbb{F}^\nu = e^{\nu \log \mathbb{F}} \quad : \quad \nu \text{ real or complex}$$

But the details relating to a specific application—classical/quantum random walks on graphs—are what interest me, and it is to those which I will restrict my remarks.

A "token" strolls randomly on a graph with N nodes (or "vertices"). The j^{th} element of the "stochastic vector"

$$\mathbf{p}_n = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_j \\ \vdots \\ p_N \end{pmatrix}_n$$

states the probability that, after n steps, the token finds itself standing on node $\#j$. "Markov processes" arise when the token is "memoryless"; *i.e.*, when the token's next step is determined probabilistically by is present location, without reference to past locations

$$\mathbf{p}_{n+1} = \mathbb{M}\,\mathbf{p}_n$$

and when

$$\mathbb{M} = \|M_{ij}\| \quad \text{with} \quad M_{ij} = \text{probability}_{i\leftarrow j}$$

remains constant throughout the process. A token standing on a given node must, after one step, find itself standing on *one or another* of the nodes, from which we conclude that each of the columns of $\mathbb{M}$ sums to unity (is, in others words, a stochastic vector): $\sum_i M_{ij} = 1$ (all j). The phrase "detailed balance" refers to situations in which

$$\text{probability}_{i\leftarrow j} = \text{probability}_{j\leftarrow i} \quad : \quad \text{all } i, j$$

In such cases the "Markov matrix" $\mathbb{M}$ is symmetric: $\mathbb{M}^{\mathsf{T}} = \mathbb{M}$.

Looking to general properties of the spectra of Markov matrices. . . the row vector $(1, 1, \ldots, 1)$ is a left eigenvector of every $\mathbb{M}$, with left eigenvalue $\lambda = 1$. But left eigenvalues are right eigenvalues, so the spectra of Markov matrices invariably contain $+1$ as an element. It is entirely possible for one or more columns of such a matrix to be identical; in such cases $\det \mathbb{M} = 0$, so we must distinguish between singular and non-singular Markov matrices: 0 appears one or more times in the spectra of singular Markov matrices. The eigenvalues of symmetric Markov matrices are invariably real, but in the absence of symmetry (*i.e.*, of detailed balance) they are typically complex (and because $\mathbb{M}$ is real they occur in conjugate pairs). Numerical experimentation suggests, and it is not difficult to prove,[5] that in all cases

$$|\lambda| \leq 1$$

I will call $\lambda_1 = 1$ the "leading eigenvector. For symmetric Markov matrices the spectrum can be displayed

$$\lambda_1 = 1 > \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_N \geq -1$$

while in the absence of symmetry it is more natural to write

$$\lambda_1 = 1 \geq |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_N| \geq 0$$

Of special interest in some respects are cases in which $\lambda_N = -1$. Look, for

---

[5] See, for example, /www.numbertheory.org/courses/MP274/markov.pdf.

example, to the Markov matrix

$$\mathbb{M}_{\text{square}} = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{pmatrix}$$

which refers to unbiased balanced random walks on the perimeter of a square, and has eigenvalues $\{\lambda_1 = 1, \lambda_2 = \lambda_3 = 0, \lambda_4 = -1\}$. This (non-invertible!) matrix has the curious property that

$$\mathbb{M}^n_{\text{square}} = \begin{cases} \begin{pmatrix} 0 & \bullet & 0 & \bullet \\ \bullet & 0 & \bullet & 0 \\ 0 & \bullet & 0 & \bullet \\ \bullet & 0 & \bullet & 0 \end{pmatrix} & n \text{ even} \\ \\ \begin{pmatrix} \bullet & 0 & \bullet & 0 \\ 0 & \bullet & 0 & \bullet \\ \bullet & 0 & \bullet & 0 \\ 0 & \bullet & 0 & \bullet \end{pmatrix} & n \text{ odd} \end{cases}$$

where $\bullet = \frac{1}{2}$. I interpolate here the observation that (pretty obviously)

Products of Markov matrices are Markovian

The matrix $\mathbb{M} = \mathbb{M}_1\mathbb{M}_2\cdots\mathbb{M}_p$ describes a Markov process in which the next-step instructions are adjusted repetitively/cyclically.

So much by way of preparation.

Standard/generalized spectral decomposition supples

$$\mathbb{M} = \mathbb{P}_1 + \lambda_2\mathbb{P}_2 + \lambda_3\mathbb{P}_3 + \cdots \lambda_N\mathbb{P}_N$$

which in non-singular cases can be used to construct

$$\mathbb{B} \equiv \log\mathbb{M} = \log(\lambda_2)\,\mathbb{P}_2 + \log(\lambda_3)\,\mathbb{P}_3 + \cdots + \log(\lambda_N)\,\mathbb{P}_N$$

and thus to assign meaning

$$\mathbb{M}^\nu = e^{\nu\mathbb{B}}$$

to the real integral/fractional powers of $\mathbb{M}$ (and even to complex powers, in which, however, I have no present interest). All of which is evidently known to *Mathematica*: when I construct the logarithm $\mathbb{B}$ of a random Markov matrix $\mathbb{M}$ I find that in all cases the commands **MatrixPower[M,** ν**]** and **MatrixExp[**ν **B]** give identical results.

Typically—in the absence of the restrictive assumptions that I will soon have occasion to install—some eigenvalues are complex, others are real but negative, and in all such cases $\log\lambda$ is complex. Moreover, the right/left eigenvectors $|r)$ and $(\ell|$—whence also the projector $\mathbb{P}$—associated with complex eigenvalues are complex. For those reasons we expect $\mathbb{B}$, in the absence of restrictive assumptions, to be complex, and the same to be true of $\mathbb{M}^\nu = e^{\nu\mathbb{B}}$. Numerical experimentation provides compelling evidence, however, that when ν is a positive integer all complexity magically disappears:

$$\mathbb{M}^\nu = e^{\nu\mathbb{B}} \quad \text{is real Markovian when } \nu = 1, 2, 3, \ldots$$

But when ν is fractional $\mathbb{M}^\nu$ is complex; its elements are no longer interpretable as transition probabilities, yet $\mathbb{M}^\nu$ remains "formally Markovian" in the sense that its columns remain "formally stochastic" (sum to $1 + 0i$).[6] Application of $\mathbb{M}^\nu$ to a stochastic vector p yields a formally stochastic complex vector p_ν.

Numerical experimentation indicates that in all cases the columns of $\mathbb{B}$ sum to zero, which is what one would expect on the grounds that when ν is infinitesimal

$$\mathbb{M}^\nu \approx \mathbb{I} + \nu\mathbb{B}$$

Since the columns of $\mathbb{I}$ sum to unity and $\mathbb{M}^\nu$ is formally Markovian, the columns of $\mathbb{B}$ must sum to $0 + 0i$.

To obtain "continuous Markov processes" we must arrange to avoid the intrusion of complex numbers: in short, we must look to the subclass of Markov matrices that have real non-negative eigenvalues. To achieve spectral reality we have only to assume detailed balancing; *i.e.*, to insist that the real matrix $\mathbb{M}$ be symmetric. Spectral positivity is, however, a bit more difficult to guarantee. It is simplest to attach our restrictive assumptions not to $\mathbb{M}$ but to its necessarily symmetric logarithm $\mathbb{B}$. If the automatically real eigenvalues of $\mathbb{B}$ are denoted $\{\beta_1, \beta_2, \ldots, \beta_N\} = \{\log\lambda_1, \log\lambda_2, \ldots, \log\lambda_N\}$ then the eigenvalues of the automatically symmetric matrix $\mathbb{M}$ become automatically positive:

$$\{\lambda_1, \lambda_2, \ldots, \lambda_N\} = \{e^{\beta_1}, e^{\beta_2}, \ldots, e^{\beta_N}\}$$

To insure that $\mathbb{M} = e^{\mathbb{B}}$ be Markovian we impose upon $\mathbb{B}$ the requirement that all rows/columns sum to zero. To construct such matrices we write

$$\mathbb{A} = \begin{pmatrix} 0 & a_{12} & a_{13} & a_{14} & \ldots & a_{1N} \\ 0 & 0 & a_{23} & a_{24} & \ldots & a_{2N} \\ 0 & 0 & 0 & a_{34} & \ldots & a_{3N} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 0 \end{pmatrix} + \text{transpose}$$

---

[6] Something quite similar can be said when $\nu = -1, -2, -3, \ldots$; the elements of $\mathbb{M}^\nu$ then tend to fall outside the allowed interval $[0, 1]$, but the elements of each column continue to sum to unity. The same can be said in particular of $\mathbb{M}^{-1}$: Markov processes are *non-invertible* (better: their inverses are non-Markovian).

$$\mathbb{D} = \begin{pmatrix} d_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & d_2 & 0 & 0 & \dots & 0 \\ 0 & 0 & d_3 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & d_N \end{pmatrix} \quad \text{with} \quad d_j = \sum_{k=1}^{N} A_{jk}$$

and from those matrices assemble

$$\mathbb{B} = \mathbb{A} - \mathbb{D}$$

Numerical experimentation establishes[7] that all such matrices are singular, and possess spectra of the form $\{\beta_1 = 0 \geq \beta_2 \geq \beta_3 \geq \cdots \geq \beta_N\}$. We establish similarly that $\mathbb{M} = e^{\mathbb{B}}$ is invariably Markovian, and that so more generally are all fractional powers of $\mathbb{M}$.

Feeding $\lambda_1 = e^0 = 1$ and $0 < \lambda_k = e^{\beta_k} < 1$ $(k = 2, 3, \dots, N)$ into the spectral decomposition of $\mathbb{M}$ we find that

$$\lim_{\nu \to \infty} \mathbb{M}^\nu = \mathbb{P}_1$$

where $\mathbb{P}_1$ projects onto the leading eigenvector $|r_1)$ of $\mathbb{M}$. By experimental evidence

$$|r_1) \sim \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbb{P}_1 = \tfrac{1}{N} \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix}$$

which establishes the sense in which such Markov processes *equilibrate*: taking p_0 to be an arbitrary initial stochastic vector, we have[8]

$$\mathsf{p}_0 \quad \longrightarrow \quad \mathsf{p}_\infty = \tfrac{1}{N} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

---

[7] Here as always—so limited are my present objectives—I omit any attempt to construct analytical demonstrations of claims the validity of which I am convinced on the basis of randomized numerical calculation.

[8] Markov matrices which, like the $\mathbb{M}_{\text{square}}$ considered previously, include -1 among their eigenvalues cannot be rendered continuous. They give rise asymptotically to states that "blink" alternately between

$$\begin{pmatrix} \bullet \\ 0 \\ \bullet \\ 0 \\ \vdots \\ \bullet \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ \bullet \\ 0 \\ \vdots \\ \bullet \\ 0 \\ \bullet \end{pmatrix} \quad \text{with} \quad \bullet = \tfrac{1}{N/2}$$

and give back the preceding p_∞ *when averaged*.

In the theory of continuous-time *quantum* walks on finite-dimensional state spaces (graphs)[9] the role of $\mathbb{M}^\nu$ is taken over by unitary matrices of the special design

$$\mathbb{U}(t) = e^{-i\,t\,\mathbb{B}}$$

where $\mathbb{B}$ retains the classical structure described above, but is thought of in this context as a real symmetric (therefore hermitian) "Hamiltonian." It serves the purposes most typical of workers in this field to assume that $\mathbb{A} = \mathbb{B} + \mathbb{D}$ has not the general structure described earlier, but is simply the "adjacency matrix" of the graph, which is the standard graph-theoretic device used to indicate which nodes (or "vertices") are linked together by "edges." Look, for example, to the adjacency matrix of a typical 5-node graph:

$$\mathbb{A} = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

which gives

$$\mathbb{B} = \begin{pmatrix} -2 & 0 & 1 & 1 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 & 0 \\ 1 & 1 & 0 & -3 & 1 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}$$

By (standard) spectral decomposition we have

$$\mathbb{B} = \omega_1\mathbb{P}_1 + \omega_2\mathbb{P}_2 + \omega_3\mathbb{P}_3 + \omega_4\mathbb{P}_4 + \omega_5\mathbb{P}_5$$

where invariably $\omega_1 = 0$ and $\mathbb{P}_1$ projects onto the ray defined by

$$|1) = \tfrac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Therefore $\mathbb{U}(t)$ is given in this instance by

$$\mathbb{U}(t) = \mathbb{P}_1 + e^{-i\,\omega_2 t}\,\mathbb{P}_2 + e^{-i\,\omega_3 t}\,\mathbb{P}_3 + e^{-i\,\omega_4 t}\,\mathbb{P}_4 + e^{-i\,\omega_5 t}\,\mathbb{P}_5$$

The complex matrix $\mathbb{U}(t)$ is "formally Markovian" in the sense that each of its rows/columns sum to complex unity. Unlike its classical counterpart, $\mathbb{U}(t)$ does

---

[9] See A. M Childs, E. Farhi & S. Gutmann, "Example of the difference between quantum and classical random walks," arXiv:quant-ph/0103020v1 (6 March 2001); D. Aharonov, A. Ambainis, J. Kempe & U. Vazirani, "Quantum walks on graphs," arXiv:quant-ph/0012090 (25 May 2002).

not equilibrate asymptotically, but instead forever buzzes (anharmonically unless the eigenvalues $\{\omega_2, \omega_3, \ldots, \omega_N\}$ are rational multiples of one another) about its steady component $\mathbb{P}_1$. A quantum allusion to the classical asymptote does, however, emerge in this time-averaged sense:

$$\lim_{T\to\infty} \frac{1}{T}\int_0^\infty \mathbb{U}(t)\,dt = \mathbb{P}_1$$

Hilbert-Schmidt orthogonalization. The space of $n \times n$ complex matrices, since closed under complex linear combination, is an n^2-dimensional vector space $\mathcal{V}_{n^2}$ which upon introduction of the Hilbert-Schmidt inner product[10]

$$(\mathbb{X}, \mathbb{Y}) \equiv \tfrac{1}{n}\,\mathrm{tr}(\mathbb{X}^+\,\mathbb{Y})$$

becomes an inner product space.

Let $\{\mathbb{X}_1, \mathbb{X}_2, \ldots, \mathbb{X}_{n^2}\}$ be an arbitrary set of linearly independent elements of $\mathcal{V}_{n^2}$, in short: a *basis* in $\mathcal{V}_{n^2}$. One then has

$$(\mathbb{X}_i, \mathbb{X}_j) = g_{ij}, \text{ elements of a hermitian “metric matrix” } \mathbb{G}$$

I describe a direct analog of the Gram-Schmidt orthogonalization process[11] that proceeds from the given basis $\{\mathbb{X}_1, \mathbb{X}_2, \ldots, \mathbb{X}_{n^2}\}$ to a basis $\{\mathbb{E}_1, \mathbb{E}_2, \ldots, \mathbb{E}_{n^2}\}$ the elements of which are Hilbert-Schmidt (or, as it has become my habit to say, “trace-wise”) orthonormal:

$$(\mathbb{E}_i, \mathbb{E}_j) = \delta_{ij}$$

By way of preparation, we observe that for all $\mathbb{Z}$ and all non-zero $\mathbb{A}$ one has the identity

$$\mathbb{Z} = \left\{\mathbb{Z} - \frac{(\mathbb{Z}, \mathbb{A})}{(\mathbb{A}, \mathbb{A})}\mathbb{A}\right\} + \left\{\frac{(\mathbb{Z}, \mathbb{A})}{(\mathbb{A}, \mathbb{A})}\mathbb{A}\right\} \equiv \mathbb{Z}_\perp + \mathbb{Z}_\parallel$$

where $\mathbb{Z}_\perp$ is orthogonal to $\mathbb{A}$

$$(\mathbb{Z}_\perp, \mathbb{A}) = 0$$

---

[10] In the following definition the prefactor $\frac{1}{n}$ has been introduced to achieve

$$(\mathbb{I},\, \mathbb{I}) = 1$$

One might, more generally, define

$$(\mathbb{X}, \mathbb{Y})_G \equiv \tfrac{1}{g}\,\mathrm{tr}(\mathbb{X}^+\mathbb{G}\,\mathbb{Y})$$

where $\mathbb{G}$ is an arbitrary non-singular $n \times n$ hermitian matrix and $g = \mathrm{tr}\,\mathbb{G}$.

[11] For a very clear account of this process—which can be traced back to Laplace and Cauchy—see http://en.wikipedia.org/wiki/Gram-Schmidt_process.

and

$$\mathbb{Z}_{\parallel} \equiv \frac{(\mathbb{Z},\mathbb{A})}{(\mathbb{A},\mathbb{A})}\,\mathbb{A} \equiv \text{proj}_{\mathbb{A}}(\mathbb{Z})$$

is the $\mathbb{A}$-component of $\mathbb{Z}$, "parallel" to $\mathbb{A}$. Finally, let us agree on occasion to write

$$\|\mathbb{A}\| \equiv \sqrt{(\mathbb{A},\mathbb{A})}$$

The orthogonalization procedure proceeds stepwise, and at each step one first orthogonalizes, then normalizes:

$$\begin{aligned}
\mathbb{X}_1 \to \mathbb{Y}_1 &= \mathbb{X}_1\\
&\downarrow\\
\mathbb{E}_1 &= \mathbb{Y}_1/\|\mathbb{Y}_1\|\\[1em]
\mathbb{X}_2 \to \mathbb{Y}_2 &= \mathbb{X}_2 - \text{proj}_{\mathbb{X}_1}(\mathbb{X}_2)\\
&\downarrow\\
\mathbb{E}_2 &= \mathbb{Y}_2/\|\mathbb{Y}_2\|\\[1em]
\mathbb{X}_3 \to \mathbb{Y}_3 &= \mathbb{X}_3 - \text{proj}_{\mathbb{X}_1}(\mathbb{X}_3) - \text{proj}_{\mathbb{X}_2}(\mathbb{X}_3)\\
&\downarrow\\
\mathbb{E}_3 &= \mathbb{Y}_3/\|\mathbb{Y}_3\|\\
&\vdots\\
\mathbb{X}_k \to \mathbb{Y}_k &= \mathbb{X}_k - \sum_{j=1}^{k-1}\text{proj}_{\mathbb{X}_j}(\mathbb{X}_k)\\
&\downarrow\\
\mathbb{E}_k &= \mathbb{Y}_k/\|\mathbb{Y}_k\| \quad : \quad k = 1,2,\ldots,n^2
\end{aligned}$$

The classic orthogonalization/orthonormalization procedure described above is recursive, but leads to a result that is easily expressed in non-recursive form. Immediately

$$\mathbb{Y}_2 = \frac{\begin{vmatrix}(\mathbb{X}_1,\mathbb{X}_1) & (\mathbb{X}_2,\mathbb{X}_1)\\ \mathbb{X}_1 & \mathbb{X}_2\end{vmatrix}}{(\mathbb{X}_1,\mathbb{X}_1)}$$

and by extension

$$\mathbb{Y}_k = \frac{\begin{vmatrix} (\mathbb{X}_1,\mathbb{X}_1) & (\mathbb{X}_2,\mathbb{X}_1) & \dots & (\mathbb{X}_k,\mathbb{X}_1) \\ (\mathbb{X}_1,\mathbb{X}_2) & (\mathbb{X}_2,\mathbb{X}_2) & \dots & (\mathbb{X}_k,\mathbb{X}_2) \\ \vdots & \vdots & \ddots & \vdots \\ (\mathbb{X}_1,\mathbb{X}_{k-1}) & (\mathbb{X}_2,\mathbb{X}_{k-1}) & \dots & (\mathbb{X}_k,\mathbb{X}_{k-1}) \\ \mathbb{X}_1 & \mathbb{X}_2 & \dots & \mathbb{X}_k \end{vmatrix}}{\begin{vmatrix} (\mathbb{X}_1,\mathbb{X}_1) & (\mathbb{X}_2,\mathbb{X}_1) & \dots & (\mathbb{X}_k,\mathbb{X}_1) \\ (\mathbb{X}_1,\mathbb{X}_2) & (\mathbb{X}_2,\mathbb{X}_2) & \dots & (\mathbb{X}_k,\mathbb{X}_2) \\ \vdots & \vdots & \ddots & \vdots \\ (\mathbb{X}_1,\mathbb{X}_{k-1}) & (\mathbb{X}_2,\mathbb{X}_{k-1}) & \dots & (\mathbb{X}_k,\mathbb{X}_{k-1}) \\ (\mathbb{X}_1,\mathbb{X}_k) & (\mathbb{X}_2,\mathbb{X}_k) & \dots & (\mathbb{X}_k,\mathbb{X}_k) \end{vmatrix}}$$

Let the denominator be denoted D_k. Then[11] the normalized matrices $\mathbb{E}_k$ acquire the non-recursive descriptions

$$\mathbb{E}_k = \frac{\text{same numerator}}{\sqrt{D_{k-1}D_k}}$$

where it is to be understood that $D_0 = 1$.

It is claimed[11] that the Gram-Schmidt orthogonalization procedure—and presumably also the Hilbert-Schmidt procedure—is numerically unstable, in the sense that orthogonality falls an easy victim to round-off errors. Refined iterative procedures have been devised to circumvent that problem. But my own numerical experimentation, in which I used randomly constructed $\mathbb{X}$ matrices of modest dimension, exposed no such problem.

Often it proves convenient to set $\mathbb{X}_1 = \mathbb{I}$; then $\mathbb{E}_1 = \mathbb{I}$ and the remaining $\mathbb{E}$-matrices are (by orthogonality) rendered traceless.

From the definition of the Hilbert-Schmidt inner product it follows readily that

$$(\mathbb{A},\mathbb{B}) = \begin{cases} (\mathbb{B},\mathbb{A})^* & : \quad \mathbb{A} \text{ and } \mathbb{B} \text{ arbitrary} \\ (\mathbb{B},\mathbb{A}) & : \quad \mathbb{A} \text{ and } \mathbb{B} \text{ hermitian} \end{cases}$$

so if the matrices $\{\mathbb{X}_1,\mathbb{X}_2,\dots,\mathbb{X}_{n^2}\}$ are hermitian then only real coefficients enter into the linear processes that send $\{\mathbb{X}_1,\mathbb{X}_2,\dots,\mathbb{X}_{n^2}\} \to \{\mathbb{E}_1,\mathbb{E}_2,\dots,\mathbb{E}_{n^2}\}$. Real linear combinations of hermitian matrices are hermitian, so we are led to the important conclusion that *the Hilbert-Schmidt process preserves hermiticity*.

Let $\{\mathbb{H}_1,\mathbb{H}_2,\dots,\mathbb{H}_{n^2}\}$ be a hermitian orthonormal basis in $\mathcal{V}_{n^2}$ and let $\mathbb{X}$ be an arbitrary element. One then has

$$\mathbb{X} = \sum_{k=1}^{n^2} x_k \mathbb{H}_k \quad \text{with} \quad x_k = (\mathbb{H}_k,\mathbb{X}) = \tfrac{1}{n}\,\text{tr}(\mathbb{H}_k\mathbb{X})$$

Suppose more particularly that $\mathbb{A}$ and $\mathbb{B}$ are hermitian. Then

$$\mathbb{AB} = \sum_{k=1}^{n^2} c_k \mathbb{H}_k \quad \text{with} \quad c_k = (\mathbb{H}_k,\mathbb{AB})$$

but products of hermitian matrices are typically *not* hermitian, which is to say: the coefficients c_k will typically not be real. Exceptions arise when (i) $\mathbb{A}$ and $\mathbb{B}$ *commute*[12] or (ii) when $\mathbb{AB}$ can be developed

$$\mathbb{AB} = \text{real linear combination of hermitian matrices}$$

We observe in this connection that we can (since $\{\mathbb{E}_1, \mathbb{E}_2, \ldots, \mathbb{E}_{n^2}\}$ is an orthonormal *basis* in $\mathcal{V}_{n^2}$) always write

$$\mathbb{E}_j\mathbb{E}_k = \sum_i c_j{}^i{}_k \mathbb{E}_i \quad \text{with} \quad c_j{}^i{}_k = (\mathbb{E}_i, \mathbb{E}_j\mathbb{E}_k)$$

whereupon the $\mathbb{E}$-matrices become elements of an *algebra*. As we have seen, only exceptionally in cases where the $\mathbb{E}$-matrices are hermitian can we expect their products to be also hermitian (all $c_j{}^i{}_k$ to be real). Look, for example, to the Pauli matrices

$$\boldsymbol{\sigma}_0 = \mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

which are seen to be hermitian and trace-wise orthonormal

$$(\boldsymbol{\sigma}_i, \boldsymbol{\sigma}_j) \equiv \tfrac{1}{2}\text{tr}\big(\boldsymbol{\sigma}_i\boldsymbol{\sigma}_j\big) = \delta_{ij}$$

and therefore to provide an orthonormal hermitian basis in $\mathcal{V}_4$.The set of Pauli matrices is algebraically closed

$$\boldsymbol{\sigma}_1^2 = \boldsymbol{\sigma}_2^2 = \boldsymbol{\sigma}_3^2 = \mathbf{I}$$

$$\begin{aligned} \boldsymbol{\sigma}_1\boldsymbol{\sigma}_2 &= i\boldsymbol{\sigma}_3 = -\boldsymbol{\sigma}_2\boldsymbol{\sigma}_1 \\ \boldsymbol{\sigma}_2\boldsymbol{\sigma}_3 &= i\boldsymbol{\sigma}_1 = -\boldsymbol{\sigma}_3\boldsymbol{\sigma}_2 \\ \boldsymbol{\sigma}_3\boldsymbol{\sigma}_1 &= i\boldsymbol{\sigma}_2 = -\boldsymbol{\sigma}_1\boldsymbol{\sigma}_3 \end{aligned}$$

but six of the sixteen possible products are seen to be *anti*-hermitian.

The seemingly weird placement of indices on the "structure constants" $c_j{}^i{}_k$ was intended to serve a purpose, which I digress now to explain. As an expression of the associativity of matrix multiplication

$$\mathbb{E}_i(\mathbb{E}_j\mathbb{E}_k) = (\mathbb{E}_i\mathbb{E}_j)\mathbb{E}_k$$

---

[12] The argument proceeds this way: we have

$$\begin{aligned} c_k = (\mathbb{H}_k, \mathbb{AB}) &= (\mathbb{AB},\ \mathbb{H}_k)^* \quad \text{in general} \\ &= (\mathbb{H}_k, \mathbb{AB})^* \quad \text{if } \mathbb{H}_k \text{ and } \mathbb{AB} \text{ are both hermitian} \\ &= c_k^* \end{aligned}$$

But if $\mathbb{A}$ and $\mathbb{B}$ are hermitian then $\mathbb{AB}$ is hermitian iff $\mathbb{A}$ and $\mathbb{B}$ commute.

we have

$$c_i{}^q{}_p\, c_j{}^p{}_k = c_i{}^p{}_j\, c_p{}^q{}_k$$

with $\sum_p$ understood. If we introduce matrices

$$\mathbb{C}_i \equiv \| c_i{}^q{}_p \|$$

the preceding equation can be formulated

$$\mathbb{C}_i\,\mathbb{C}_j = \textstyle\sum_p c_i{}^p{}_j\,\mathbb{C}_p \quad : \quad \text{compare} \quad \mathbb{E}_i\,\mathbb{E}_j = \textstyle\sum_p c_i{}^p{}_j\,\mathbb{E}_p$$

from which it becomes clear that the $\mathbb{C}$-matrices provide an $n^2 \times n^2$ matrix representation of the $\mathbb{E}$-algebra.

The Hilbert-Schmidt inner product is a standard mathematical device, and trace-wise orthonormality is a notion that I have encountered in a variety of contexts over the years (though in the past it has always arisen as a *discovered* property of matrices that recommended themselves to my attention for other reasons, not a property that I set out intentionally to construct). My interest in the construction of trace-wise orthonormal *hermitian* bases was sparked by the discovery that such things enter critically into the derivation and final statement of the Lindblad equation.[13]

Trace-wise orthonormal unitary bases. Hermiticity is defined by an additive condition $\mathbb{H}^+ - \mathbb{H} = \mathbb{O}$: (real) linear combinations of hermitian matrices are hermitian, and we are therefore not surprised by the discovery that Hilbert-Schmidt orthogonalization—an additive procedure—serves to construct trace-wise orthonormal hermitian bases $\{\mathbb{H}_1, \mathbb{H}_2, \dots, \mathbb{H}_{n^2}\}$ in $\mathcal{V}_{n^2}$. Unitarity, on the other hand, is defined by a *multiplicative* condition $\mathbb{U}^+\mathbb{U} = \mathbb{I}$: linear combinations of unitary matrices are, in general, *not* unitary. It is, therefore, somewhat surprising that it is nevertheless possible to construct complete sets of trace-wise orthonormal unitary matrices $\{\mathbb{U}_1, \mathbb{U}_2, \dots, \mathbb{U}_{n^2}\}$ in terms of which every matrix $\mathbb{X} \in \mathcal{V}_{n^2}$—and, more particularly, every unitary matrix in $\mathcal{V}_{n^2}$—can be developed

$$\mathbb{X} = x^k\,\mathbb{U}_k \quad : \quad \textstyle\sum_k \text{ understood}$$

where $(\mathbb{U}_j, \mathbb{U}_k) = \delta_{jk}$ entails $x^k = (\mathbb{U}_k, \mathbb{X})$.

It is clear that the construction principle must in this instance be not additive but multiplicative. One might contemplate starting from a (randomly?) prescribed set $\{\mathbb{W}_1, \mathbb{W}_2, \dots, \mathbb{W}_{n^2}\}$ of multiplicatively independent[14] unitary matrices and proceeding to "orthogonalize" them, but I have discovered—neither in the literature nor in the limited reach of my own imagination—no unitarity-preserving way to accomplish that objective. I am brought to the

[13] See Heinz-Peter Breuer & Francesco Petruccione, *The Theory of Open Quantum Systems* (2006), page 74 and §3.2.2, pages 115-120.

[14] . . . in the sense that none can be assembled multiplicatively from the others.

conclusion that orthogonality cannot be achieved "after the fact," but must be built in from the outset. And so it is in the schemes devised by Schwinger (1960) and by Werner (2000), of which I have written out detailed accounts in "Relationships among the unitary bases of Weyl, Schwinger, Werner and Oppenheim" (March 2012). Those schemes—which are, however different their superficial appearance, fundamentally identical—both proceed along lines first laid down by Weyl (1930), and both achieve their success by what (at least on first encounter) appears to be magical indirection. My objective here will be to try to remove some of the magic.

I begin with some trivial observations:

- All unitary matrices are automatically normalized:

$$(\mathbb{U},\mathbb{U}) = \tfrac{1}{n}\mathrm{tr}(\mathbb{U}^{+}\,\mathbb{U}) = \tfrac{1}{n}\mathrm{tr}(\mathbb{I}) = 1$$

- $\mathbb{I}$ is unitary, and matrices orthogonal to $\mathbb{I}$ are invariably traceless:

$$(\mathbb{I},\mathbb{U}) = 0 = \tfrac{1}{n}\mathrm{tr}(\mathbb{U})$$

Let the elements of the unitary basis be tentatively denoted

$$\{\mathbb{U}\} \equiv \{\mathbb{U}_0,\mathbb{U}_1,\mathbb{U}_2,\ldots,\mathbb{U}_{n^2-1}\}$$

where we have opted to include $\mathbb{I}$ as an element of the basis, writing $\mathbb{U}_0 = \mathbb{I}$. Let us agree, moreover, to write

$$\mathbb{U}_k^{+} = \mathbb{U}_{-k}$$

In that notation

$$(\mathbb{U}_i,\mathbb{U}_j) = \tfrac{1}{n}\mathrm{tr}(\mathbb{U}_{-i}\mathbb{U}_j)$$

Notice now that if the set $\{\mathbb{U}\}$ were *multiplicatively closed* in—tentatively—the weak sense that

$$\mathbb{U}_{-i}\mathbb{U}_j = \chi(i,j)\cdot\mathbb{U}_{k(i,j)} \;:\; \chi(i,i) = 1 \quad\text{and}\quad k(i,j)\begin{cases} = 0 \;:\; i=j \\ \neq 0 \;:\; \text{otherwise}\end{cases}$$

then <u>orthonormality would be automatic</u>:

$$(\mathbb{U}_i,\mathbb{U}_j) = \tfrac{1}{n}\chi(i,j)\,\mathrm{tr}(\mathbb{U}_{k(i,j)}) = \begin{cases} 1 \;:\; i=j \\ 0 \;:\; \text{otherwise}\end{cases}$$

The "multiplicative closure" notion must, however, be subjected to some refinement if it is to be rendered consistent with certain basic facts. First off, the forced unitarity of the expression on the right side of $\mathbb{U}_{-i}\mathbb{U}_j = \chi(i,j)\cdot\mathbb{U}_{k(i,j)}$ requires that in all cases the complex multiplier χ must have unit modulus:

$$\chi(i,j) = e^{i\,\varphi(i,j)}$$

From $(\mathbb{U}_{-i}\mathbb{U}_j)^{+} = \mathbb{U}_{-j}\mathbb{U}_i$ it follows moreover that

$$e^{-i\,\varphi(i,j)}\,\mathbb{U}_{-k(i,j)} = e^{i\,\varphi(j,i)}\,\mathbb{U}_{k(j,i)}$$

and therefore that both $\varphi(i,j)$ and $k(i,j)$ must be antisymmetric functions of their arguments

$$\varphi(i,j) = -\varphi(j,i) \qquad \text{and} \qquad k(i,j) = -k(j,i)$$

from which previous stipulations $\chi(i,i) = e^{i\,\varphi(i,i)} = e^0 = 1$ and $k(i,i) = 0$ now follow as consequences.

The simplest and most obvious way to comply with those conditions is to set $\varphi(i,j) = 0$ (all i,j) and $k(i,j) = j - i$, which is in effect to require of the elements of $\{\mathbb{U}\}$ that

$$\mathbb{U}_i\mathbb{U}_j = \mathbb{U}_{i+j}$$

And the simplest way to realize that condition is to assume the existence of a unitary matrix $\mathbb{V}$ such that

$$\mathbb{U}_k = \mathbb{V}^k \quad : \quad k = 0, 1, 2, \ldots, n^2 - 1$$

Since our objective is to span $\mathcal{V}_{n^2}$ we require that those powers of $\mathbb{V}$ be distinct, and to truncate the set of powers we impose the requirement that

$$\mathbb{V}^k = \mathbb{V}^{k+n^2}$$

Thus are we led to contemplate constructions of the form

$$\mathbb{U}_k = \mathbb{V}^{k \bmod n^2} \qquad \text{with} \qquad \mathbb{V}^p \begin{cases} = \mathbb{I} \ : \ p = 0 \\ \neq \mathbb{I} \ : \ p = 1, 2, \ldots, n^2 - 1 \\ = \mathbb{I} \ : \ p = n^2 \end{cases}$$

But it is not immediately immediately evident how an $n \times n$ matrix $\mathbb{V}$ with those properties is to be constructed, or even whether such a matrix can exist.[15]

This development motivates us to contemplate a fundamental revision of our program. We take the elements of the unitary basis $\{\mathbb{U}\}$ to be *doubly* indexed matrices of the form

$$\mathbb{U}_{ij} = \mathbb{V}^i\mathbb{W}^j \quad : \quad i, j \in \{0, 1, 2, \ldots, n-1\}$$

where the unitary "generators" $\mathbb{V}$ and $\mathbb{W}$ satisfy the cyclicity conditions

$$\mathbb{V}^n = \mathbb{I} \qquad \text{and} \qquad \mathbb{W}^n = \mathbb{I}$$

Normality $(\mathbb{U}_{ij}, \mathbb{U}_{ij}) = 1$ is again automatic, and if orthogonality

$$(\mathbb{U}_{ij}, \mathbb{U}_{kl}) = \delta_{ik}\delta_{jl} \equiv \delta_{ij,kl}$$

had been established it would again follow from $\mathbb{U}_{00} = \mathbb{I}$ that all the other $\mathbb{U}_{ij}$-matrices are traceless. But to achieve orthogonality we must work a bit.

---

[15] Every $n \times n$ matrix $\mathbb{V}$ is known by the Cayley-Hamilton theorem to satisfy a polynomial equation of degree n. How, therefore, can it come to satisfy a polynomial condition $\mathbb{V}^{n^2} = \mathbb{I}$ of degree n^2? This is a question to which I will return.

We have

$$(\mathbb{U}_{ij}, \mathbb{U}_{kl}) = \tfrac{1}{n}\,\mathrm{tr}(\mathbb{W}^{-j}\mathbb{V}^{-i}\mathbb{V}^{j}\mathbb{W}^{l}) = \tfrac{1}{n}\,\mathrm{tr}(\mathbb{W}^{-j}\mathbb{V}^{k-i}\mathbb{W}^{l})$$

but to progress beyond this point we need to be in position to bring the $\mathbb{V}^{k-i}$ factor to the left of the $\mathbb{W}^{-j}$ factor. We would like, more specifically, to be in position to write $\mathbb{W}^{-j}\mathbb{V}^{k-i} = e^{i\Omega}\,\mathbb{V}^{P}\,\mathbb{W}^{Q}$, where Ω, P and Q depend in presently unspecified ways upon i, j and k. This objective would be served most simply by positing a commutation rule of the form

$$\mathbb{W}\mathbb{V} = \chi{\cdot}\mathbb{V}\mathbb{W} \quad\Longrightarrow\quad \mathbb{W}^{p}\,\mathbb{V}^{q} = \chi^{pq}{\cdot}\,\mathbb{V}^{q}\,\mathbb{W}^{p} \quad : \quad \chi \equiv e^{i\varphi}$$

which, we notice, would have the attractive consequence of establishing *multiplicative closure* within the set of $\mathbb{U}_{ij}$-matrices:

$$\mathbb{U}_{ij}\mathbb{U}_{kl} = \mathbb{V}^{i}\mathbb{W}^{j}\mathbb{V}^{k}\mathbb{W}^{l} = \chi^{jk}\cdot\mathbb{V}^{i+k}\mathbb{W}^{j+l} = \chi^{jk}\cdot\mathbb{U}_{i+k,j+l}$$

It would then follow that

$$(\mathbb{U}_{ij}, \mathbb{U}_{kl}) = \chi^{-j(k-i)}\cdot\tfrac{1}{n}\,\mathrm{tr}\,\mathbb{U}_{k-i,l-j}$$

Looking to the traces of the left and right sides of $\mathbb{W}^{p}\mathbb{V}^{q} = \chi^{pq}{\cdot}\,\mathbb{V}^{q}\,\mathbb{W}^{p}$ we discover that

$$\mathrm{tr}\,\mathbb{U}_{pq} = 0 \quad\text{provided}\quad \chi^{pq} \neq 1,\ \textit{i.e.},\ \varphi\,pq \neq 0 \bmod 2\pi$$

while trivially

$$\mathrm{tr}\,\mathbb{U}_{00} = n$$

Deeper analysis is required, however, (i) to discover the trace implications of the statements $\mathbb{V}^{n} = \mathbb{W}^{n} = \mathbb{I}$ and (ii) to identify the conditions that must be imposed upon $\mathbb{V}$ and $\mathbb{W}$ to achieve $\mathrm{tr}\,\mathbb{V}^{k} = \mathrm{tr}\,\mathbb{W}^{k} = 0$ $(k = 1, 2, \ldots, n-1)$. We have at our disposal the following

> Little Theorem: Let $\mathbb{X}$ be an $n\times n$ matrix with $n \geq 2$. The characteristic polynomial of $\mathbb{X}$ has the form
>
> $$\begin{aligned} p(x) &= (-)^{n}\det(\mathbb{X} - x\,\mathbb{I}) \\ &= x^{n} - x^{n-1}(\mathrm{tr}\mathbb{X}) + \cdots + (-)^{n}\det\mathbb{X} \end{aligned}$$
>
> so $\mathrm{tr}\mathbb{X} = 0$ if and only if the characteristic polynomial presents no term of order x^{n-1}.

but confront the fact that the Cayley-Hamilton theorem cannot be used "in reverse": if , for example, it were known that the characteristic polynomial of $\mathbb{X}$ were $p(x) = x^{n} - 1$ then we would assuredly have $p(\mathbb{X}) = \mathbb{X}^{n} - \mathbb{I} = \mathbb{O}$, but from $\mathbb{X}^n - \mathbb{I} = \mathbb{O}$ it does not follow that the characteristic polynomial of $\mathbb{X}$ reads $p(x) = x^n - 1$, as I demonstrate. Look, for example, to

$$\mathbb{X} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad : \quad \text{satisfies } \mathbb{X}^4 - \mathbb{I} = \mathbb{O}$$

of which the characteristic polynomial is

$$\begin{aligned} p(x) &= x^4 - 2x^2 + 1 \\ &= (x^2 - 1)^2 \\ &= (x^4 - 1) - 2(x^2 - 1) \\ &\neq x^4 - 1 \end{aligned}$$

The obvious tracelessness of $\mathbb{X}$ conforms to the fact that $p(x)$ presents no term of order x^3. The matrix $\mathbb{X}$ does in fact satisfy $p(\mathbb{X}) = \mathbb{O}$, but it is seen to be a root also of the *reduced* characteristic polynomial $r(x) = x^2 - 1$, and it is by implication of the latter circumstance that it comes also to satisfy $\mathbb{X}^4 - \mathbb{I} = \mathbb{O}$. Turning our attention now to the powers of $\mathbb{X}$, we find[16]

$$\begin{aligned} p(\mathbb{X}^2, x) &= x^4 - 4x^3 + 6x^2 - 4x + 1 = (x-1)^4 \\ r(\mathbb{X}^2, x) &= x - 1 \end{aligned}$$

$$\begin{aligned} p(\mathbb{X}^3, x) &= x^4 - 2x^2 + 1 = (x^2 - 1)^2 \\ r(\mathbb{X}^3, x) &= x^2 - 1 \end{aligned}$$

from which we conclude on the basis of the LITTLE THEOREM that $\mathrm{tr}\,\mathbb{X}^2 = 4$, $\mathrm{tr}\,\mathbb{X}^3 = 0$. The reduced polynomials supply $(\mathbb{X}^2)^1 - \mathbb{I} = \mathbb{O}$ and $(\mathbb{X}^3)^2 - \mathbb{I} = \mathbb{O}$ from which we extract $\mathbb{X}^0 = \mathbb{X}^2 = \mathbb{X}^4 = \mathbb{I}$ and $\mathbb{X}^1 = \mathbb{X}^3 \neq \mathbb{I}$, whence

$$\begin{aligned} \mathrm{tr}\,\mathbb{X}^0 = \mathrm{tr}\,\mathbb{X}^2 = \mathrm{tr}\,\mathbb{X}^4 &= 4 \\ \mathrm{tr}\,\mathbb{X}^1 = \mathrm{tr}\,\mathbb{X}^3 &= 0 \end{aligned}$$

These results are further illuminated when one looks to the spectral properties of $\mathbb{X}$. Quite generally, if

$$\mathbb{X} \text{ has eigenvalues } \{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$$

then
$$\mathbb{X}^k \text{ has eigenvalues } \{\lambda_1^k, \lambda_2^k, \lambda_3^k, \lambda_4^k\} \quad : \quad k = 2, 3, \ldots$$

and
$$\mathrm{tr}\,\mathbb{X}^k = \lambda_1^k + \lambda_2^k + \lambda_3^k + \lambda_4^k$$

[16] I find it convenient in this expanded context to borrow a notational convention from *Mathematica*.

In the present instance

$$\begin{array}{lcl} \mathbb{X} \text{ has eigenvalues } \{+1,+1,-1,-1\} & \Longrightarrow & \mathrm{tr}\,\mathbb{X}\ = 0 \\ \mathbb{X}^2 \text{ has eigenvalues } \{+1,+1,+1,+1\} & \Longrightarrow & \mathrm{tr}\,\mathbb{X}^2 = 4 \\ \mathbb{X}^3 \text{ has eigenvalues } \{+1,+1,-1,-1\} & \Longrightarrow & \mathrm{tr}\,\mathbb{X}^3 = 0 \\ \mathbb{X}^4 \text{ has eigenvalues } \{+1,+1,+1,+1\} & \Longrightarrow & \mathrm{tr}\,\mathbb{X}^4 = 4 \end{array}$$

Our assignment now is to *exhibit* unitary $n \times n$ matrices $\mathbb{V}$ and $\mathbb{W}$ with the properties sufficient to ensure the trace-wise orthonormality of the unitary matrices $\mathbb{U}_{ij} = \mathbb{V}^i\mathbb{W}^j$. Schwinger and (in effect) Werner were content to pluck such matrices from their respective hats. I will attempt to proceed by a series of motivated steps along a path that is by now almost obvious. Taking clue from the cyclic implications of the equations $\mathbb{V}^n = \mathbb{I}$, $\mathbb{W}^n = \mathbb{I}$ and from the familiar fact that cycles loom large in the theory of permutation groups, we look to...

Some properties of permutation matrices and their characteristic polynomials. The matrix $\mathbb{P}$ that sends the top row (displayed as a column vector) of the permutation symbol

$$\mathcal{P} = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ i_1 & i_2 & i_3 & \cdots & i_n \end{pmatrix}$$

to the second row (similarly displayed) is a "permutation matrix." Such matrices show a solitary 1 in every row/column, with all other elements 0. For example, as an expression of

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 4 & 6 & 1 & 5 & 8 & 3 & 7 \end{pmatrix}$$

we have

$$\begin{pmatrix} 2\\4\\6\\1\\5\\8\\3\\7 \end{pmatrix} = \mathbb{P} \begin{pmatrix} 1\\2\\3\\4\\5\\6\\7\\8 \end{pmatrix} \quad \text{with} \quad \mathbb{P} = \begin{pmatrix} 0 & \bullet & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \bullet & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \bullet & 0 & 0 \\ \bullet & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \bullet & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \bullet \\ 0 & 0 & \bullet & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \bullet & 0 \end{pmatrix}$$

where for visual emphasis I have adopted the convention $\bullet = 1$.

It is clear by inspection that invariably $\mathbb{P}^{-1} = \mathbb{P}^{\mathsf{T}}$, which is to say: every permutation matrix is a *rotation* matrix (a real *unitary* matrix), proper or improper according as the permutation to which it refers is even or odd.

Every permutation can be resolved into disjoint "cycles." In our example, we have[17]

$$1 \to 2 \to 4 \to 1 \quad : \quad 3 \to 6 \to 8 \to 7 \to 3 \quad : \quad 5 \to 5$$

To emphasize its cycle structure we adjust the order of the terms in the original permutation symbol, writing

$$\left(\begin{array}{ccc|cccc|c} 1 & 2 & 4 & 3 & 6 & 8 & 7 & 5 \\ 2 & 4 & 1 & 6 & 8 & 7 & 3 & 5 \end{array}\right)$$

which after some simple relabeling (which itself is a kind of permutation) becomes

$$\left(\begin{array}{ccc|cccc|c} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 3 & 1 & 5 & 6 & 7 & 4 & 8 \end{array}\right)$$

and causes the associated permutation matrix to asssume the structure

$$\begin{pmatrix} 0 & \bullet & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \bullet & 0 & 0 & 0 & 0 & 0 \\ \bullet & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \bullet & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \bullet & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \bullet & 0 \\ 0 & 0 & 0 & \bullet & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \bullet \end{pmatrix} = \mathcal{B}\{\mathbb{C}_3, \mathbb{C}_4, \mathbb{C}_1\}$$

where

$$\mathcal{B}\{\mathbb{C}_3, \mathbb{C}_4, \mathbb{C}_1\} \text{ denotes the } \textit{block diagonal} \text{ matrix } \begin{pmatrix} \mathbb{C}_3 & 0 & 0 \\ 0 & \mathbb{C}_4 & 0 \\ 0 & 0 & \mathbb{C}_1 \end{pmatrix}$$

and in this instance

$$\mathbb{C}_3 = \begin{pmatrix} 0 & \bullet & 0 \\ 0 & 0 & \bullet \\ \bullet & 0 & 0 \end{pmatrix}, \quad \mathbb{C}_4 = \begin{pmatrix} 0 & \bullet & 0 & 0 \\ 0 & 0 & \bullet & 0 \\ 0 & 0 & 0 & \bullet \\ \bullet & 0 & 0 & 0 \end{pmatrix}, \quad \mathbb{C}_1 = (\,\bullet\,)$$

$\mathbb{C}_3$ refers to a permutation $1 \to 2 \to 3 \to 1$ that repeats with period $3 = \dim \mathbb{C}_3$, from which it follows that

$$\mathbb{C}_3^{\,3} = \mathbb{I}_3\,; \text{ similarly } \mathbb{C}_4^{\,4} = \mathbb{I}_4 \text{ and } \mathbb{C}_1^{\,1} = \mathbb{I}_1$$

---

[17] To the command **ToCycles**[{2, 4, 6, 1, 5, 8, 3, 7}] *Mathematica* responds

$$\{\{2, 4, 1\}, \{6, 8, 7, 3\}, \{5\}\}$$

where $\mathbb{I}_d$ refers to the d-dimensional identity matrix. We agree henceforth to reserve the term "cycle" (or "cyclic") for permutations $\mathcal{C}$ that cannot be resolved into subcycles, and the notation $\mathbb{C}$ for the matrix representations of such permutations—matrices which can by simple relabeling be brought to the "canonical form"

$$\mathbb{C}_n = \begin{pmatrix} 0 & \bullet & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \bullet & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \bullet & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \bullet & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \bullet \\ \bullet & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

The p^{th} power $(p = 2, 3, \ldots, n-1)$ of $\mathbb{C}_n$ is cyclic iff $\{p, n\}$ are relatively prime (*i.e.*, $\text{GCD}[p, n] = 1$), which will invariably be the case if n is prime. But more generally—typically but not invariably—powers of cyclic matrices are *not* cyclic:

$$\mathbb{C}_n{}^p \hookrightarrow \mathcal{B}\{g \text{ copies of } \mathbb{C}_{n/g}\} \quad : \quad g = \text{GCD}[p, n]$$

where $\hookrightarrow$ signifies "can by a relabeling be rendered." Thus[18]

$$\mathbb{C}_6{}^2 = \begin{pmatrix} 0 & 0 & \bullet & 0 & 0 & 0 \\ 0 & 0 & 0 & \bullet & 0 & 0 \\ 0 & 0 & 0 & 0 & \bullet & 0 \\ 0 & 0 & 0 & 0 & 0 & \bullet \\ \bullet & 0 & 0 & 0 & 0 & 0 \\ 0 & \bullet & 0 & 0 & 0 & 0 \end{pmatrix} \hookrightarrow \mathcal{B}\{\mathbb{C}_3, \mathbb{C}_3\}$$

$$\begin{aligned} \mathbb{C}_6{}^3 &\hookrightarrow \mathcal{B}\{\mathbb{C}_2, \mathbb{C}_2, \mathbb{C}_2\} \\ \mathbb{C}_6{}^4 &\hookrightarrow \mathcal{B}\{\mathbb{C}_3, \mathbb{C}_3\} \\ \mathbb{C}_6{}^5 &\hookrightarrow \mathcal{B}\{\mathbb{C}_6\} \\ \mathbb{C}_6{}^6 &\hookrightarrow \mathcal{B}\{\mathbb{C}_1, \mathbb{C}_1, \mathbb{C}_1, \mathbb{C}_1, \mathbb{C}_1, \mathbb{C}_1\} = \mathbb{I}_6 \end{aligned}$$

If a general/unspecialized permutation matrix is rendered

$$\mathbb{P} \hookrightarrow \mathcal{B}\{\mathbb{C}_a, \mathbb{C}_b, \ldots, \mathbb{C}_z\} \quad : \quad a + b + \cdots + z = n$$

then

$$\mathbb{P}^p \hookrightarrow \mathcal{B}\{\mathbb{C}_a{}^p, \mathbb{C}_b{}^p, \ldots, \mathbb{C}_z{}^p\}$$

where the matrices on the diagonal are subject to the resolution principle just described.

---

[18] To conduct experiments in this connection, command

RotateLeft$[\{1, 2, 3, \ldots, n\}, p]$
ToCycles$[\%]$
GCD$[p, n]$

All permutation matrices—whether cyclic or not—are periodic. The period of $\mathcal{B}\{\mathbb{C}_a, \mathbb{C}_b, \ldots, \mathbb{C}_z\}$ is

$$\pi\big(\mathcal{B}\{\mathbb{C}_a, \mathbb{C}_b, \ldots, \mathbb{C}_z\}\big) = \text{LCM}(a, b, \ldots, z)$$

Thus (returning to the example introduced on page 24) the 8×8 permutation matrix $\mathbb{P} = \mathcal{B}\{\mathbb{C}_3, \mathbb{C}_4, \mathbb{C}_1\}$ has period $\text{LCM}(3, 4, 1) = 12 > 8$.

All of which constitutes simply a permutationally particularized instance of the more general ideas/results sketched at the end of the preceding section. But the particularization permits a sharpeining of some of those results. To wit:

The (monic) characteristic polynomial of every $\mathbb{C}$-matrix is irreducible, and possesses the cyclotomic structure

$$r(\mathbb{C}_n, x) = p(\mathbb{C}_n, x) = x^n - 1$$

from which it follows by our Little Theorem that—which is clear already by inspection—every such matrix is traceless:

$$\text{tr}\,\mathbb{C}_n = 0 \quad : \quad n = 2, 3, \ldots$$

Immediately

$$\text{tr}\,\mathcal{B}\{\mathbb{C}_a, \mathbb{C}_b, \ldots, \mathbb{C}_z\} = \text{number of unit subscripts}$$

and

$$\text{tr}\,(\mathbb{C}_n{}^p) = \text{tr}\,\mathcal{B}\{g \text{ copies of } \mathbb{C}_{n/g}\} = \begin{cases} 0 & : \quad n/g \neq 1 \\ n & : \quad n/g = 1 \end{cases}$$

which is to say

$$= \begin{cases} 0 & : \quad p \bmod n \neq 0 \\ n & : \quad p \bmod n = 0 \end{cases}$$

There are, in this connection, lessons yet to be learned from the example cited for a third time at the top of this page:

$$\mathbb{P} = \begin{pmatrix} 0 & \bullet & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \bullet & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \bullet & 0 & 0 \\ \bullet & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \bullet & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \bullet \\ 0 & 0 & \bullet & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \bullet & 0 \end{pmatrix} \hookrightarrow \mathcal{B}\{\mathbb{C}_3, \mathbb{C}_4, \mathbb{C}_1\}$$

We have

$$\begin{aligned} r(\mathbb{P}, x) = p(\mathbb{P}, x) &= (x^3 - 1)(x^4 - 1)(x - 1) \\ &= x^8 - x^7 - x^5 + x^3 + x - 1 \end{aligned}$$

so by the Cayley-Hamilton theorem expect to have (and verify by computation that we do in fact have)

$$\mathbb{P}^8 - \mathbb{P}^7 - \mathbb{P}^5 + \mathbb{P}^3 + \mathbb{P}^1 - \mathbb{P}^0 = \mathbb{O} \tag{ii}$$

On the other hand, we have argued that $\pi(\mathbb{P}) = 12$ and verify by computation that indeed

$$\mathbb{P}^{12} - \mathbb{I} = \mathbb{O} \tag{iii}$$

I discuss now how it comes about that

$$(ii) \Longrightarrow (iii), \text{ but not conversely}$$

The roots of the cyclotomic polynomial $x^n - 1$ are n^{th} roots of unity:

$$x^n - 1 = \prod_{k=0}^{n-1}(x - \omega^k) \quad : \quad \omega = e^{i2\pi/n}$$

Noting once again that $\text{LCM}(3,4,1) = 12$, we observe that the factored form of (i) presents some but not all of the 12^{th} roots of unity:

$$\begin{aligned}
&(x^3-1)(x^4-1)(x-1)\\
&\quad = [(x-\Omega^0)(x-\Omega^4)(x-\Omega^8)]\\
&\qquad \cdot [(x-\Omega^0)(x-\Omega^3)(x-\Omega^6)(x-\Omega^9)] \cdot [(x-\Omega^0)] \quad : \quad \Omega = e^{i2\pi/12}
\end{aligned}$$

The missing roots are $\{\Omega^1, \Omega^2, \Omega^5, \Omega^7, \Omega^{10}, \Omega^{11}\}$. Clearly

$$\begin{aligned}
(x^3-1)(x^4-1)(x-1) &= 0\\
&\Downarrow\\
(x-\Omega^1)(x-\Omega^2)(x-\Omega^5)(x-\Omega^7)(x-\Omega^{10})(x-\Omega^{11})&\\
\cdot\, (x^3-1)(x^4-1)(x-1) &= (x^{12}-1)\\
&= 0
\end{aligned}$$

It is for this reason that

$$(\mathbb{P}^3 - \mathbb{I})(\mathbb{P}^4 - \mathbb{I})(\mathbb{P} - \mathbb{I}) = \mathbb{O} \quad \Longrightarrow \quad \mathbb{P}^{12} - \mathbb{I} = \mathbb{O}$$

A more heroic illustration of the same point follows from a fact that of all the 490 partitions of 19 the one with greatest LCM is[19]

$$19 = 3 + 4 + 5 + 7 \quad : \quad \text{LCM}(3,4,5,7) = 420$$

So if we construct the 19×19 block matrix (composite permutation matrix)

$$\mathbb{P}_{19} = \mathcal{B}(\mathbb{C}_3, \mathbb{C}_4, \mathbb{C}_5, \mathbb{C}_7)$$

we expect to have (and verify by calculation that we do in fact have)

$$\begin{aligned}
p(\mathbb{P}_{19}, x) &= (x^3-1)(x^4-1)(x^5-1)(x^7-1)\\
&\Downarrow\\
(\mathbb{P}_{19}^3 - \mathbb{I})(\mathbb{P}_{19}^4 - \mathbb{I})(\mathbb{P}_{19}^5 - \mathbb{I})(\mathbb{P}_{19}^7 - \mathbb{I}) &= \mathbb{O}\\
&\Downarrow\\
(\mathbb{P}_{19}^{420} - \mathbb{I}) &= \mathbb{O} \quad \Longrightarrow \quad \pi(\mathbb{P}_{19}) = 420
\end{aligned}$$

[19] I borrow this information from "The maximal LCM problem" (April 2012).

Looking now to the traces of cyclic matrices, and of permutation matrices assembled from them, and of the powers of such matrices. . . we have

$$\mathrm{tr}\,\mathbb{C}_n = \sum_{k=0}^{n-1}(\omega_n)^k = 0 \quad : \quad \omega_n = e^{i2\pi/n}$$

This elementary property of the roots of unity—made no less striking by its familiarity—becomes obvious when (as is standard practice in textbooks) the roots are displayed as symmetrically arranged points (unit vectors) on the complex plane, but can be attributed also—by appeal to our LITTLE THEOREM—to the fact that $p(\mathbb{C}_n, x) = x^n - 1$ $(n > 1)$ presents no term of order x^{n-1}. Those same arguments supply

$$\begin{aligned}\mathrm{tr}\,\mathbb{C}_n^p &= \begin{cases} n & : \quad p = 0 \\ 0 & : \quad p = 1, 2, \dots, n-1 \end{cases} \\ &= n\,\delta_{n,\,p \bmod n}\end{aligned}$$

Looking more generally to

$$\mathbb{P} = \mathcal{B}\{\mathbb{C}_a, \mathbb{C}_b, \dots, \mathbb{C}_z\}$$

we have

$$\begin{aligned}\mathrm{tr}\,\mathbb{P}^p &= \mathrm{tr}\,\mathbb{C}_a^p + \mathrm{tr}\,\mathbb{C}_b^p + \cdots + \mathrm{tr}\,\mathbb{C}_z^p \\ &= a\,\delta_{a,\,p \bmod a} + b\,\delta_{b,\,p \bmod b} + \cdots + z\,\delta_{z,\,p \bmod z} \\ &= \sum \text{dimensions of } \mathbb{C}\text{-factors with completed periods}\end{aligned}$$

In the following table I show how this works in the case $\mathbb{P} = \mathcal{B}\{\mathbb{C}_2, \mathbb{C}_3, \mathbb{C}_4\}$, which is 9×9 and has period $\mathrm{LCM}(2, 3, 4) = 12$:

p	$\mathrm{tr}\,\mathbb{C}_2^p$	$\mathrm{tr}\,\mathbb{C}_3^p$	$\mathrm{tr}\,\mathbb{C}_4^p$	$\mathrm{tr}\,\mathbb{P}^p$
1	0	0	0	0
2	2	0	0	2
3	0	3	0	3
4	2	0	4	6
5	0	0	0	0
6	2	3	0	5
7	0	0	0	0
8	2	0	4	6
9	0	3	0	3
10	2	0	0	2
11	0	0	0	0
12	2	3	4	9

We saw a simpler example of the same phenomenon already on page 22. It is a phenomenon of musical simplicity, familiar in every clockshop, or to everyone (of whom there cannot be many!) who has attended a performance of Steve Reich's "Clapping Music"[20] (1972): each $\mathbb{C}$-matrix is in effect a "drummer" ("clock,' "clapper"), who drums (ticks, claps) with his own characteristic frequency. From time to time the drum beats of this or that pair of drummers coincide, and less frequently (but periodically) they all coincide.

To expose the interrelationships among drummers—stripped of everything having to do explicitly with "permutations"—we deploy the eigenvalues of $\mathbb{C}_n$ on the principal diagonal of an $n \times n$ diagonal matrix $\mathbb{D}_n$:

$$\mathbb{C}_n \equiv \begin{pmatrix} 0 & \bullet & 0 & 0 & \cdots & 0 \\ 0 & 0 & \bullet & 0 & \cdots & 0 \\ 0 & 0 & 0 & \bullet & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \bullet \\ \bullet & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$\Updownarrow$$

$$\mathbb{D}_n \equiv \begin{pmatrix} \omega^0 & 0 & 0 & \cdots & 0 \\ 0 & \omega^1 & 0 & \cdots & 0 \\ 0 & 0 & \omega^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \omega^{n-1} \end{pmatrix} \quad : \quad \omega = e^{i\,2\pi/n}$$

The matrices $\mathbb{C}_n$ and $\mathbb{D}_n$ have—by design—identical spectra, therefore have identical characteristic polynomials, identical algebraic properties, identical traces. The same can be said of the matrices $\mathbb{P}$ and $\mathbb{Q}$ that result from the association

$$\begin{gathered} \mathbb{P} = \mathcal{B}(\mathbb{C}_1, \mathbb{C}_2, \ldots, \mathbb{C}_\nu) \\ \Updownarrow \\ \mathbb{Q} = \mathcal{B}(\mathbb{D}_1, \mathbb{D}_2, \ldots, \mathbb{D}_\nu) \end{gathered}$$

We observe that the matrices $\mathbb{D}_n$—whence also matrices of type $\mathbb{Q}$—are manifestly unitary.

Multiplication—whether from right or left—of a $\mathbb{C}$-matrix by any diagonal matrix serves simply to "decorate the ones" and to leave all zeros in place, so the trace properties of $\mathbb{C}_n\mathbb{D}_m$ and $\mathbb{D}_m\mathbb{C}_n$ are identical to those of $\mathbb{C}_n$. Comparing

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \omega^0 & 0 & 0 & 0 \\ 0 & \omega^1 & 0 & 0 \\ 0 & 0 & \omega^2 & 0 \\ 0 & 0 & 0 & \omega^3 \end{pmatrix} = \begin{pmatrix} 0 & \omega^1 & 0 & 0 \\ 0 & 0 & \omega^2 & 0 \\ 0 & 0 & 0 & \omega^3 \\ \omega^0 & 0 & 0 & 0 \end{pmatrix}$$

[20] See (!) the video at http://www.youtube.com/watch?v=dXhBti6256_s.

with

$$\omega\begin{pmatrix}\omega^0 & 0 & 0 & 0\\ 0 & \omega^1 & 0 & 0\\ 0 & 0 & \omega^2 & 0\\ 0 & 0 & 0 & \omega^3\end{pmatrix}\begin{pmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 1 & 0 & 0 & 0\end{pmatrix} = \omega\begin{pmatrix}0 & \omega^0 & 0 & 0\\ 0 & 0 & \omega^1 & 0\\ 0 & 0 & 0 & \omega^2\\ \omega^3 & 0 & 0 & 0\end{pmatrix}$$

we infer that

$$\mathbb{C}_n\mathbb{D}_n = \omega\cdot\mathbb{D}_n\mathbb{C}_n \quad : \quad \text{all } n>1$$

from which it follows that

$$\mathbb{C}_n^a\,\mathbb{D}_n^b = \omega^{ab \bmod n}\cdot\mathbb{D}_n^b\,\mathbb{C}_n^a$$

where $\omega = e^{i2\pi/n}$ and where the matrix products on both left and right are clearly unitary.

Schwinger's construction. Our attempt to solve the "orthonormal unitary basis problem" had, by page 21, brought us to constructions of the form $\mathbb{U}_{ij} = \mathbb{V}^i\mathbb{W}^j$ where unitary generators $\mathbb{V}$ and $\mathbb{W}$ were required to satisfy the cyclicity conditions $\mathbb{V}^n = \mathbb{W}^n = \mathbb{I}$, to have traceless non-trivial powers, and to satisfy a commutation relation of the form $\mathbb{W}\,\mathbb{V} = e^{i\varphi}\cdot\mathbb{V}\,\mathbb{W}$. Drawing motivation from he first of those conditions, we reviewed aspects of the theory of permutation matrices and were led (in the n-dimensional case) to matrices $\mathbb{C}_n$ and $\mathbb{D}_n$ that satisfy *all* of the required conditions... and thus to Schwinger's construction

$$\mathbb{U}_{ij} = \mathbb{C}^i\mathbb{D}^j$$

of manifestly unitary matrices that satisfy

$$(\mathbb{U}_{ij},\mathbb{U}_{kl}) \equiv \tfrac{1}{n}\mathrm{tr}(\mathbb{U}_{ij}^+\,\mathbb{U}_{kl}) = \delta_{ik}\delta_{jl}$$

In the case $n=3$ those matrices read

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix},\quad \begin{pmatrix}1&0&0\\0&\omega&0\\0&0&\omega^2\end{pmatrix},\quad \begin{pmatrix}1&0&0\\0&\omega^2&0\\0&0&\omega\end{pmatrix}$$

$$\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix},\quad \begin{pmatrix}0&\omega&0\\0&0&\omega^2\\1&0&0\end{pmatrix},\quad \begin{pmatrix}0&\omega^2&0\\0&0&\omega\\1&0&0\end{pmatrix}$$

$$\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix},\quad \begin{pmatrix}0&0&\omega^2\\1&0&0\\0&\omega&0\end{pmatrix},\quad \begin{pmatrix}0&0&\omega\\1&0&0\\0&\omega^2&0\end{pmatrix}$$

with $\omega = e^{i2\pi/3}$.

The matrices $\mathbb{U}_{ij}$ provide a (unitary) *basis* in the space of $n\times n$ matrices, which is to say: every such matrix $\mathbb{A}$ can be developed

$$\mathbb{A} = \sum_{i,j=0}^{n-1} a_{ij}\mathbb{U}_{ij} \quad \text{with} \quad a_{pq} = (\mathbb{U}_{ij},\mathbb{A})$$

which entails

$$\mathbb{A}^+ = \sum_{i,j=0}^{n-1} \bar{a}_{ij}\mathbb{U}_{ij}^+$$

where $\mathbb{U}_{ij}^+ = \mathbb{D}^{-j}\mathbb{C}^{-i} = \omega^{-ij \bmod n}\cdot\mathbb{C}^{-i}\mathbb{D}^{-j} = \omega^{-ij \bmod n}\,\mathbb{U}_{n-i,n-j}$. So we have

$$\begin{aligned}\mathbb{A}^+ &= \sum_{i,j=0}^{n-1} \bar{a}_{ij}\,\omega^{-ij \bmod n}\,\mathbb{U}_{n-i,n-j}\\ &= \sum_{p,q=0}^{n-1} \bar{a}_{n-p,n-q}\,\omega^{-(n-p)(n-q) \bmod n}\,\mathbb{U}_{pq}\\ &\equiv \sum_{p,q=0}^{n-1} \mathrm{A}_{pq}\,\mathbb{U}_{pq}\end{aligned}$$

Hermiticity therefore imposes upon the $\mathbb{U}$-coordinates of $\mathbb{A}$ the conditions

$$a_{pq} = \bar{a}_{n-p,n-q}\,\omega^{-(n-p)(n-q) \bmod n}$$

where interpretation of the expression on the right is subject to the understanding that the subscripts on $\bar{a}_{n-p,n-q}$ are to be read mod n. It is striking that the hermiticity condition—so easy to describe in terms of coordinates that refer to a hermitian basis—becomes so awkward when the coordinates refer to a Schwinger unitary basis. The awkwardness derives largely from the intrusion of modular arithmetic, which *Mathematica* takes in easy stride; I have shown by that means that the result stated above works out correctly in a randomly-constructed 3-dimensional case.

The conditions imposed upon the $\mathbb{U}$-coordinates of $\mathbb{A}$ by the requirement that $\mathbb{A}$ be unitary

$$\mathbb{A}\,\mathbb{A}^+ = \mathbb{U}_{00}$$

are even more intricate, but merit study not least because they possess at least one simply-stated but very pretty property. It proves advantageous to split the problem into two parts: we look first to the coordinates of the matrix $\mathbb{B}$ that satisfies

$$\mathbb{B}\,\mathbb{A}^+ = \mathbb{U}_{00}$$

(*i.e.*, of the inverse of $\mathbb{A}^+$) and then to implications of the condition $\mathbb{B} = \mathbb{A}$.

We have

$$\mathbb{B}\,\mathbb{A}^+ = \sum_{i,j,p,q=0}^{n-1} b_{ij}\mathrm{A}_{pq}\mathbb{U}_{ij}\mathbb{U}_{pq}$$

which by $\mathbb{U}_{ij}\,\mathbb{U}_{pq} = \mathbb{C}^i\,\mathbb{D}^j\,\mathbb{C}^p\,\mathbb{D}^q = \omega^{-jp \bmod n}\cdot\mathbb{C}^{i+p}\,\mathbb{D}^{j+q} = \omega^{-jp \bmod n}\,\mathbb{U}_{i+p,j+q}$ (here again, $i+p$ and $j+q$ are to be evaluated mod n) becomes

$$= \sum_{i,j,p,q=0}^{n-1} b_{ij}\mathrm{A}_{pq}\omega^{-jp \bmod n}\,\mathbb{U}_{(i+p) \bmod n,\,(j+q) \bmod n} \qquad (iv)$$

We would like to be in position to write

$$= \sum_{r,s=0}^{n-1} Z_{rs}(\bar{a}, b)\,\mathbb{U}_{rs} \quad : \quad Z_{rs}(\bar{a}, b) \text{ bilinear}$$

and then to discover the b-coordinates that satisfy

$$Z_{rs}(\bar{a}, b) = \begin{cases} 1 & : \quad r, s = 0, 0 \\ 0 & : \quad \text{otherwise} \end{cases}$$

but are frustrated by the modular arithmetic. To illustrate how that program works out I look to the 2-dimensional case, where things can be written out in detail, the modular arithmetic takes care of itself, and only incidental machine assistance is required.

In the 2-dimensional case one has

$$\mathbb{C} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \omega = e^{i2\pi/2} = -1, \quad \mathbb{D} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

giving[20]

$$\begin{aligned} \mathbb{U}_{00} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbb{U}_{01} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ \mathbb{U}_{10} &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mathbb{U}_{11} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

so

$$\begin{aligned} \mathbb{A} &= \sum_{p,q=0}^{1} a_{pq}\,\mathbb{U}_{pq} = \begin{pmatrix} a_{00} + a_{10} & a_{10} - a_{11} \\ a_{10} + a_{11} & a_{00} - a_{10} \end{pmatrix} \\ \mathbb{B} &= \sum_{p,q=0}^{1} b_{pq}\,\mathbb{U}_{pq} = \begin{pmatrix} b_{00} + b_{10} & b_{10} - b_{11} \\ b_{10} + b_{11} & b_{00} - b_{10} \end{pmatrix} \\ \mathbb{A}^{+} &= \sum_{p,q=0}^{1} \bar{a}_{pq}\,\mathbb{U}^{+}_{pq} = \begin{pmatrix} \bar{a}_{00} + \bar{a}_{10} & \bar{a}_{10} + \bar{a}_{11} \\ \bar{a}_{10} - \bar{a}_{11} & \bar{a}_{00} - \bar{a}_{10} \end{pmatrix} \end{aligned}$$

which supply

$$\mathbb{B}\,\mathbb{A}^{+} = \begin{pmatrix} ba_{00} & ba_{01} \\ ba_{10} & ba_{11} \end{pmatrix}$$

with

$$\begin{aligned} ba_{00} &= (b_{00} + b_{10})(\bar{a}_{00} + \bar{a}_{10}) + (b_{10} - b_{11})(\bar{a}_{10} - \bar{a}_{11}) \\ ba_{01} &= (b_{00} + b_{10})(\bar{a}_{10} + \bar{a}_{11}) + (b_{10} - b_{11})(\bar{a}_{00} - \bar{a}_{10}) \\ ba_{10} &= (b_{10} + b_{11})(\bar{a}_{00} + \bar{a}_{10}) + (b_{00} - b_{10})(\bar{a}_{10} - \bar{a}_{11}) \\ ba_{11} &= (b_{10} + b_{11})(\bar{a}_{10} + \bar{a}_{11}) + (b_{00} - b_{10})(\bar{a}_{00} - \bar{a}_{10}) \end{aligned}$$

The $\mathbb{U}$-expansion of $\mathbb{B}\,\mathbb{A}^{+}$ reads

$$\begin{aligned} \mathbb{B}\,\mathbb{A}^{+} &= \sum_{r,s=0}^{1} Z_{rs}\,\mathbb{U}_{rs} \\ Z_{rs} &= \tfrac{1}{2}\mathrm{tr}\left(\mathbb{U}^{+}_{rs}\mathbb{B}\,\mathbb{A}^{+}\right) \end{aligned}$$

---

[20] This is the only case in which all elements of all Schwinger matrices are real. Werner's construction will—granted the validity of Hadamard's conjecture—present additional cases when n is a multiple of 4.

where *Mathematica* supplies

$$\begin{aligned}
Z_{00} &= b_{00}\bar{a}_{00} + b_{01}\bar{a}_{01} + b_{10}\bar{a}_{10} + b_{11}\bar{a}_{11}\\
Z_{01} &= b_{00}\bar{a}_{01} + b_{01}\bar{a}_{00} - b_{10}\bar{a}_{11} - b_{11}\bar{a}_{10}\\
Z_{10} &= b_{00}\bar{a}_{10} + b_{01}\bar{a}_{11} + b_{10}\bar{a}_{00} + b_{11}\bar{a}_{01}\\
Z_{11} &= -b_{00}\bar{a}_{11} - b_{01}\bar{a}_{10} + b_{01}\bar{a}_{01} + b_{11}\bar{a}_{00}
\end{aligned}$$

which can be written

$$\mathbb{Z}\begin{pmatrix} b_{00}\\ b_{01}\\ b_{10}\\ b_{11}\end{pmatrix} = \begin{pmatrix} Z_{00}\\ Z_{01}\\ Z_{10}\\ Z_{11}\end{pmatrix} \quad\text{with}\quad \mathbb{Z} = \begin{pmatrix} \bar{a}_{00} & \bar{a}_{01} & \bar{a}_{10} & \bar{a}_{11}\\ \bar{a}_{01} & \bar{a}_{00} & -\bar{a}_{11} & -\bar{a}_{10}\\ \bar{a}_{10} & \bar{a}_{11} & \bar{a}_{00} & \bar{a}_{01}\\ -\bar{a}_{11} & -\bar{a}_{10} & \bar{a}_{01} & \bar{a}_{00}\end{pmatrix}$$

Calculation supplies

$$\det\mathbb{Z} = (\det\bar{\mathbb{A}})^2 \quad\text{with}\quad \det\mathbb{A} = a_{00}^2 - a_{01}^2 - a_{10}^2 + a_{11}^2$$

so when $\mathbb{A}$ is non-singular it makes sense to write

$$\begin{pmatrix} b_{00}\\ b_{01}\\ b_{10}\\ b_{11}\end{pmatrix} = \mathbb{Z}^{-1}\begin{pmatrix} Z_{00}\\ Z_{01}\\ Z_{10}\\ Z_{11}\end{pmatrix}$$

and in such non-singular cases we are led to the $\mathbb{U}$-coordinates of $\mathbb{B}$, the inverse of $\mathbb{A}^{+}$:

$$\begin{pmatrix} Z_{00}\\ Z_{01}\\ Z_{10}\\ Z_{11}\end{pmatrix} = \begin{pmatrix} 1\\ 0\\ 0\\ 0\end{pmatrix} \Longrightarrow \begin{pmatrix} b_{00}\\ b_{01}\\ b_{10}\\ b_{11}\end{pmatrix} = (\bar{a}_{00}^2 - \bar{a}_{10}^2 - \bar{a}_{01}^2 + \bar{a}_{11}^2)^{-1}\begin{pmatrix} \bar{a}_{00}\\ -\bar{a}_{01}\\ -\bar{a}_{10}\\ \bar{a}_{11}\end{pmatrix}$$

The 2-dimensional unitarity condition therefore reads

$$\begin{pmatrix} a_{00}\\ a_{01}\\ a_{10}\\ a_{11}\end{pmatrix} = (\bar{a}_{00}^2 - \bar{a}_{10}^2 - \bar{a}_{01}^2 + \bar{a}_{11}^2)^{-1}\begin{pmatrix} \bar{a}_{00}\\ -\bar{a}_{01}\\ -\bar{a}_{10}\\ \bar{a}_{11}\end{pmatrix}$$

I do not possess a parameterized description of all possible solutions of the preceding condition, but note this very pretty immediate corollary:

> If the complex numbers a_{ij} are Schwinger coordinates of a 2-dimensional unitary matrix, then
>
> $$\sum_{i,j=0}^{1} \bar{a}_{ij}a_{ij} = 1$$

Note, however, that <u>the converse is *not* valid</u>.

The equations developed in the preceding paragraph comprise an elaborately disguised variant of some quaternionic algebra that was familiar already to Hamilton. Similar discussion in higher-dimensional cases—if written onto the page, rather than into the memory of a computer—is scarcely feasible (and certainly not informative), since all expressions expand by a factor of $4 \to n^2$. I have, however, managed—with *Mathematica*'s assistance—to establish similar results in randomly constructed cases of of dimension $n = 3$. I am confident that it must be possible to extract such results in full generality from (*iv*), but have at present no clue as to how that might be accomplished.

Werner's construction. Schwinger's construction[21] builds upon ideas that had been developed by Hermann Weyl[22] nearly thirty years previously—a debt which Schwinger acknowledges. Reinhardt Werner, on the other hand, proceeds along a seemingly quite different fresh path to a similar objective. Werner's construction[23] makes essential use of two novel devices: Latin squares and Hadamard matrices.

LATIN, GRAECO-LATIN & MAGIC SQUARES

Latin squares are square arrangements of symbols (call them $1, 2, \ldots, n$ in which each symbol appears exactly once in each row and column:

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \\ 3 & 4 & 1 & 2 \\ 4 & 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \\ 3 & 1 & 4 & 2 \\ 4 & 3 & 2 & 1 \end{pmatrix}$$

Some Latin squares (such as those shown above) can be read as group tables, but others cannot; the following is the smallest example of a Latin square that can be interpreted to refer not to a group but to a "quasi-group" (non-associative group, or "loop"):

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 1 & 5 & 3 \\ 3 & 5 & 4 & 2 & 1 \\ 4 & 1 & 5 & 3 & 2 \\ 5 & 3 & 2 & 1 & 4 \end{pmatrix}$$

But the latter serves Werner's purpose just as well as the others. The literature describes criteria with respect to which Latin squares of the same dimension become "equivalent/inequivalent." The number of inequivalent Latin squares is

---

[21] "Unitary operator bases," PNAS **45**, 570 (1960).

[22] *The Theory of Groups & Quantum Mechanics* (1930), Chapter 4, §14 ("Quantum kinematics as an Abelian group of rotations").

[23] "All teleportation and dense coding schemes," arXiv:quant-ph/003070v1, 17 Mar 2000. See especially §4 "Construcing bases of unitaries." Some of that work was anticipated in K. G. H. Vollbrecht & R. F. Werner, "Why two qubits are special," arXiv:quanat-ph/9910064v1 (14 Oct 1999).

a very rapidly increasing function of dimension: at $n = 2^3$ it has become 283657, and by n = 10 it has reportedly grown to $34817397894749939 \approx 3.48 \times 10^{16}$, which affords Werner plenty of selection options!

We note in passing that[24] that $\mathbb{Z}$ (neglect the signs) is a 4$^{\text{th}}$-order Latin square, of the same design as the first of those shown on the preceding page.

Graeco-Latin squares—sometimes called "Euler squares"—are Latin with respect to each of two distinct symbol-sets, assembled subject to the constraint that every symbol pair appears once and only once (*i.e.*, that the respective Latin squares are "orthogonal"):

$$\begin{pmatrix} Aa & Bc & Cb \\ Bb & Ca & Ac \\ Cc & Ab & Ba \end{pmatrix} \qquad \begin{pmatrix} Aa & Bd & Cb & De & Ec \\ Bb & Ce & Dc & Ea & Ad \\ Cc & Da & Ed & Ab & Be \\ Dd & Eb & Ae & Bc & Ca \\ Ee & Ac & Ba & Cd & Db \end{pmatrix}$$

Noting that Graeco-Latin squares of order 2 are clearly impossible, and finding himself unable to consruct one of order 6, Euler (1782) conjectured that such squares are impossible if $n = 2 \bmod 4$ (*i.e.*, if $n = 2, 6, 10, 14, 18, 22, \ldots$). Euler's conjecture was refuted when, in 1959, counterexamples of order 22 were constructed. Shortly thereafter, UNIVAC produced a square of order 10, and it was established that Euler squares exist for all $n > 2$ except $n = 6$. Little is known about the number of permutationally distinct Euler squares of order $n > 6$ beyond the fact that it is typically *much* smaller than the number of Latin squares of the same order. Graeco-Latin squares play a prominent role in theory and practice having to do with the *design of experiments* (as do "hyper-Graeco-Latin squares," in which more than two distinct symbol-sets are employed), and enter also into some of the many methods for constructing magic squares.

Magic squares are square matrices in which the elements are (typically but not invariably) real numbers, arranged in such a way that all rows, columns and principal diagonals sum to the same number (the "magic constant"). In "normal" magic squares the elements are drawn from the list $\{1, 2, \ldots, n^2\}$ and the magic constant is $\frac{1}{2}n(n^2 + 1)$. Here is the magic square that appears in Albrecht Dürer's *Melencolia* (1514: see the center of the bottom row)

$$\begin{pmatrix} 16 & 3 & 2 & 13 \\ 5 & 10 & 11 & 8 \\ 9 & 6 & 7 & 12 \\ 4 & 15 & 14 & 1 \end{pmatrix}$$

which possesses some unexpected magical features: adding the four elements in the central square—or the four elements at the corners—also produces the magic constant 34. Magic squares have fascinated people of many cultures since

[24] See again page 33.

deep antiquity,[25] and exist in a great variety of variant forms. Here is a non-normal magic square in which all the elements are prime: [26]

$$\begin{pmatrix} 17 & 89 & 71 \\ 113 & 59 & 5 \\ 47 & 29 & 101 \end{pmatrix}$$

To illustrate the simple method by which ordinary "additive magic squares" can be converted into "multiplicative magic squares" I introduce the "Lo Shu square" (purportedly discovered on the back of a turtle in 650 BC, and of mystical importance in Chinese history)

$$\begin{pmatrix} 4 & 9 & 2 \\ 3 & 5 & 7 \\ 8 & 1 & 6 \end{pmatrix}$$

and *use its elements as exponents* to construct

$$\mathbb{M}(x) = \begin{pmatrix} x^4 & x^9 & x^2 \\ x^3 & x^5 & x^7 \\ x^8 & x^1 & x^6 \end{pmatrix} \quad\Longrightarrow\quad \mathbb{M}(2) = \begin{pmatrix} 16 & 512 & 4 \\ 8 & 32 & 128 \\ 256 & 2 & 64 \end{pmatrix}$$

Clearly it is now the *products* of all rows/columns/diagonals that are equal (equal in fact to $2^{15} = 32768$). Notice, however, that while (for example)

$$\begin{pmatrix} 12 & 9 & 12 \\ 36 & 6 & 1 \\ 3 & 4 & 18 \end{pmatrix}$$

is multiplicative it clearly did *not* arise from a construction of the type just described.

HADAMARD MATRICES, REAL & COMPLEX

The theory of Hadamard matrices originates in a paper[27] by J. J. Sylvester, and acquired its name from a paper published twenty-six years later by Jacques

---

[25] Benjamin Franklin (see Google) was a prolific inventor of magic squares, an activity with which he says (1771) he busied himself while serving as secretary to meetings at which he could perforce not speak. Actually, some of his ingenious designs date from his youth, though he could, when called upon, reproduce them at advanced age; it appears to have been a life-long addiction.

[26] It is known that such squares exist of all orders. The magic constant in this instance is 177.

[27] "Thoughts on inverse orthogonal matrices, simultaneous sign successions, and tessellated pavements in two or more colors, with applications to Newton's rule, ornamental tile-work, and the theory of numbers," Phil. Mag. **34**, 461–475, (1867).

Hadamard. Hadamard matrices are square matrices with all elements equal to ± 1 and with the further property that all rows/columns are orthogonal, which forces the dimenson of such matrices to be even, and entails

$$\mathbb{H}\,\mathbb{H}^{\mathsf{T}} = n\,\mathbb{I}$$

In the simplest instance one has

$$\mathbb{H}_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Sylvester himself contemplated matrices of progressively higher order

$$\mathbb{H}_4 = \mathbb{H}_2 \otimes \mathbb{H}_2 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix},\ \mathbb{H}_8 = \mathbb{H}_2 \otimes \mathbb{H}_4\,,\ \textit{etc.}$$

Such matrices have dimension $\{2, 4, 8, 16, 32, 64, \ldots, 2^n\}$. The still-unproven *Hadamard conjecture* asserts that real Hadamard matrices exist in all dimensions that are multiples of 4, which would fill in these gaps in Sylvester's list: $\{12, \star, 20, 24, 28, \star, 36, 40, 44, 48, 52, 56, 60, \star, \ldots\}$. As of 2008, the least value of n for which Hadamard's conjecture has not been confirmed is $n = 688 = 4 \times 172$, and there were a total of thirteen such cases with $n < 2000$. The real Hadamard matrices are (given the natural interpretation of "equivalence" supplied by the literature) unique through $n = 2, 4, 6, 12$, but 5 inequivalent Hadamard matrices exist for $n = 16$, and millions are known for $n \geq 32$. This again provides Werner with plenty of room to wiggle, at least in dimensions that are multiples of four.[28]

Complex Hadamard matrices—which satisfy the complexified condition

$$\mathbb{H}\,\mathbb{H}^{+} = n\,\mathbb{I}$$

[28] A word about how Hadamard's motivation, which differed markedly from Sylvester's: Let $\{x_0, y_0, z_0\}, \{x_1, y_1, z_1\}, \{x_2, y_2, z_2\}, \{x_3, y_3, z_3\}$ be Cartesian coordinates of the vertices of a 3-simplex (tetrahedron) contained within a unit cube (all coordinates bounded by ± 1). One has

$$\text{Volume} = \tfrac{1}{3!}\det\begin{pmatrix} 1 & x_0 & y_0 & z_0 \\ 1 & x_1 & y_1 & z_1 \\ 1 & x_2 & y_2 & z_2 \\ 1 & x_3 & y_3 & z_3 \end{pmatrix}$$

which according to *Hadamard's determinant theorem* assumes its maximal value when the matrix is equivalent to $\mathbb{H}_4$.

—exist in *all* dimensions, including those that are *not* multiples of four. The most important class of such matrices are those of "Butson type,"[29] which in n-dimensions possess the "Fourier structure"

$$\mathbb{F}_n = \|F_{n,jk}\| \quad \text{with} \quad F_{n,jk} = \omega^{jk \bmod n}$$

where $\omega = e^{i\,2\pi/n}$ and $j, k \in \{0, 1, 2, \ldots, n-1\}$. Low-dimensional examples look like this:

$$\begin{aligned}
\mathbb{F}_2 &= \begin{pmatrix} 1 & 1 \\ 1 & \omega \end{pmatrix} && \text{with} \quad \omega = e^{i\,2\pi/2} = -1 \\
\mathbb{F}_3 &= \begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega^4 \end{pmatrix} && \text{with} \quad \omega = e^{i\,2\pi/3} \\
&= \begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{pmatrix} \\
\mathbb{F}_4 &= \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & \omega & \omega^2 & \omega^3 \\ 1 & \omega^2 & \omega^4 & \omega^6 \\ 1 & \omega^3 & \omega^6 & \omega^9 \end{pmatrix} && \text{with} \quad \omega = e^{i\,2\pi/4} = i \\
&= \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & \omega & \omega^2 & \omega^3 \\ 1 & \omega^2 & 1 & \omega^2 \\ 1 & \omega^3 & \omega^2 & \omega \end{pmatrix}
\end{aligned}$$

Notice that in each case the elements of every row/column (except the 0^{th}) sum to zero; the $\mathbb{F}$-matrices can in that sense be said (not very usefully) to be "semi-magical."[30] From $\mathbb{H}\,\mathbb{H}^+ = n\,\mathbb{I}$ we see that

$$\mathbb{U}_n = \tfrac{1}{\sqrt{n}}\mathbb{H}_n$$

is unitary. Familiarly, such matrices are central to the theory of discrete Fourier transforms.[31] Notice finally that the complex numbers that appear on the k^{th}

---

[29] See http://en.wikipedia.org/wiki/Butson-type_Hadamard_matrices. The original reference is A. T. Butson, "Generalized Hadamard matrices," Proc. Amer. Math. Soc. **13**, 894-898 (1962).

[30] Properly complex magic squares present no mystery; look, for instance, to these Lo Shu examples (which might have appeared on the back of a complex turtle):

$$\begin{pmatrix} 4+\alpha & 9+\alpha & 2+\alpha \\ 3+\alpha & 5+\alpha & 7+\alpha \\ 8+\alpha & 1+\alpha & 6+\alpha \end{pmatrix} + i\beta \begin{pmatrix} 2 & 7 & 6 \\ 9 & 5 & 1 \\ 4 & 3 & 8 \end{pmatrix}$$

where α and β are any real or complex numbers. The magic constant in this instance is $15 + (3\alpha + i15\beta)$.

[31] See, for example, http://en.wikipedia.org/wiki/Discrete_Fourier_transform.

row/column of $\mathbb{F}_n$ are precisely the numbers that appear on the diagonal of

$$\mathbb{D}_n{}^k = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \omega & 0 & 0 & \dots & 0 \\ 0 & 0 & \omega^2 & 0 & \dots & 0 \\ 0 & 0 & 0 & \omega^3 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \omega^{n-1} \end{pmatrix}^k$$

To create the $n \times n$ elements of a unitary basis, Werner would have us select *any* n^{th}-order Hadamard matrix $\mathbb{H}$ and *any* n^{th}-order Latin square $\mathbb{L}$, and from that material assemble

$$\begin{aligned} \mathbb{W}_{ij} &= \|W_{ij,pq}\| \\ &\qquad W_{ij,pq} = H_{ip} \cdot \text{KroneckerDelta}[q, L_{jp}] \end{aligned}$$

Evidently, $\mathbb{H}$ supplies the numerical data that will be displayed as elements of the matrices $\mathbb{W}_{ij}$ while $\mathbb{L}$ controls the deployment of that data. To illustrate how this works we look to the case $n = 3$, in which, as it happens, we are essentially forced to select[32]

$$\mathbb{H} = \begin{pmatrix} H_{00} & H_{01} & H_{02} \\ H_{10} & H_{11} & H_{12} \\ H_{20} & H_{21} & H_{22} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega^4 \end{pmatrix}$$

$$\mathbb{L} = \begin{pmatrix} L_{00} & L_{01} & L_{02} \\ L_{10} & L_{11} & L_{12} \\ L_{20} & L_{21} & L_{22} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

Thus

$$\begin{aligned} \mathbb{W}_{00} = \|H_{0p}\Delta(q, L_{0p})\| &= \begin{pmatrix} H_{00}\Delta(0, L_{00}) & H_{00}\Delta(1, L_{00}) & H_{00}\Delta(2, L_{00}) \\ H_{01}\Delta(0, L_{01}) & H_{01}\Delta(1, L_{01}) & H_{01}\Delta(2, L_{01}) \\ H_{02}\Delta(0, L_{02}) & H_{02}\Delta(1, L_{02}) & H_{02}\Delta(2, L_{02}) \end{pmatrix} \\ &= \begin{pmatrix} H_{00}\delta_{00} & H_{00}\delta_{10} & H_{00}\delta_{20} \\ H_{01}\delta_{01} & H_{01}\delta_{11} & H_{01}\delta_{21} \\ H_{02}\delta_{02} & H_{02}\delta_{12} & H_{02}\delta_{22} \end{pmatrix} \\ &= \begin{pmatrix} H_{00} & 0 & 0 \\ 0 & H_{01} & 0 \\ 0 & 0 & H_{02} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

[32] Here again $\omega = e^{i2\pi/3}$, and in the description of $\mathbb{L}$ we have adopted a symbol-set that conforms to our convention that indices range on $\{0, 1, 2\}$.

$$\mathbb{W}_{01} = \|H_{0p}\Delta(q, L_{1p})\| = \begin{pmatrix} H_{00}\Delta(0,L_{10}) & H_{00}\Delta(1,L_{10}) & H_{00}\Delta(2,L_{10}) \\ H_{01}\Delta(0,L_{11}) & H_{01}\Delta(1,L_{11}) & H_{01}\Delta(2,L_{11}) \\ H_{02}\Delta(0,L_{12}) & H_{02}\Delta(1,L_{12}) & H_{02}\Delta(2,L_{12}) \end{pmatrix}$$
$$= \begin{pmatrix} H_{00}\delta_{01} & H_{00}\delta_{11} & H_{00}\delta_{21} \\ H_{01}\delta_{02} & H_{01}\delta_{12} & H_{01}\delta_{22} \\ H_{02}\delta_{00} & H_{02}\delta_{10} & H_{02}\delta_{20} \end{pmatrix}$$
$$= \begin{pmatrix} 0 & H_{00} & 0 \\ 0 & 0 & H_{01} \\ H_{02} & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

The preceding statements illustrate how it comes about that the 0^{th} row of $\mathbb{L}$ controls the design of $\{\mathbb{W}_{i0} : i = 0, 1, \ldots\}$, the 1^{th} row controls the design of $\{\mathbb{W}_{i1} : i = 0, 1, \ldots\}$, *etc.* Since $\mathbb{W}_{01}$ and $\mathbb{W}_{11}$ (see below) are under identical control we are not surprised to find that they display the same essential structure (same deployment of 0s):

$$\mathbb{W}_{11} = \|H_{1p}\Delta(q, L_{1p})\| = \begin{pmatrix} H_{10}\Delta(0,L_{10}) & H_{10}\Delta(1,L_{10}) & H_{10}\Delta(2,L_{10}) \\ H_{11}\Delta(0,L_{11}) & H_{11}\Delta(1,L_{11}) & H_{11}\Delta(2,L_{11}) \\ H_{12}\Delta(0,L_{12}) & H_{12}\Delta(1,L_{12}) & H_{12}\Delta(2,L_{12}) \end{pmatrix}$$
$$= \begin{pmatrix} H_{10}\delta_{01} & H_{10}\delta_{11} & H_{10}\delta_{21} \\ H_{11}\delta_{02} & H_{11}\delta_{12} & H_{11}\delta_{22} \\ H_{12}\delta_{00} & H_{12}\delta_{10} & H_{12}\delta_{20} \end{pmatrix}$$
$$= \begin{pmatrix} 0 & H_{10} & 0 \\ 0 & 0 & H_{11} \\ H_{12} & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & \omega \\ \omega^2 & 0 & 0 \end{pmatrix}$$

From

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & \omega \\ \omega^2 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \omega^2 \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix}$$
$$\Downarrow$$
$$\mathbb{W}_{11} = \omega^2 \cdot \mathbb{U}_{11} \quad , \ \textit{etc.}$$

it becomes clear (i) why Werner speaks of "shift-and-multiply bases" and (ii) that <u>Werner's unitary bases differ only cosmetically from Schwinger's</u>. We can therefore consider unitarity, tracelessness (except in the case $i = j = 0$) and trace-wise orthonormality to have been already established. Werner's construction—precisely because it makes such clever use of Latin squares and Hadamard matrices—is so convoluted as to make the direct *demonstration* of such properties (ditto the formulation of hermiticity, inversion and unitarity conditions) relatively awkward; in those respects Schwinger's construction presents distinct advantages. On the other hand, the fact that (when n is not too small) Latin squares and Hadamard matrices come in so many flavors lends a remarkable—and potentially useful—range and degree of flexibility to the Werner formalism. Note particularly that, when n is either 2 or a multiple of 4, Werner's construction can be used to create unitary bases all elements of which are *real* (*i.e.*, rotational).