Nonstandard applications of

MOHR'S CONSTRUCTION

Nicholas Wheeler, Reed College Physics Department
June 1998

Introduction. Many years ago I was motivated by adventures in the laboratory to pose this question:[1]

> *Given that the angular velocity* $\boldsymbol{\omega}$ *and angular momentum (spin)* $\boldsymbol{S}$ *of a rigid body stand in this familiar relationship*
>
> $$\boldsymbol{S} = \mathbb{I}\,\boldsymbol{\omega} \quad : \quad \mathbb{I} \textit{ is the symmetric } \text{moment of inertia } \textit{matrix}$$
>
> *what is the maximal angle that* $\boldsymbol{\omega}$ *and* $\boldsymbol{S}$ *can subtend?*

My solution of that problem hinged on a simple geometrical construction (described below) which, as I subsequently discovered, was original not to me but to Christian Otto Mohr (1882), who had himself built upon a suggestion of Karl Culmann (1866). That Culmann and Mohr were concerned not with the dynamics of tops but with stress analysis and the fracture of brittle materials[2] —and yet managed to anticipate me—can be a source of no deep surprise; the equation $\boldsymbol{y} = \mathbb{M}\,\boldsymbol{x}$ ($\mathbb{M}$ symmetric) is so primitive, and encountered in so many physical settings, that interesting remarks derived from the *structure* of that

---

[1] See CLASSICAL GYRODYNAMICS (1976), pp. 92–106.

[2] Culmann (1821–1881) was a German professor of civil engineering who is remembered today mainly for his contributions—some of which had been anticipated by Maxwell—to "graphical statics." Mohr (1835–1918) taught civil engineering first in Stuttgart and then (from 1873 until his retirement in 1900) in Dresden. He was said by his student A. Föppl (who himself figures importantly in the history of electrodynamics, and whose texts influenced the development of the young Einstein) to have been an outstanding teacher: a tall, proud and taciturn man who spoke and wrote with simplicity, clarity and conciseness. "Mohr's stress circle"known—provided the basis for his theory of stress failure; for an account of the "Coulomb-Mohr fracture criterion" see (for example) C. C. Mei, *Mathematical Analysis in Engineering* (1995), p. 150.

equation are destined to be repeatedly rediscovered, and pressed into service in a great variety of applications.[3]

Two circumstances motivate my return to this topic:

The "density matrix" (which in recent literature is more properly called the "density *operator*")

$$\boldsymbol{\rho} = \sum\!\!\!\!\!\!\int |\psi_\nu)p_\nu(\psi_\nu| \quad : \quad \text{the } p_\nu \text{ are probabilities, and sum to unity}$$

serves in orthodox quantum mechanics to describe a "mixture" of quantum states $|\psi_\nu)$, which (since no orthogonality or linear independence requirement attaches to the concoction of such mixtures) may be either finite or discretely/continuously infinite in number, but which I will assume collectively span a finite-dimensional subspace $\mathcal{H}_n$ in the space $\mathcal{H}$ of states. The density matrix is manifestly hermitian, so admits of unique spectral representation

$$\boldsymbol{\rho} = \sum_{k=1}^{n} |k)\rho_k(k|$$

But the latter expression can be considered to describe a mixture of finitely many *orthogonal* states. In some previous work[4] I was led thus to the realization that an element of ambiguity attaches to the "mixed state" concept; distinct mixtures can give rise to the same density matrix $\boldsymbol{\rho}$, and have therefore to be considered physically equivalent. The question now arises: How does one most informatively describe (i.e., how does one *understand*) the conditions under which ostensibly distinct mixtures which are, in this sense, "equivalent"? When this question was posed to Tom Wieting he provided—almost instantly—an elegant response which is, however, special to the case $n = 2$. My efforts to generalize "Wieting's construction" have at several points acquired the scent of Mohr's construction, which I review now with these questions foremost in mind:

- Can Mohr's construction be used to construct a generalizable reformulation of Wieting's construction? If not,
- Can Mohr's construction be used to establish that (and why) Wieting's construction does not admit of generalization?

Oz Bonfim has directed my attention to a paper[5] which is of interest to him in connection with his own research, and which—when scanned with squinted eyes, and to my predisposed nose—has again the "scent of Mohr's construction." I have secondary interest in discovering whether that hunch can be supported.

---

[3] For an electrodynamical application, having nothing at all to do either with tops or with fracture, see my CLASSICAL ELECTRODYNAMNICS (1980), p. 127.

[4] "Status and some ramifications of Ehrenfest's theorem" (1998), §9.

[5] S. Habib & R. D. Ryne, "Symplectic Calculation of Lyapunov Exponents," Phys. Rev. Letters **74**, 70 (1995).

Those two lines of motivation have one feature in common: each requires some enlargement of the setting within which Mohr's construction is usually contemplated. The organization of my remarks will reflect my desire to render that process—"enlargement of the setting"—as smooth and natural as possible.

1. Point of departure: Mohr's construction in the 2-dimensional case. One natural approach to the "top problem" posed earlier would be to (i) go to the principal axis frame of the body (with respect to which the moment of inertia matrix is diagonal), (ii) let $\boldsymbol{\omega}$ assume all possible orientations, and (iii) study the placement, relative to $\boldsymbol{\omega}$, of

$$\boldsymbol{S} = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \boldsymbol{\omega} \quad \text{as } \boldsymbol{\omega} \text{ ranges over the } \omega\text{-sphere}$$

Alternatively—and for present purposes more usefully—one might (i) assign $\boldsymbol{\omega}$ any convenient fixed value, (ii) let the body (which is to say: the principal axis frame) assume all possible orientations with respect to some fixed frame, and (iii) study the relative placement of

$$\boldsymbol{S} = \mathbb{R}^{-1} \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \mathbb{R} \cdot \begin{pmatrix} \omega \\ 0 \\ 0 \end{pmatrix} \quad \text{as } \mathbb{R} \text{ ranges over } O(3) \tag{1}$$

It becomes in this light natural to look, by way of preliminary orientation, to this dimensionally reduced analog of the preceding problem: study the relative placement of

$$\boldsymbol{S} = \mathbb{R}^{-1} \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \mathbb{R} \cdot \begin{pmatrix} \omega \\ 0 \end{pmatrix} \quad \text{as } \mathbb{R} \text{ ranges over } O(2) \tag{2}$$

and it is the latter problem that gives rise to Mohr's construction in its most familiar form.

It is interesting to note in passing that what I have called elsewhere[6] the "method of dimensional reduction" is inapplicable to the theory of tops. Only in the 3-dimensional case does rotational kinematics lead to a "angular velocity (psuedo) vector;" only in that case is the angular momentum concept captured by a construction of the design $\boldsymbol{r} \times \boldsymbol{p}$. The physicists of Flatland may have instructive things to say about the fracture of brittle materials (if such can even exist in two dimensions), but their theory of tops is a pallid affair: a "theory of wheels." One can, however, recover (2) from (1) by confining $\boldsymbol{\omega}$ to the plane normal to the 3^{rd} principal axis; i.e., by considering the 2-dimensional problem to be a *constrained instance* of the 3-dimensional problem.

---

[6] "Electrodynamics in 2-dimensional spacetime" (1997).

We have

$$\mathbb{R}^{-1}\begin{pmatrix} a & 0\\ 0 & b\end{pmatrix}\mathbb{R} = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}^{\mathsf{T}}\begin{pmatrix} a & 0\\ 0 & b\end{pmatrix}\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$$
$$= \begin{pmatrix} a\cos^2\theta + b\sin^2\theta & (b-a)\cos\theta\sin\theta\\ (b-a)\cos\theta\sin\theta & b\cos^2\theta + a\sin^2\theta\end{pmatrix}$$

and, drawing upon the identities $\cos^2 = \frac{1}{2}(1+\cos 2\theta)$ and $\sin^2 = \frac{1}{2}(1-\cos 2\theta)$, obtain

$$= \begin{pmatrix} \frac{1}{2}(a+b) + \frac{1}{2}(a-b)\cos 2\theta & -\frac{1}{2}(a-b)\sin 2\theta\\ -\frac{1}{2}(a-b)\sin 2\theta & \frac{1}{2}(a+b) - \frac{1}{2}(a-b)\cos 2\theta\end{pmatrix} \tag{3}$$
$$\equiv \mathbb{M}(\theta)$$

which I will—for lack of any standard terminology, and since it lies at the algebraic heart of Mohr's construction—call the "Mohr matrix;" it is, in the present application, just the moment of intertia matrix, referred to an arbitrarily-oriented frame,[7] but in other applications (Mohr's own, for example) acquires other interpretations. I will return in a moment to discussion of some of the distinctive properties of Mohr matrices.

As θ ranges on $[0, 2\pi]$ the vector

$$\boldsymbol{s}(\theta) \equiv \mathbb{M}(\theta)\,\hat{\boldsymbol{\omega}} \quad \text{with} \quad \hat{\boldsymbol{\omega}} \equiv \begin{pmatrix}1\\0\end{pmatrix}$$
$$= \begin{pmatrix} \frac{1}{2}(a+b) + \frac{1}{2}(a-b)\cos 2\theta\\ -\frac{1}{2}(a-b)\sin 2\theta\end{pmatrix} \tag{4}$$

traces a closed curve on the s-plane; specifically, it traces—twice!—a circle, the so-called "Mohr circle," shown in Figure 1.

It is clear from the figure, and illustrative of its utility, that

$$\sigma \equiv \text{maximal angular separation between } \boldsymbol{s}(\theta) \text{ and } \hat{\boldsymbol{\omega}}$$
$$= \arcsin\left\{\frac{\frac{1}{2}(a-b)}{\frac{1}{2}(a+b)}\right\} \tag{5}$$

and that maximality is achieved when $\cos(\pi - 2\theta_{\max}) = -\cos 2\theta_{\max} = \frac{a-b}{a+b}$, which entails

$$\sin^2\theta_{\max} = \frac{a}{a+b} \qquad \text{and} \qquad \cos^2\theta_{\max} = \frac{b}{a+b}$$

Every real symmetric 2×2 matrix $\mathbb{M} = \begin{pmatrix} A & C\\ C & B\end{pmatrix}$ is latently a "Mohr matrix," and can be brought to "Mohr form" by inverted use of Mohr's construction, as described by Figure 2. The resulting compass-&-ruler construction of the spectrum and eigenvectors of $\mathbb{M}$ is in some respects reminiscent of a spectral estimation technique (more familiar to engineers than to physicists) which derives from "Geršgorin's theorem." For details, see an appendix to the CLASSICAL GYRODYNAMICS notes cited previously.

[7] I cannot write $\mathbb{I}(\theta)$ because $\mathbb{I}$ is a reserved symbol

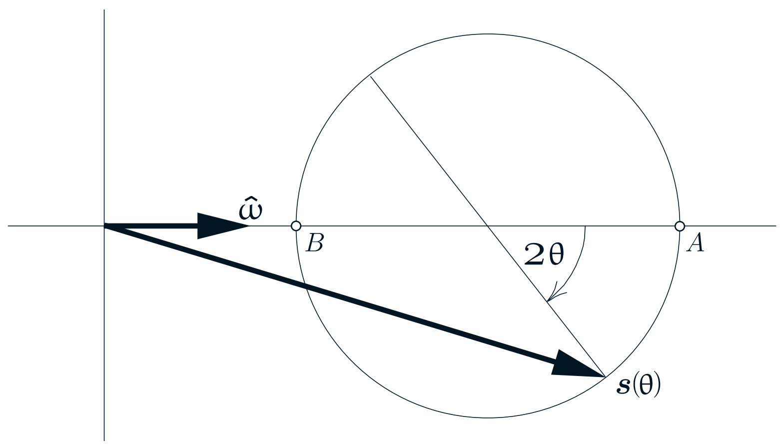

FIGURE 1: *Graph of (4), drawn in the presumption that* $a > b > 0$. *The "Mohr circle" is traced by* $\boldsymbol{s}(\theta)$ *as* θ *ranges on* $[0, 2\pi]$*; it is, that is to say, traced (twice) by*

$$\boldsymbol{s} \equiv \mathbb{R}^{-1}\left(\begin{smallmatrix} a & 0 \\ 0 & b \end{smallmatrix}\right)\mathbb{R}\cdot\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \quad \textit{as } \mathbb{R} \textit{ ranges over } O(2)$$

The circle is centered at $(\frac{a+b}{2}, 0)$*, has radius* $\frac{a-b}{2}$*, and intercepts the horizontal axis at points* $s_1 = a$ *and* $s_1 = b$ *which mark the obvious* eigenvalues *of the "Mohr matrix"* $\mathbb{M} \equiv \mathbb{R}^{-1}\left(\begin{smallmatrix} a & 0 \\ 0 & b \end{smallmatrix}\right)\mathbb{R}$.

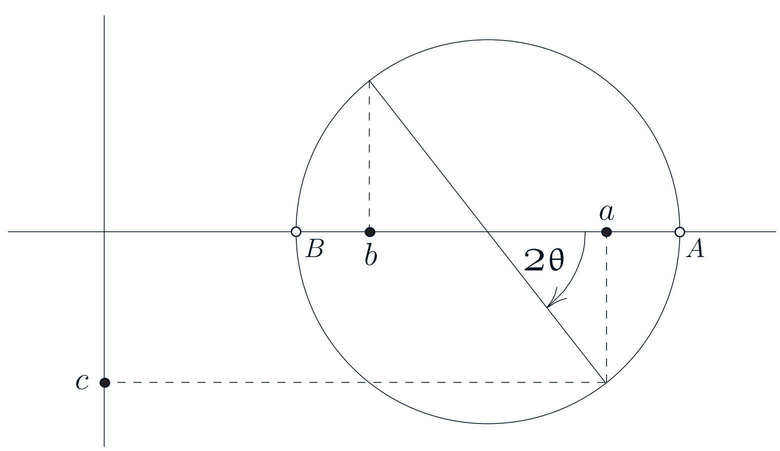

FIGURE 2: *Compass-&-ruler construction of the spectral properties of* $\mathbb{M} = \left(\begin{smallmatrix} A & C \\ C & B \end{smallmatrix}\right)$. *One uses Mohr's construction "backwards" to read off the eigenvalues* a *and* b *and the angular parameter* θ *that fixes the locations of the associated eigenvectors:*

$$\mathbb{M}\boldsymbol{e}_1 = a\boldsymbol{e}_1 \quad \textit{with} \quad \boldsymbol{e}_1 = \mathbb{R}^{-1}\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} +\cos\theta \\ -\sin\theta \end{smallmatrix}\right)$$
$$\mathbb{M}\boldsymbol{e}_2 = b\boldsymbol{e}_2 \quad \textit{with} \quad \boldsymbol{e}_2 = \mathbb{R}^{-1}\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} +\sin\theta \\ +\cos\theta \end{smallmatrix}\right)$$

We note in passing that "rotating a vector $\boldsymbol{\omega}$ with respect to a fixed blob" is the same as "*counterrotating* the blob with respect to a fixed vector," and that this simple fact accounts for the *sense* of the angular advancement in Mohr's construction, as I have presented it.

Working from (3) we have

$$\mathbb{M}(\theta) = a\,\mathbb{P}_1 + b\,\mathbb{P}_2 \tag{6}$$

where

$$\mathbb{P}_1 \equiv \tfrac{1}{2}\begin{pmatrix} 1+\cos 2\theta & -\sin 2\theta \\ -\sin 2\theta & 1-\cos 2\theta \end{pmatrix} \tag{7.1}$$

$$\mathbb{P}_2 \equiv \tfrac{1}{2}\begin{pmatrix} 1-\cos 2\theta & +\sin 2\theta \\ +\sin 2\theta & 1+\cos 2\theta \end{pmatrix} \tag{7.2}$$

are seen to comprise a complete system of orthogonal projection operators:

$$\mathbb{P}_1 + \mathbb{P}_2 = \mathbb{I}\,, \quad \mathbb{P}_1\cdot\mathbb{P}_2 = \mathbb{O}, \quad \text{and} \quad \begin{cases} \mathbb{P}_1^2 = \mathbb{P}_1 \\ \mathbb{P}_2^2 = \mathbb{P}_2 \end{cases} \tag{8}$$

The right side of (6) achieves the "spectral resolution" of $\mathbb{M}(\theta)$, but conceals no deep mystery: the spectral resolution of $\left(\begin{smallmatrix} a & 0 \\ 0 & b \end{smallmatrix}\right)$ is trivial

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} = a\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

and one has only to multiply that equation by $\mathbb{R}^{-1}$ on the left and $\mathbb{R}$ on the right to recover precisely (6). From

$$\mathrm{tr}\,\mathbb{P}_1 = \mathrm{tr}\,\mathbb{P}_2 = 1$$

we learn that $\mathbb{P}_1$ and $\mathbb{P}_2$ project onto one-dimensional spaces; in fact

$$\left.\begin{array}{l} \mathbb{P}_1 \text{ projects onto the } \boldsymbol{e}_1\text{-ray} \\ \mathbb{P}_2 \text{ projects onto the } \boldsymbol{e}_2\text{-ray} \end{array}\right\} \tag{9}$$

and both can be recovered as instances of the general proposition that if $\boldsymbol{e} \equiv \left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right)$ is a unit vector then the projector onto the $\boldsymbol{e}$-ray can be described $\left(\begin{smallmatrix} pp & pq \\ qp & qq \end{smallmatrix}\right)$. All such projectors are, by the way, manifestly symmetric. We observe finally that $\mathbb{P}_1$ can be written

$$\mathbb{P}_1 = \tfrac{1}{2}\big\{\mathbb{I} + p^1\mathbb{S}_1 + p^2\mathbb{S}_2\big\} \tag{10}$$

with

$$\boldsymbol{p} \equiv \begin{pmatrix} p^1 \\ p^2 \end{pmatrix} \equiv \begin{pmatrix} +\cos 2\theta \\ -\sin 2\theta \end{pmatrix}, \quad \mathbb{S}_1 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad \mathbb{S}_2 \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{11}$$

and that $\boldsymbol{p} \longrightarrow -\boldsymbol{p}$ sends $\mathbb{P}_1 \longrightarrow \mathbb{P}_2 \perp \mathbb{P}_1$:

$$\mathbb{P}_2 = \tfrac{1}{2}\big\{\mathbb{I} - p^1\mathbb{S}_1 - p^2\mathbb{S}_2\big\} \tag{12}$$

While it might seem most natural to associate projectors $\mathbb{P}$ with the vectors $\boldsymbol{e}$ onto which they project, we have at (10) been led to assign $\mathbb{P}$ a vectorial "address" $\boldsymbol{p}$ which is distinct from $\boldsymbol{e}$. The distinction is signaled by the presence of certain factors of 2. Clearly we have been led to a kind of "toy instance" of the $SU(2)$ representation of $O(3)$, and it is with Pauli matrices in mind ("sigma matrices") that—lacking a double-stroke version of σ—I have adopted my $\mathbb{S}$ notation.

In Figure 3 I show the Mohr constructions associated with the equations

$$\boldsymbol{s}_1(\theta) = \mathbb{P}_1(\theta)\hat{\boldsymbol{\omega}} \quad \text{and} \quad \boldsymbol{s}_2(\theta) = \mathbb{P}_2(\theta)\hat{\boldsymbol{\omega}}$$

My several attempts to illustrate the geometrical mechanism by which

$$\boldsymbol{s}(\theta) = a\boldsymbol{s}_1(\theta) + b\boldsymbol{s}_2(\theta)$$

reproduces Figure 1 were all too confusingly complex to be useful, but it was that very failure which led me to the realization that a simple "complexification trick" (see the caption of Figure 3) reduces the point at issue almost to a triviality: if we make the associations

$$\begin{aligned}\boldsymbol{s}_1(\theta) &\longleftrightarrow z_1(\theta) \equiv \tfrac{1}{2}(1 + e^{-2i\theta})\\ \boldsymbol{s}_2(\theta) &\longleftrightarrow z_2(\theta) \equiv \tfrac{1}{2}(1 - e^{-2i\theta})\end{aligned}$$

it is then immediate that

$$\boldsymbol{s}(\theta) \longleftrightarrow z(\theta) = az_1(\theta) + bz_2(\theta) = \tfrac{a+b}{2} + \tfrac{a-b}{2}e^{-2i\theta}$$

2. First approach to the density matrix problem. We were led just above to

- associate projectors $\mathbb{P}$ with points on the "unit Mohr circle" (Figure 3)
- associate points on the "unit Mohr circle" with complex numbers

and to the observation that weighted sums of projectors (of which we have so far considered only one—exceptionally simple—single instance) are thus made susceptible to analysis as weighted sums of complex numbers. Enlarging upon that observation, we look to symmetric matrices (toy density matrices) of the design

$$\mathbb{D} = \sum_k w_k\mathbb{P}_k$$

where the weights w_k are positive real numbers subject (we may without loss of generality assume) to the constraint $\sum_k w_k = 1$. From

$$\mathbb{D} \quad \longleftrightarrow \quad \text{weighted set of points on the unit Mohr circle}$$

we are led to write

$$D = \sum_k w_k \tfrac{1}{2}\big\{1 + e^{-2i(\theta+\alpha_k)}\big\}$$

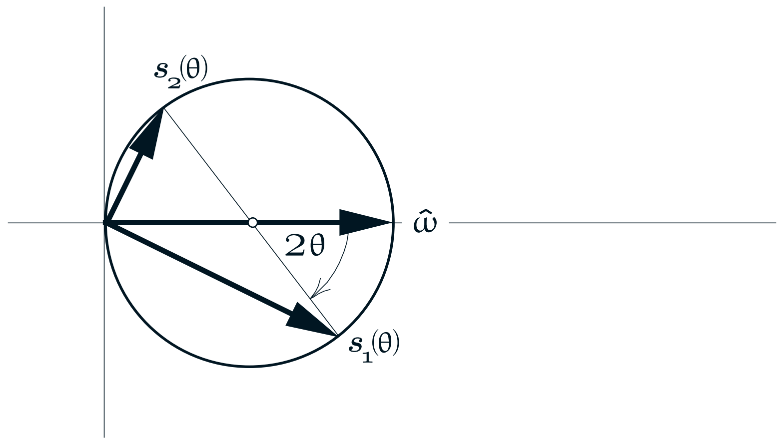

FIGURE 3: *Superimposed Mohr constructions associated with the equations*

$$\boldsymbol{s}_1(\theta) = \mathbb{P}_1(\theta)\hat{\boldsymbol{\omega}} \quad \text{and} \quad \boldsymbol{s}_2(\theta) = \mathbb{P}_2(\theta)\hat{\boldsymbol{\omega}}$$

It is geometrically evident, and easy to prove analytically, that

$$\boldsymbol{s}_1(\theta) \perp \boldsymbol{s}_2(\theta)$$

It becomes natural to associate the s-plane with the complex plane, writing

$$\begin{aligned} \boldsymbol{s}_1(\theta) &\longleftrightarrow z_1(\theta) \equiv \tfrac{1}{2}(1 + e^{-2i\theta}) \\ \boldsymbol{s}_2(\theta) &\longleftrightarrow z_2(\theta) \equiv \tfrac{1}{2}(1 - e^{-2i\theta}) \end{aligned}$$

But $\sum w_k e^{-2i\alpha_k} = re^{-2i\alpha}$ with $0 \leq r \leq 1$ so (writing $\vartheta \equiv \theta + \alpha$) we have this "Mohr representation" of D

$$D = \tfrac{1}{2}(1 + re^{-2i\vartheta}) \tag{13}$$

and this "spectral representation:"

$$\begin{aligned} D = w \cdot \tfrac{1}{2}(1 + e^{-2i\vartheta}) + (1 - w) \cdot \tfrac{1}{2}(1 - e^{-2i\vartheta}) \\ \Big| \qquad\qquad\qquad\qquad \\ w \equiv \tfrac{1}{2}(1 + r) \qquad\qquad \end{aligned} \tag{14}$$

These results are illustrated in Figures 4 & 5. They arise from what I consider to be a "non-standard application of Mohr's construction," and cast in new light—I am not yet prepared to say in improved light—the essence of "Wieting's construction," at least as it pertains to our toy density matrices.

Figure 4 goes here

FIGURE 4: *Weighted points decorate the unit Mohr circle. Each represents a weighted projector* $w_k\mathbb{P}_k$*. Their "center of mass" lies at the point* $re^{-2i\vartheta}$ *representative of the symmetric density matrix* $\mathbb{D} = \sum w_k\mathbb{P}_k$*. The associated Mohr circle crosses the horizontal axix at points which mark the*

$$\text{eigenvalues of } \mathbb{D} = \tfrac{1}{2}(1 \pm r)$$

Clearly, diverse weighted distributions can share the same center of mass; diverse mixtures can give rise to the same density matrix.

Figure 5 goes here

FIGURE 5: *"Spectral representation" of the preceding* $\mathbb{D}$*-matrix. The projectors stand diametrically opposite one another, and (therefore) project onto orthogonal vectors. From the "teeter-totter condition"* $w(1-r) = (1-w)(1+r)$ *we obtain* $w = \frac{1}{2}(1+r)$*.*

3. General principles preliminary to generalization. Let $\boldsymbol{x}, \boldsymbol{y}, \ldots$ be real column vectors

$$\boldsymbol{x} = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}, \text{ etc.}$$

let $\mathbb{G}$ be an invertible $n \times n$ matrix, and agree to write

$$\big\{\boldsymbol{x}, \boldsymbol{y}\big\} \equiv \boldsymbol{x}^{\mathsf{T}}\mathbb{G}\,\boldsymbol{y} \tag{15}$$

Imposition of a requirement that the linear transformation $\boldsymbol{x} \to \mathbb{T}\boldsymbol{x}$ preserve curly brackets

$$\big\{\mathbb{T}\boldsymbol{x}, \mathbb{T}\boldsymbol{y}\big\} = \big\{\boldsymbol{x}, \boldsymbol{y}\big\} \quad : \quad \text{all } \boldsymbol{x} \text{ and } \boldsymbol{y} \tag{16}$$

entails

$$\mathbb{T}^{\mathsf{T}}\mathbb{G}\mathbb{T} = \mathbb{G} \quad \text{i.e.,} \quad \mathbb{T}^{-1} = \mathbb{G}^{-1}\mathbb{T}^{\mathsf{T}}\mathbb{G} \tag{17}$$

Assume it possible to write

$$\mathbb{T} = e^{\mathbb{L}} \tag{18.1}$$

where $\mathbb{L}$ is the "logarithm" of $\mathbb{T}$, and subject to the proviso that

$$\mathbb{T} \to \mathbb{I} \quad \text{entails} \quad \mathbb{L} \to \mathbb{O} \tag{18.2}$$

Then

$$\det \mathbb{T} = e^{\text{trace}\,\mathbb{L}} \tag{19}$$

which serves to sharpen the condition $(\det \mathbb{T})^2 = 1$ implicit in (17); the same argument, run backwards, informs us that *in real theory* the condition

$$\text{trace}\,\mathbb{L} = 0 \tag{20}$$

is in fact universal: "improper" $\mathbb{T}$-matrices—those with $\det \mathbb{T} = -1$ (the only other possibility)—do not possess logarithms. From the multiplicative condition (17) we obtain the additive condition

$$\mathbb{G}\mathbb{L} + \mathbb{L}^{\mathsf{T}}\mathbb{G} = \mathbb{O} \tag{21}$$

Look now to the (spectrum-preserving) similarity transformation

$$\mathbb{M} \longrightarrow \mathbb{M}' \equiv \mathbb{T}^{-1}\mathbb{M}\mathbb{T} \tag{22}$$

By (17) we have

$$\mathbb{G}\mathbb{M}' = \mathbb{T}^{\mathsf{T}} \cdot \mathbb{G}\mathbb{M} \cdot \mathbb{T} \tag{23}$$

from which it becomes clear that

$$\text{if } \mathbb{G}\mathbb{M} \text{ is } \begin{Bmatrix} \text{symmetric} \\ \text{antisymmetric} \end{Bmatrix} \text{ then so also is } \mathbb{G}\mathbb{M}' \tag{24}$$

From this bland stock we can make several kinds of soup, depending upon what additional ingredients and seasoning we toss into the pot. Let us, before we proceed any further, agree in place of the generic $\{\bullet,\bullet\}$ to

$$\text{write}\begin{cases}(\bullet,\bullet)\text{ when }\mathbb{G}\text{ is symmetric:} & \mathbb{G}^{\mathsf{T}}=+\mathbb{G}\\ [\bullet,\bullet]\text{ when }\mathbb{G}\text{ is antisymmetric:} & \mathbb{G}^{\mathsf{T}}=-\mathbb{G}\end{cases}$$

and to restrict our attention to those two complementary cases. In place of (21) we then have this sharper statement:

$$\mathbb{GL}\text{ is }\begin{cases}\text{antisymmetric if }\mathbb{G}\text{ is symmetric}\\ \text{symmetric if }\mathbb{G}\text{ is antisymmetric}\end{cases}\tag{25}$$

Look to the simple Euclidean case

$$\mathbb{G}=\begin{pmatrix}1&0\\0&1\end{pmatrix}\quad:\quad\text{symmetric}$$

In that case $\mathbb{T}$ (called $\mathbb{R}$) by (17) satisfies $\mathbb{R}^{-1}=\mathbb{R}^{\mathsf{T}}$; it preserves the value of $(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{x}^{\mathsf{T}}\boldsymbol{y}$ and is called a "rotation matrix." Its logarithm $\mathbb{L}$ (called $\mathbb{A}$) is by (25) literally antisymmetric; writing

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

one by quick calculation obtains

$$\mathbb{R}=e^{\mathbb{A}}=\begin{pmatrix}\cos\theta&-\sin\theta\\ \sin\theta&\cos\theta\end{pmatrix}$$

I need not rehearse the familiar details, but do want to emphasize that it was precisely that body of detail—especially this instance[8] of (24):

$$\mathbb{R}^{-1}\cdot\text{symmetric}\cdot\mathbb{R}=\text{symmetric}$$

—that led us in §1 to Mohr's construction.

4. Mohr's construction in the 2-dimensional Lorentzian case. Look now to the (only slightly less familiar) Lorentzian case

$$\mathbb{G}=\begin{pmatrix}1&0\\0&-1\end{pmatrix}\quad:\quad\text{symmetric (but indefinite)}$$

---

[8] Curiously, we did not have occasion to draw upon the companion statement

$$\mathbb{R}^{-1}\cdot\text{antisymmetric}\cdot\mathbb{R}=\text{antisymmetric}$$

but for this there is a good explanation: the matrix on the right is the *same* matrix; nothing is going on; antisymmetric matrices are *invariant* with respect to rotational similarity transformation. A similar remark pertains to (27.2) below.

The antisymmetry of $\mathbb{GL}$ forces $\mathbb{L}$ itself to assume the form

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

and by a slight variant of the "quick calculation" to which I just referred[9] one obtains the "Lorentz matrix"

$$\mathbb{T} = \begin{pmatrix} \cosh\psi & \sinh\psi \\ \sinh\psi & \cosh\psi \end{pmatrix} \tag{26}$$

where in the relativistic application the parameter ψ—sometimes called the "rapidity"—acquires kinematic meaning from the equation $\tanh\psi = \beta \equiv v/c$. Equation (24) asserts that

$$\begin{aligned} (\text{lorentz})^{-1}\cdot \begin{pmatrix} a & c \\ -c & b \end{pmatrix} \cdot (\text{lorentz}) &= \text{matrix of that same structure} \qquad (27.1)\\ &= \begin{pmatrix} A & C \\ -C & B \end{pmatrix} \\ (\text{lorentz})^{-1}\cdot \begin{pmatrix} 0 & d \\ d & 0 \end{pmatrix} \cdot (\text{lorentz}) &= \text{unchanged}^{\,8} \qquad (27.2) \end{aligned}$$

where by computation we find

$$\begin{aligned} A &= a\cosh^2\psi - b\sinh^2\psi + 2c\,\cosh\psi\sinh\psi \\ B &= b\cosh^2\psi - a\sinh^2\psi - 2c\,\cosh\psi\sinh\psi \\ C &= (a-b)\cosh\psi\sinh\psi + c\,(\cosh^2\psi + \sinh^2\psi) \end{aligned}$$

which with the aid of some elementary identities

$$\begin{aligned} \cosh^2\psi &= \tfrac{1}{2}(\cosh 2\psi + 1) \\ \sinh^2\psi &= \tfrac{1}{2}(\cosh 2\psi - 1) \\ 2\cosh\psi\sinh\psi &= \sinh 2\psi \end{aligned}$$

become

$$\left.\begin{aligned} A &= \tfrac{a+b}{2} + \tfrac{a-b}{2}\cosh 2\psi + c\cdot\sinh 2\psi \\ B &= \tfrac{a+b}{2} - \tfrac{a-b}{2}\cosh 2\psi - c\cdot\sinh 2\psi \\ C &= \phantom{\tfrac{a+b}{2} + {}} \tfrac{a-b}{2}\sinh 2\psi + c\cdot\cosh 2\psi \end{aligned}\right\} \tag{28}$$

In the argument which led to (3) we drew tacitly upon the familiar fact that every real symmetric matrix $\left(\begin{smallmatrix} a & c \\ c & b \end{smallmatrix}\right)$ can be diagonalized by rotation. Equation (28) exposes a less familiar fact: the Lorentzian diagonalization of a matrix of the "Lorentz symmetric" structure $\left(\begin{smallmatrix} a & c \\ -c & b \end{smallmatrix}\right)$ entails $\tanh 2\psi = -\frac{2c}{a-b}$, and this

[9] See CLASSICAL ELECTRODYNAMICS (1980), p. 196.

can (with real ψ) be achieved if an only if $-1 \le -\frac{2c}{a-b} \le +1$. The latter condition can conveniently be formulated

$$\Delta \geq 0 \quad \text{with} \quad \Delta \equiv \big(\tfrac{a-b}{2}\big)^2 - c^2 \tag{29}$$

We note in this connection that the eigenvalues of $\left(\begin{smallmatrix} a & c \\ -c & b \end{smallmatrix}\right)$ can be described

$$\left.\begin{aligned} \lambda_1 &= \tfrac{a+b}{2} + \sqrt{\big(\tfrac{a-b}{2}\big)^2 - c^2} \quad &: \quad \text{becomes } a \text{ as } c^2 \downarrow 0 \\ \lambda_2 &= \tfrac{a+b}{2} - \sqrt{\big(\tfrac{a-b}{2}\big)^2 - c^2} \quad &: \quad \text{becomes } b \text{ as } c^2 \downarrow 0 \end{aligned}\right\} \tag{30}$$

and are
- imaginary if $\Delta < 0$;
- coincident if $\Delta = 0$;
- real and distinct if $\Delta > 1$.

I am concerned here with certain geometrical constructions, and it is hard to draw diagrams when things become imaginary, so it is to the class of cases (29) that I henceforth restrict my remarks. *Within such a restricted setting* it becomes possible to mimic the derivation of (3), writing

$$\mathbb{M}(\psi) \equiv (\text{lorentz})^{-1} \cdot \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \cdot (\text{lorentz}) = \begin{pmatrix} A & C \\ -C & B \end{pmatrix} \tag{31.1}$$

with

$$\left.\begin{aligned} A &= \tfrac{a+b}{2} + \tfrac{a-b}{2}\cosh 2\psi \\ B &= \tfrac{a+b}{2} - \tfrac{a-b}{2}\cosh 2\psi \\ C &= \qquad\;\; \tfrac{a-b}{2}\sinh 2\psi \end{aligned}\right\} \tag{31.2}$$

Borrowing terminology from relativity, I will say that a vector

$$\boldsymbol{x} \text{ is } \left\{\begin{matrix} \text{timelike} \\ \text{null} \\ \text{spacelike} \end{matrix}\right\} \text{ according as } (\boldsymbol{x}, \boldsymbol{x}) \equiv \boldsymbol{x}^{\mathsf{T}}\mathbb{G}\,\boldsymbol{x} \text{ is } \left\{\begin{matrix} > 0 \\ = 0 \\ < 0 \end{matrix}\right.$$

Less standardly, I will say that

$$\boldsymbol{x} \text{ is a } \left\{\begin{matrix} \text{timelike unit vector} \\ \text{spacelike unit vector} \end{matrix}\right\} \text{ if } \left\{\begin{matrix} (\boldsymbol{x}, \boldsymbol{x}) = +1 \\ (\boldsymbol{x}, \boldsymbol{x}) = -1 \end{matrix}\right.$$

Proceeding in imitation of Mohr, we (in a notation which has now to be considered vestigal) have

$$\begin{aligned} \boldsymbol{s}_1(\psi) = \begin{pmatrix} s_{11}(\psi) \\ s_{12}(\psi) \end{pmatrix} &\equiv \mathbb{M}(\psi)\hat{\boldsymbol{\omega}}_1 \quad \text{with} \quad \hat{\boldsymbol{\omega}}_1 \equiv \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)\text{: timelike unit vector} \\ &= \begin{pmatrix} A \\ -C \end{pmatrix} = \begin{pmatrix} \frac{a+b}{2} + \frac{a-b}{2}\cosh 2\psi \\ -\frac{a-b}{2}\sinh 2\psi \end{pmatrix} \end{aligned} \tag{32.1}$$

$$\begin{aligned} \boldsymbol{s}_2(\psi) = \begin{pmatrix} s_{21}(\psi) \\ s_{22}(\psi) \end{pmatrix} &\equiv \mathbb{M}(\psi)\hat{\boldsymbol{\omega}}_2 \quad \text{with} \quad \hat{\boldsymbol{\omega}}_2 \equiv \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)\text{: timelike unit vector} \\ &= \begin{pmatrix} C \\ B \end{pmatrix} = \begin{pmatrix} +\frac{a-b}{2}\sinh 2\psi \\ \frac{a+b}{2} - \frac{a-b}{2}\cosh 2\psi \end{pmatrix} \end{aligned} \tag{32.2}$$

where we have been obliged to write a *duplex* variant of (4) because the timelike and spacelike sectors of vector space are transformationally distinct.[10]

Eliminating ψ between the first pair of equations gives an equation of the form

$$\big(s_1 - \tfrac{a+b}{2}\big)^2 - s_2^2 = \Delta$$
$$\Delta \equiv \big(\tfrac{a-b}{2}\big)^2$$

while elimination between the latter pair gives

$$\big(s_2 - \tfrac{a+b}{2}\big)^2 - s_1^2 = \Delta$$

These equations describe a pair of hyperbolas, of which

- the first is centered at $(\frac{a+b}{2}, 0)$ and opens left/right;
- the second is centered at $(0, \frac{a+b}{2})$ and opens up/down;

But in each case, one branch is spurious, an artifact of the ψ-elimination procedure; the $\boldsymbol{s}_1(\psi)$ given by (31.1) glides along the right/left branch according as $\frac{a-b}{2} \gtrless 0$, and (when $\psi = 0$) crosses the axis at $s_1 = a$, while the $\boldsymbol{s}_2(\psi)$ given by (31.2) glides along the lower/upper branch according (again) as $\frac{a-b}{2} \gtrless 0$, and crosses the axis at $s_2 = b$. See Figure 6.

To what extent can the train of thought which flowed from (6) be carried over into the Lorentzian setting? Does "spectral resolution" remain available as a tool? If

$$\boldsymbol{x} = \begin{pmatrix} x^1 \\ x^2 \end{pmatrix}$$

then the "projector onto the $\boldsymbol{x}$-ray" can be described

$$\mathbb{P} = \tfrac{1}{(\boldsymbol{x},\boldsymbol{x})} \cdot \begin{pmatrix} x^1 x_1 & x^1 x_2 \\ x^2 x_1 & x^2 x_2 \end{pmatrix} \quad \text{where} \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \equiv \mathbb{G} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \begin{pmatrix} x^1 \\ -x^2 \end{pmatrix}$$

We have

$$\begin{pmatrix} \cosh\psi \\ \sinh\psi \end{pmatrix} \quad : \quad \text{representation of all "timelike unit vectors"}$$
$$\begin{pmatrix} \sinh\psi \\ \cosh\psi \end{pmatrix} \quad : \quad \text{representation of all "spacelike unit vectors"}$$

and notice that the former is orthogonal (Lorentzian sense) to the latter. The projector onto the timelike $\big(\begin{smallmatrix} \cosh\psi \\ \sinh\psi \end{smallmatrix}\big)$-ray can be described

$$\begin{aligned} \mathbb{P}_1 &= +\begin{pmatrix} \cosh^2\psi & -\cosh\psi\sinh\psi \\ \sinh\psi\cosh\psi & -\sinh^2\psi \end{pmatrix} \\ &= \tfrac{1}{2}\begin{pmatrix} 1+\cosh 2\psi & -\sinh 2\psi \\ +\sinh 2\psi & 1-\cosh 2\psi \end{pmatrix} \end{aligned}$$

---

[10] In the interest of expository clarity I have elected to omit discussion of the case $\boldsymbol{s}(\psi) = \mathbb{M}\boldsymbol{\omega}_{\text{null}}$.

Figure 6 goes here

FIGURE 6:

Figure 7 goes here

FIGURE 7:

while the projector onto the spacelike $\left(\begin{smallmatrix}\sinh\psi\\ \cosh\psi\end{smallmatrix}\right)$-ray becomes

$$\begin{aligned}\mathbb{P}_2 &= -\begin{pmatrix}\sinh^2\psi & -\sinh\psi\cosh\psi\\ \cosh\psi\sinh\psi & -\cosh^2\psi\end{pmatrix}\\ &= \tfrac{1}{2}\begin{pmatrix}1-\cosh 2\psi & +\sinh 2\psi\\ -\sinh 2\psi & 1+\cosh 2\psi\end{pmatrix}\end{aligned}$$

These matices (compare (7)) are demonstrably projective ($\mathbb{P}_1^2 = \mathbb{P}_1$, $\mathbb{P}_2^2 = \mathbb{P}_2$), orthogonal ($\mathbb{P}_1 \cdot \mathbb{P}_2 = \mathbb{O}$) and complementary ($\mathbb{P}_1 + \mathbb{P}_2 = \mathbb{I}$). And each possess the (additively closed) Lorentz-symmetric structure of $\left(\begin{smallmatrix}a & c\\ -c & b\end{smallmatrix}\right)$. Having established those elementary facts, I find it convenient in the application at hand[11] to

$$\text{flip from prograde to retrograde parameterization: } \psi \longrightarrow -\psi$$

i.e., to modify the definitions of $\mathbb{P}_1$ and $\mathbb{P}_2$, writing

$$\begin{aligned}\mathbb{P}_1 &\equiv \tfrac{1}{2}\begin{pmatrix}1+\cosh 2\psi & +\sinh 2\psi\\ -\sinh 2\psi & 1-\cosh 2\psi\end{pmatrix} &&: \quad \text{projects onto timelike } \left(\begin{smallmatrix}\cosh\psi\\ -\sinh\psi\end{smallmatrix}\right)\\ \mathbb{P}_2 &\equiv \tfrac{1}{2}\begin{pmatrix}1-\cosh 2\psi & -\sinh 2\psi\\ +\sinh 2\psi & 1+\cosh 2\psi\end{pmatrix} &&: \quad \text{projects onto spacelike } \left(\begin{smallmatrix}-\sinh\psi\\ \cosh\psi\end{smallmatrix}\right)\end{aligned}$$

Then (31) becomes

$$\mathbb{M}(\psi) = a\mathbb{P}_1 + b\mathbb{P}_2 \tag{33}$$

which precisely mimics (6). We observe that (compare (10)) $\mathbb{P}_1$ can be written

$$\mathbb{P}_1 = \tfrac{1}{2}\big\{\mathbb{I} + p^1\mathbb{S}_1 + p^3\mathbb{S}_3\big\} \tag{34}$$

with

$$\boldsymbol{p} \equiv \begin{pmatrix}p^1\\ p^3\end{pmatrix} \equiv \begin{pmatrix}\cosh 2\psi\\ \sinh 2\psi\end{pmatrix}, \quad \mathbb{S}_1 \equiv \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \quad \text{and} \quad \mathbb{S}_3 \equiv \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix} \tag{35}$$

and that $\boldsymbol{p} \longrightarrow -\boldsymbol{p}$ sends $\mathbb{P}_1 \longrightarrow \mathbb{P}_2$. The vector $\boldsymbol{p}$ ranges on the right branch (and its negative on the left branch) of a hyperbola which is centered at the origin of the (p^1, p^3)-plane, and which opens left/right; it is, in all cases (i.e., for all ψ), a timelike unit vector. See Figure 7.

It becomes natural to consider a "Lorentzian analog of the density matrix problem." Distribute weighted points on the "unit Mohr hyperbola," associate such a distribution with a weighted sum of projectors, and ask for the "spectral representation" of the resulting matrix:

$$\text{given } \mathbb{D} = \sum_k w_k\mathbb{P}_k, \text{ achieve the display } a\mathbb{P}_1 + b\mathbb{P}_2$$

[11] Recall from §1 that the Mohr construction proceeds from a presumption that "blobs *counter*rotate."

If all the points lie on the *same branch* of the unit Mohr hyperbola then (as is clear from the figure) the centroid of the distribution will lie necessarily in the timelike sector, and a Lorentzian analog of Wieting's construction becomes immediately available. But if points are distributed on both branches then the centroid *may lie in the spacelike sector*, and the possibility of spectral representation (at least along the lines of the present discussion, which presumes $\Delta \geq 0$) is lost. I lack physical motivation to pursue this matter in greater detail.