BULL. AUSTRAL. MATH. SOC. 47A60, 47B15
VOL. 55 (1997) [321–325]

# THE WEYL FUNCTIONAL CALCULUS AND TWO-BY-TWO SELFADJOINT MATRICES

WERNER J. RICKER

Let $D$ be a $(2 \times 2)$ matrix with distinct eigenvalues $\lambda_1$ and $\lambda_2$. There is a basic and well known functional equation which provides a formula for constructing the matrix $g(D)$, for any $\mathbb{C}$-valued function $g$ defined on a subset of $\mathbb{C}$ containing $\{\lambda_1, \lambda_2\}$, namely

$$g \mapsto g(D) = (\lambda_1 - \lambda_2)^{-1}\{g(\lambda_1) \cdot (D - \lambda_2 I) - g(\lambda_2) \cdot (D - \lambda_1 I)\}.$$

This equation is used to give a direct and transparent proof of the following fact due to Anderson: A pair of $(2 \times 2)$ selfadjoint matrices $A_1$ and $A_2$ commute if and only if the Weyl functional calculus of the pair $(A_1, A_2)$, which is a matrix-valued distribution, has order zero (that is, is a measure).

Given two selfadjoint matrices in $\mathcal{H} = \mathbb{C}^2$, say $A_1, A_2$, the Weyl calculus for the pair $A = (A_1, A_2)$ is an $L(\mathcal{H})$-valued distribution which is a particular rule allowing the construction of certain functions of the pair $(A_1, A_2)$. For $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$, the matrix $\langle \xi, A \rangle = \xi_1 A_1 + \xi_2 A_2$ is again selfadjoint and hence $\left\| e^{i\langle \xi, A \rangle} \right\| = 1$. Let $\mathcal{S}(\mathbb{R}^2)$ denote the Schwartz space of $\mathbb{C}$-valued, rapidly decreasing functions on $\mathbb{R}^2$. More precisely then, the Weyl calculus for $A$, [**1**, **6**, **7**], is the $L(\mathcal{H})$-valued distribution $T(A)$ defined by

$$T(A)f = \frac{1}{2\pi} \int_{\mathbb{R}^2} e^{i\langle \xi, A \rangle} \widehat{f}(\xi) d\xi, \qquad f \in \mathcal{S}(\mathbb{R}^2); \tag{1}$$

here $\widehat{f}$ denotes the Fourier transform of $f$ and $L(\mathcal{H})$ is the space of all $(2 \times 2)$ matrices over $\mathbb{C}$. The following result connects an analytic property of $T(A)$ with a purely algebraic property of $A$.

**THEOREM 1.** *Given a pair $A = (A_1, A_2)$ of selfadjoint matrices in $\mathcal{H} = \mathbb{C}^2$ the following statements are equivalent.*

(i) *The matrices $A_1$ and $A_2$ commute.*
(ii) *The associated Weyl calculus $T(A) : \mathcal{S}(\mathbb{R}^2) \to L(\mathcal{H})$ is a distribution of order zero.*

Received 4th June, 1996

There are several proofs of this theorem in the literature. The first proof given of this result is due to Anderson, [**2**, Theorem 2], and applies in $\mathbb{C}^m$, not just $\mathbb{C}^2$. It is based on properties of the numerical range and the theory of multivariable differential calculus. A completely different proof (also applying in $\mathbb{C}^m$), which is based on certain aspects of matrix-valued harmonic analysis in $L^p$-spaces (see [**3**]), is given in [**5**]. A third proof, specific to the case of $\mathbb{C}^2$, was given in [**4**]. This proof is essentially computational and is based on an elegant formula of Anderson, [**1**, Theorem 4.1], which expresses the Weyl calculus $T(J)$ of the triple $J = (J_1, J_2, J_3)$ whose entries are the classical spin 1/2-matrices in $L(\mathbb{C}^2)$, in terms of an integral formula over the unit sphere $S^2$ (in $\mathbb{R}^3$) with respect to normalised surface measure $\mu$.

The aim of this note is to present another proof of Theorem 1. The proof is again computational in nature, but has the advantage over [**4**] in that it is based on a much more elementary and very well known functional equation. Namely, for a $(2 \times 2)$-matrix $D$ with distinct eigenvalues $\lambda_1$ and $\lambda_2$ and any $\mathbb{C}$-valued function $g$ defined on a subset of $\mathbb{C}$ containing $\sigma(D) = \{\lambda_1, \lambda_2\}$, the matrix $g(D)$ is given by the formula

$$(2) \qquad g(D) = \frac{g(\lambda_1)}{(\lambda_1 - \lambda_2)} \cdot (D - \lambda_2 I) - \frac{g(\lambda_2)}{(\lambda_1 - \lambda_2)} \cdot (D - \lambda_1 I).$$

In particular, the proof given below provides an interesting and non-trivial application of (2).

To establish (i) $\Rightarrow$ (ii) is elementary and can be found in [**4**], for example. So let $A_1$ and $A_2$ be selfadjoint matrices in $L(\mathcal{H})$ which do *not* commute. To establish (ii) $\Rightarrow$ (i) it is to be shown that the distribution $T(A) : \mathcal{S}(\mathbb{R}^2) \to L(\mathcal{H})$ has positive order. If $U$ is any orthogonal $(2 \times 2)$-matrix, define $UAU^{-1} = (UA_1U^{-1}, UA_2U^{-1})$. Then $T(UAU^{-1})f = U(T(A)f)U^{-1}$, for every $f \in \mathcal{S}(\mathbb{R}^2)$, [**1**, Theorem 2.9(e)]. So, choose for $U$ an orthogonal transformation such that the matrix $B_1$ of $UA_1U^{-1}$ with respect to the basis of $\mathcal{H}$ consisting of the orthonormal eigenvectors of $A_1$ is diagonal, say $\begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}$. Then the matrix $B_2$ of $UA_2U^{-1}$ with respect to this basis is of the form $\begin{pmatrix} \beta_1 & w \\ \overline{w} & \beta_2 \end{pmatrix}$ for some $w \in \mathbb{C}$ and $\beta_1, \beta_2 \in \mathbb{R}$. Since $A_1A_2 \neq A_2A_1$ it follows that $B_1B_2 \neq B_2B_1$ and moreover, that $\alpha_1 \neq \alpha_2$ (with $\alpha_1, \alpha_2 \in \mathbb{R}$) and $w \neq 0$. Since the order of the distribution $T(B)$ is the same as that of $T(A)$ it suffices to show that $T(B)$ has positive order.

Fix $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$. For each $\lambda \in \mathbb{C}$, it follows that

$$(3) \qquad \det(\lambda I - \langle \xi, B \rangle) = \lambda^2 - (\xi_1\alpha_2 + \xi_2\beta_2 + \xi_1\alpha_1 + \xi_2\beta_1)\lambda + (\xi_1\alpha_1 + \xi_2\beta_1).(\xi_1\alpha_2 + \xi_2\beta_2) - |w|^2 \xi_2^2.$$

https://doi.org/10.1017/S0004972700033980 Published online by Cambridge University Press

Let $h = (\alpha_1 - \alpha_2)/2$ and $k = (\beta_1 - \beta_2)/2$, in which case $k \in \mathbb{R}$ and $h \in \mathbb{R}\backslash\{0\}$. Direct calculation shows that the solutions of (3) are given by

$$\lambda(\xi) = \frac{1}{2}(\xi_1[\alpha_1 + \alpha_2] + \xi_2[\beta_1 + \beta_2]) \pm \{\Delta(\xi)\}^{1/2}, \tag{4}$$

where $\Delta(\xi) = (h\xi_1 + k\xi_2)^2 + |w|^2 \xi_2^2$. Since $\Delta(\xi) = 0$ if and only if $\xi = 0$, it follows from (4) that $\langle \xi, B\rangle$ has two distinct eigenvalues, say $\lambda_1(\xi)$ and $\lambda_2(\xi)$, whenever $\xi \neq 0$. The identity (2), with $D = \langle \xi, B\rangle$ and $g(z) = e^{iz}$, implies that

$$e^{i\langle \xi, B\rangle} = \frac{e^{i\lambda_1(\xi)}}{(\lambda_1(\xi) - \lambda_2(\xi))}.(\langle \xi, B\rangle - \lambda_2(\xi)I) - \frac{e^{i\lambda_2(\xi)}}{(\lambda_1(\xi) - \lambda_2(\xi))}.(\langle \xi, B\rangle - \lambda_1(\xi)I), \tag{5}$$

for every $\xi \neq 0$. Of course, $e^{i\langle 0, B\rangle} = I$. Substituting (5) into (1), with $A$ replaced by $B$, shows that the (1,2)-entry of the matrix $T(B)f$ is given by

$$L(f) = \frac{w}{2\pi}\int_{\mathbb{R}^2} \frac{\left(e^{i\lambda_1(\xi)} - e^{i\lambda_2(\xi)}\right)\xi_2 \widehat{f}(\xi)}{(\lambda_1(\xi) - \lambda_2(\xi))}\, d\xi, \qquad f \in \mathcal{S}(\mathbb{R}^2). \tag{6}$$

If $\lambda_1(\xi)$ denotes the eigenvalue of $\langle \xi, B\rangle$ corresponding to the $+$ sign in front of $\{\Delta(\xi)\}^{1/2}$ in (4), then it follows from (4) that (6) simplifies to

$$L(f) = \frac{iw}{2\pi}\int_{\mathbb{R}^2} \frac{\xi_2 e^{i\langle \xi, u\rangle}\widehat{f}(\xi)\sin\{\Delta(\xi)\}^{1/2}}{\{\Delta(\xi)\}^{1/2}}\, d\xi, \qquad f \in \mathcal{S}(\mathbb{R}^2),$$

where $u = (\alpha_1 + \alpha_2, \beta_1 + \beta_2)/2$. If $f_u(\eta) = f(u + \eta)$, for $\eta \in \mathbb{R}^2$, then $\widehat{f_u}(\xi) = e^{i\langle \xi, u\rangle}\widehat{f}(\xi)$ and so

$$L(f) = \frac{iw}{2\pi}\int_{\mathbb{R}^2} \frac{\xi_2 \widehat{f_u}(\xi)\sin\{\Delta(\xi)\}^{1/2}}{\{\Delta(\xi)\}^{1/2}}\, d\xi = \frac{w}{2\pi}\int_{\mathbb{R}^2} \frac{(D_2 f_u)\widehat{\ }(\xi)\sin\{\Delta(\xi)\}^{1/2}}{\{\Delta(\xi)\}^{1/2}}\, d\xi,$$

where $D_2$ denotes differentiation with respect to the second variable. By making the linear change of variables in $\mathbb{R}^2$ given by $y = M\xi$, where $M = \begin{pmatrix} h & k \\ 0 & |w| \end{pmatrix}$ and elements of $\mathbb{R}^2$ are interpreted as column vectors, it follows that

$$L(f) = \frac{h\,|w|\,w}{2\pi}\int_{\mathbb{R}^2} \frac{(D_2 f_u)\widehat{\ }(M^{-1}y)\sin\left(y_1^2 + y_2^2\right)^{1/2}}{\left(y_1^2 + y_2^2\right)^{1/2}}\, dy, \qquad f \in \mathcal{S}(\mathbb{R}^2). \tag{7}$$

The Fourier-Stieltjes transform $\widehat{\mu}$ of the measure $\mu$ (recall that $\mathrm{supp}(\mu) = S^2 \subseteq \mathbb{R}^3$) is easily computed via spherical polar coordinates and is given by

$$\widehat{\mu}(\gamma) = \frac{\sin\left(\gamma_1^2 + \gamma_2^2 + \gamma_3^2\right)^{1/2}}{(2\pi)^{3/2}\left(\gamma_1^2 + \gamma_2^2 + \gamma_3^2\right)^{1/2}}, \qquad \gamma \in \mathbb{R}^3\backslash\{0\},$$

with $\widehat{\mu}(0) = (2\pi)^{-3/2}$. Let $\mathbb{D}$ be the closed unit disc in $\mathbb{R}^2$. Define a measure $\nu$ on the Borel subsets $\mathcal{B}(\mathbb{R}^2)$ of $\mathbb{R}^2$ by

$$\nu(E) = \mu((E \cap \mathbb{D}) \times \mathbb{R}), \qquad E \in \mathcal{B}(\mathbb{R}^2).$$

Given a function $\varphi : \mathbb{R}^2 \to \mathbb{C}$, let $\widetilde{\varphi} : \mathbb{R}^3 \to \mathbb{C}$ be the function defined by $\widetilde{\varphi}(x,y,z) = \varphi(x,y)$. A routine calculation shows that $\int_{\mathbb{R}^2} s\, d\nu = \int_{\mathbb{R}^3} \widetilde{s}\, d\mu$, for every $\mathcal{B}(\mathbb{R}^2)$-simple function $s : \mathbb{R}^2 \to \mathbb{C}$. It follows from the dominated convergence theorem that $\int_{\mathbb{R}^2} \varphi\, d\nu = \int_{\mathbb{R}^3} \widetilde{\varphi}\, d\mu$ for every bounded Borel function $\varphi : \mathbb{R}^2 \to \mathbb{C}$. In particular, putting $\varphi_\xi(x) = e^{i\langle \xi, x\rangle}$, for each fixed $\xi \in \mathbb{R}^2$, it follows that $\widehat{\nu}(\xi) = \widehat{\mu}(\xi, 0)$. That is,

$$\widehat{\nu}(\xi) = \frac{\sin\left(\xi_1^2 + \xi_2^2\right)^{1/2}}{(2\pi)^{3/2}(\xi_1^2 + \xi_2^2)^{1/2}}, \qquad \xi \in \mathbb{R}^2 \backslash \{0\},$$

with $\widehat{\nu}(0) = (2\pi)^{-3/2}$.

Now, the function $\Phi = \widehat{\nu}$ is locally integrable (as it is a continuous function vanishing at $\infty$) and hence, can be interpreted as a distribution in the usual way, that is, $\langle g, \Phi\rangle = \int_{\mathbb{R}^2} g(\xi)\Phi(\xi)\, d\xi$, for $g \in \mathcal{S}(\mathbb{R}^2)$. Accordingly, the distributional Fourier transform $\widehat{\Phi}$ of $\Phi$ is given by

$$\langle p, \widehat{\Phi}\rangle = \langle \widehat{p}, \Phi\rangle = \int_{\mathbb{R}^2} \widehat{\nu}(\xi)\widehat{p}(\xi)\, d\xi = \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}^2} e^{i\langle \xi, x\rangle} d\nu(x)\right) \widehat{p}(\xi)\, d\xi,$$

for each $p \in \mathcal{S}(\mathbb{R}^2)$. Applying Fubini's theorem and the Fourier inversion formula $\int_{\mathbb{R}^2} e^{i\langle \xi, x\rangle}\widehat{p}(\xi)\, d\xi = 2\pi p(x)$ shows that

$$\langle p, \widehat{\Phi}\rangle = 2\pi \int_{\mathbb{R}^2} p(x)\, d\nu(x), \qquad p \in \mathcal{S}(\mathbb{R}^2). \tag{8}$$

Accordingly, the Fourier transform of $\Phi$ is the measure $2\pi\nu$ (acting on $\mathcal{S}(\mathbb{R}^2)$ via the right-hand-side of (8)).

For $g \in \mathcal{S}(\mathbb{R}^2)$, let $g \circ M^t \in \mathcal{S}(\mathbb{R}^2)$ denote the function $x \mapsto g(M^t x)$, for each $x \in \mathbb{R}^2$, where $M^t$ is the transpose of the matrix $M$. Direct calculation shows that

$$(D_2 f_u)\widehat{\ }(M^{-1}y) = \frac{1}{h\,|w|}\cdot(D_2 f_u \circ M^t)\widehat{\ }(y), \qquad y \in \mathbb{R}^2,$$

for each $f \in \mathcal{S}(\mathbb{R}^2)$. It follows from (7), (8) and the definition of distributional Fourier transforms that

$$L(f) = w(2\pi)^{3/2} \int_{\mathbb{R}^2} (D_2 f_u \circ M^t)(x)\, d\nu(x), \qquad f \in \mathcal{S}(\mathbb{R}^2). \tag{9}$$

https://doi.org/10.1017/S0004972700033980 Published online by Cambridge University Press

Since $f \mapsto f_u$ and $f \mapsto f \circ M^t$ are bicontinuous isomorphisms of $\mathcal{S}(\mathbb{R}^2)$ onto itself, it is clear from (9) that the distribution $L(f)$ has positive order. Since $L$ is the (1,2)-entry of $T(B)f$, for each $f \in \mathcal{S}(\mathbb{R}^2)$, it follows that $T(B)$ also has positive order. □

The identity (9) shows that the support of $L$ is a translate of the image of $\mathbb{D}$ under a non-singular transformation in $\mathbb{R}^2$ (with positive determinant). In particular, $\operatorname{supp}(L)$ is an infinite subset of $\mathbb{R}^2$. Since $\operatorname{supp}(L) \subseteq \operatorname{supp}(T(B)) = \operatorname{supp}(T(A))$ we have also given an alternative proof of the fact that $A_1A_2 = A_2A_1$ if and only if $\operatorname{supp}(T(A))$ is a finite subset of $\mathbb{R}^2$, [**4, 5**].

## References

[1] R.F.V. Anderson, 'The Weyl functional calculus', *J. Funct. Anal.* **4** (1969), 240–267.

[2] R.F.V. Anderson, 'On the Weyl functional calculus', *J. Funct. Anal.* **6** (1970), 110-115.

[3] P. Brenner, 'The Cauchy problem for symmetric hyperbolic systems in $L_p$', *Math. Scand.* **19** (1966), 27–37.

[4] B.R.F. Jefferies and W.J. Ricker, 'Commutativity for systems of $(2 \times 2)$ selfadjoint matrices', *Linear and Multilinear Algebra* **35** (1993), 107–114.

[5] W.J. Ricker, 'The Weyl calculus and commutativity for systems of selfadjoint matrices', *Arch. Math.* **61** (1993), 173–176.

[6] M.E. Taylor, 'Functions of several selfadjoint operators', *Proc. Amer. Math. Soc.* **19** (1968), 91–98.

[7] H. Weyl, *The theory of groups and quantum mechanics* (Dover Publ., New York, 1950).

School of Mathematics
The University of New South Wales
Sydney NSW 2052
Australia