Zeitschrift: Helvetica Physica Acta
Band: 45 (1972)
Heft: 2

Artikel: Quaternion determinants
Autor: Dyson, Freeman J.
DOI: https://doi.org/10.5169/seals-114385

Helvetica Physica Acta
Vol. 45, 1972.

Birkhäuser Verlag Basel

Quaternion Determinants[1])

by **Freeman J. Dyson**

The Institute for Advanced Study, Princeton, New Jersey 08540

(30. IX. 71)

Abstract. Many years ago, E. H. Moore proposed a definition for a determinant with non-commuting elements. It applies in particular to determinants with quaternion elements. The definition is here generalized and its properties studied in detail. The determinant is a multilinear polynomial in its elements. In a wide class of cases, it reduces to a Pfaffian. It possesses the property of multiplicativity only to a limited extent. It finds an application in the statistical theory of energy-levels of complex systems represented by an ensemble of random matrices.

'We find therefore, that in Equations, whether Lateral or Quadratick, which in the strict Sense, and first Prospect, appear Impossible; some mitigation may be allowed to make them Possible; and in such a mitigated interpretation they may yet be useful.'

John Wallis, *Treatise of Algebra* (London 1685), p. 272, quoted by William Hamilton in the Preface to his *Lectures on Quaternions* (Dublin 1853), p. 34.

«Vier Elemente
Innig gesellt
Bilden das Leben,
Bauen die Welt.»

Friedrich Schiller, quoted by Markus Fierz in his essay *Die Vier Elemente* (Schweizerische Gesellschaft für Analytische Psychologie, 1963).

I. History

Hamilton never mentioned determinants in his massive books [1, 2] on the theory of quaternions, although Cayley had published a paper [3] on quaternion-determinants many years earlier. In the second edition of Hamilton's work, published long after his death, the editor added a brief appendix [4] on determinants with a reference to Cayley. Hamilton expounded with loving care and in minutest detail every ramification of the quaternion algebra except this one. Why was he silent on the subject of quaternion determinants?

The reason for Hamilton's silence is presumably that he believed a satisfactory theory of quaternion-determinants to be impossible. In modern language, one can construct a simple proof of the impossibility along the following lines. Let R be a ring, let A be the ring of square matrices ($n \times n$) with elements in R, and let W be the set of single-column matrices with elements in R. By a 'determinant' we mean a mapping D from A into R which satisfies the following three axioms.

Axiom 1. *For any a in A, $D(a) = 0$ if and only if there is a non-zero w in W with $aw = 0$.*

---

[1]) Dedicated to Markus Fierz on the occasion of his sixtieth birthday.

Axiom 2. $D(a)\,D(b) = D(ab)$. (1.1)

Axiom 3. *Let the elements of a be* a_{ij}, $i, j = 1, \ldots, n$, *and similarly for b and c. If for some row-index k we have*

$$a_{ij} = b_{ij} = c_{ij}, \quad i \neq k, \tag{1.2}$$

$$a_{ij} + b_{ij} = c_{ij}, \quad i = k, \tag{1.3}$$

then

$$D(a) + D(b) = D(c). \tag{1.4}$$

Of these axioms, only Axiom 1 is indispensable for the utility of the notion of a determinant. But both 2 and 3 are desirable if one wishes to preserve the customary rules for manipulation and computation of determinants. The difficulty of keeping all three axioms is shown by

Theorem 1. *Let R be a ring with a unit element and without divisors of zero. If, on the matrix ring A with $n > 1$, a mapping D exists satisfying Axioms 1–3, then R is commutative.*

This means that quaternion-determinants cannot be constructed unless one is prepared to abandon one or more of the axioms.

Proof of Theorem 1. To save writing, we take $n = 2$. The proof for $n > 2$ goes in exactly the same way. For any r in R, we write

$$f(r) = D\left(\begin{bmatrix} 1 & 0 \\ 0 & r \end{bmatrix}\right), \tag{1.5}$$

where 1 means the unit element in R. Since R has no zero-divisors, Axioms 1, 2, 3 together imply that the mapping $r \to f(r)$ is an isomorphism of R into a sub-ring $R' \subseteq R$. Next, Axiom 3 gives

$$D\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right) = D\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) + D\left(\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}\right) + D\left(\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}\right) + D\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right),$$

and by Axiom 1 this reduces to

$$D\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right) + 1 = 0, \tag{1.6}$$

or

$$D\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right) = -1. \tag{1.7}$$

Hence, by Axiom 2

$$D\left(\begin{bmatrix} r & 0 \\ 0 & 1 \end{bmatrix}\right) = (-1) \cdot f(r) \cdot (-1) = f(r). \tag{1.8}$$

Finally by Axiom 2 again, (1.5) and (1.8) imply

$$D\left(\begin{bmatrix} r & 0 \\ 0 & s \end{bmatrix}\right) = f(r)f(s) = f(s)f(r). \tag{1.9}$$

But $f(r)$ and $f(s)$ may be any two elements of R'. Hence R' is commutative. But R is isomorphic to R' and is therefore also commutative. End of proof.

The subject of quaternion-determinants entered a long sleep, from the time of Cayley's 1844 paper until Study [5] attacked it again in 1920. Study constructed a determinant which he denoted by $\nabla(a)$, where a is a matrix over R and R is the ring of quaternions with real or complex coefficients. The values of $\nabla(a)$ are real or complex numbers as the case may be, not quaternions. Study proved that his ∇ satisfies Axioms 1 and 2. Axiom 3 is not satisfied. In fact, Axiom 3 states that the determinant is a linear function of the elements in each row of the matrix, whereas Study's ∇ is explicitly given as a quadratic function of the elements. Without Axiom 3 there is no convenient way to compute the value of a determinant of large order. Study's definition, while solving in principle the problem of the consistency of sets of linear quaternionic equations, has never found any practical application.

Independently of Study, E. H. Moore [6, 7] worked out a theory of determinants whose elements belong to a quasi-field (i.e., a non-commutative ring in which every element has an inverse). During his lifetime Moore published only a brief announcement [6] of his work. His notes were edited and published [7] after his death under the title 'General Analysis'. By 'General Analysis' Moore meant a redevelopment of large parts of classical mathematics using more general concepts than the classical system of real and complex numbers. Trying to carry out this grandiose project single-handed, Moore was overtaken by younger men, many of them his own pupils, who were generalizing mathematics in far more radical ways. When his work finally appeared in print [7], encumbered with an elaborate symbolic-logic notation, it looked like a throwback to the nineteenth century rather than the wave of the future. The Bourbaki group later adopted Moore's program and succeeded where he had failed. While Bourbaki set the style for the mathematics of the second half of the twentieth century, Moore's pioneering efforts were buried in books which the new generation found antiquated and irrelevant.

The theory of determinants which we shall describe in this paper is essentially due to Moore. We have only translated his definitions into modern notation and generalized them where appropriate. The elements out of which the determinants are constructed will belong to a ring R possessing an involution. This general frame can then be successively specialized according to the scheme: Ring-with-involution $\rightarrow$ Composition algebra $\rightarrow$ Quasi-field $\rightarrow$ Quaternion algebra. We find that some properties of the determinant hold for a general ring-with-involution while others require more special assumptions.

A more abstract approach to the construction of determinants with non-commuting elements was later worked out by Dieudonné [8] and is summarized in a book by Artin [9]. The Dieudonné determinant det (a) is defined only when R is a division ring (i.e., when every non-zero element of R has an inverse). The case when R consists of quaternions with complex coefficients, which caused no difficulty to Study, cannot be handled by Dieudonné. When R is a division ring, det (a) always exists and satisfies Axioms 1 and 2. Axiom 3 has no meaning because the values of det (a) do not lie in R. The values of det (a) lie in a semigroup $\bar{R}$ which is defined by adjoining a zero element to the quotient of R (regarded as a multiplicative group) by its commutator subgroup. For example, when R is the division ring of real quaternions, the quotient is the group of positive real numbers, and $\bar{R}$ is the semigroup of non-negative real numbers. Instead of Axiom 3, det (a) satisfies only the much weaker statement that equations (1.2) and (1.3) imply

$$\det(c) = \det(a) \cdot u + \det(b) \cdot v, \tag{1.10}$$

where u and v are elements of the commutator subgroup of R. When R is commutative equation (1.10) reduces to (1.4), and by Theorem 1 Axiom 3 can never hold otherwise.

II. Determinants of Self-Adjoint Matrices

In a recent investigation of the correlations between eigenvalues of random matrices [10, 11], it was found unexpectedly that the correlation-functions could be conveniently expressed as determinants of quaternion matrices, constructed according to the rules laid down by Moore [7]. These determinants avoid the awkward features that made Study's and Dieudonné's determinants unpractical. They have most of the desirable properties possessed by determinants of commuting elements. Why should Moore's definition of a determinant be superior to the others in this application? The reason is that, quite fortuitously, all the matrices for which determinants were required were self-adjoint in the quaternionic sense. By a self-adjoint matrix we mean one whose elements satisfy the conditions

$$a_{ij} = a_{ji}^{\dagger}, \tag{2.1}$$

where $q^{\dagger}$ denotes the quaternion adjoint to q.

To give the discussion greater generality, we suppose that R is not necessarily a quaternion algebra but a general ring with an involution. That is to say, we assume that every element q in R has a conjugate $q^{\dagger}$ with the properties

$$q^{\dagger\dagger} = q, \tag{2.2}$$

$$(q + r)^{\dagger} = q^{\dagger} + r^{\dagger}, \tag{2.3}$$

$$(qr)^{\dagger} = r^{\dagger} q^{\dagger}. \tag{2.4}$$

Elements with $q = q^{\dagger}$ are called scalars. We assume that R is a *ring with commuting scalars*, that is to say, we assume that every scalar in R commutes with all elements in R. In particular, the scalars form a commutative ring $S \subseteq R$. When R is the ring of real or complex quaternions, S is the ring of real or complex numbers. Finally, we assume that R has also the *scalar product property*. By this we mean that the scalar product

$$(q, r) = qr + r^{\dagger} q^{\dagger} \tag{2.5}$$

is symmetric, that is to say

$$(q, r) = (r, q) \tag{2.6}$$

for all q, r in R. The general symmetry (2.6) is a consequence of the special case

$$qq^{\dagger} = q^{\dagger} q. \tag{2.7}$$

If R is a quasi-field as Moore [7] assumed, the scalar product property follows from the commuting-scalar property.

The real and complex quaternions are examples of rings with commuting scalars and the scalar product property. Another example is the ring R whose elements are (2×2) matrices with conjugation defined by

$$\begin{bmatrix} x & y \\ z & w \end{bmatrix}^{\dagger} = \begin{bmatrix} w & -y \\ -z & x \end{bmatrix}. \tag{2.8}$$

When x, y, z, w are complex numbers, this R is equivalent to the ring of complex quaternions. But when x, y, z, w are real numbers, R is distinct from both the real and the complex quaternion ring.

We do not know whether the class of rings with commuting scalars and the scalar product property is exhausted by the rings of quaternions and (2×2) matrices with coefficients in suitable commutative rings. It would be interesting either to find other examples or to prove that none exist. A complete classification theory has been worked out only for a more restricted class of objects called *Composition Algebras*. We say that R is a composition algebra if 1. R is a ring with an involution, 2. R has commuting scalars and the scalar product property, 3. the scalar product is non-degenerate (i.e., $(q,r) = 0$ for all r in R implies $q = 0$), and 4. the ring S of scalars is a field (i.e., every non-zero element of S has an inverse). Jacobson [12, 13] has classified these objects without assuming the associative law of multiplication in R. Since we are only interested in associative algebras, we may state Jacobson's main result [14] as follows: Every (associative) composition algebra is either commutative or is a quaternion or (2×2) matrix algebra over the field S. So, to find new examples of rings satisfying our postulates, we must look outside the class of composition algebras.

It turns out that Moore's definition of determinants applies not only to self-adjoint matrices satisfying equation (2.1) but also to matrices which are self-adjoint except for a single row or column. A matrix a with elements in R is defined to be *almost self-adjoint* [15] if there is an integer k such that

$$a_{ij} = a^{\dagger}_{ji} \quad \text{when} \quad i \neq k, j \neq k. \tag{2.9}$$

The definition implies that a self-adjoint matrix is also almost self-adjoint. We define the determinant $Q\det(a)$ of an almost self-adjoint $(n \times n)$ matrix a by induction on n. Namely, for $n = 1$,

$$Q\det(a) = a_{11}, \tag{2.10}$$

and for $n > 1$

$$Q\det(a) = \sum_{l=1}^{n} \epsilon_{kl}\, a_{kl}\, Q\det(a(k,l)). \tag{2.11}$$

Here k is the integer singled out by equation (2.9),

$$\epsilon_{kl} = -1, \quad l \neq k; \quad \epsilon_{kk} = +1, \tag{2.12}$$

and $a(k,l)$ is the matrix obtained from a by first replacing the elements of the lth column by the corresponding elements of the kth column and then deleting both the kth row and the kth column. This recipe agrees with the usual definition of a determinant if R is commutative. It works as an inductive definition because $a(k,l)$ is an almost self-adjoint $(n-1) \times (n-1)$ matrix.

When a is almost self-adjoint but not self-adjoint, the definition (2.11) is unambiguous, since k is uniquely determined by equation (2.9). However, when a is self-adjoint, any value of k may be used in equation (2.11). To make the definition unique, we must prove

Lemma 1. *Let R be a ring with commuting scalars and the scalar product property, and let a be a self-adjoint matrix with elements in R. Then the value of $Q\det(a)$ defined by Equation (2.11) is independent of k. Also, $Q\det(a)$ is a scalar.*

Proof of Lemma 1. When a is any almost self-adjoint matrix, we define the *cycle-sum* $C(a)$ by an induction similar to equation (2.11). For $n = 1$,

$$C(a) = a_{11}, \tag{2.13}$$

and for $n > 1$

$$C(a) = \sum_{l \neq k} a_{kl}\, C(a(k,l)), \tag{2.14}$$

with $a(k,l)$ defined as before. By iterating the definition (2.14) $(n-1)$ times, we arrive at the result

$$C(a) = \sum_{P} a_{kl}\, a_{lm} \dots a_{uk}, \tag{2.15}$$

where P is summed over the $(n-1)!$ cyclic permutations

$$P = (k \to l \to m \to \cdots \to u \to k) \tag{2.16}$$

of the integers $(1, \dots, n)$. At the various stages of the iteration of equation (2.14), the column which was originally in position k is moved successively to the positions l, m, $\dots$, u. The *column-index* k can never reappear at any stage of the process until the last, when equation (2.13) is used. At the final stage, the single element a_{uk} survives in row u and column u. If a is not self-adjoint, equation (2.15) defines $C(a)$ uniquely. The first step in the proof of Lemma 1 is to show that for *self-adjoint* a the value of $C(a)$ is independent of k.

Let a be self-adjoint, and let $C'(a)$ be the sum (2.15) with some other integer j replacing k. The same cyclic permutations P appear in $C'(a)$ as in $C(a)$. Any term in $C(a)$ is of the form qr, with

$$q = a_{kl}\, a_{lm} \dots a_{tj}, \quad r = a_{js} \dots a_{uk}. \tag{2.17}$$

The term in $C'(a)$ corresponding to the same P is then rq. Now consider the permutation $\bar{P}$ obtained from P by reversing the cyclic order. When $n > 2$ $\bar{P}$ is distinct from P. In $C(a)$ the term arising from $\bar{P}$ is

$$a_{ku} \dots a_{ml}\, a_{lk} = (qr)^{\dagger} = r^{\dagger} q^{\dagger} \tag{2.18}$$

by virtue of equation (2.1). In $C'(a)$, the term arising from $\bar{P}$ is $q^{\dagger} r^{\dagger}$. Thus, P and $\bar{P}$ together contribute to $C(a)$ the scalar product (q, r), and to $C'(a)$ the scalar product (r, q). The scalar product property (2.6) then implies

$$C(a) = C'(a). \tag{2.19}$$

When $n = 2$ the same result follows trivially from equation (2.7). So we have proved that $C(a)$ is independent of k, and also that $C(a)$ is a scalar.

Now let G denote any division of the integers $(1, \dots, n)$ into m subsets $(G_1, \dots, G_m)$. If a is any almost self-adjoint matrix, we denote by $a(G_r)$ the submatrix of a formed by the elements a_{ij} with both i and j in G_r. Without loss of generality, we suppose that G_1

contains the index k defined by equation (2.9). Then $a(G_1)$ is almost self-adjoint and all the other $a(G_r)$ are self-adjoint. We now show that the definition (2.11) implies [16]

$$Q\det(a) = \sum_G (-1)^{n-m} \prod_{r=1}^{m} C(a(G_r)). \tag{2.20}$$

Since all factors but one in the product on the right of equation (2.20) are scalars, and since R is a ring with commuting scalars, the order of factors in the product is immaterial.

The proof of equation (2.20) goes by induction. For $n = 1$, equation (2.20) holds by virtue of equation (2.10) and (2.13). For $n > 1$, we suppose that equation (2.20) holds for the matrices $a(k, l)$ which have $(n - 1)$ rows and columns. The sum (2.20) can be divided into two parts, $\sum_1$ containing terms for which the set G_1 consists of the integer k alone, and $\sum_2$ containing terms for which G_1 contains other integers besides k. In $\sum_1$ we have by equation (2.13).

$$C(a(G_1)) = a_{kk}. \tag{2.21}$$

Since equation (2.20) holds for $a(k, k)$, $\sum_1$ is equal to the term with $l = k$ in equation (2.11). In $\sum_2$ we expand the factor $C(a(G_1))$ in each term using equation (2.14), and pick out the sum $\sum_{2l}$ of contributions arising from a particular value of l. The G which appear in $\sum_{2l}$ are those which have the integers k and l together in G_1, and these G correspond in one-to-one fashion with the divisions G of the integers $(1, \ldots, n)$ omitting k which appear when $Q\det(a(k, l))$ is expanded according to equation (2.20). Since equation (2.20) holds for $a(k, l)$, the sum $\sum_{2l}$ reproduces exactly the term in l on the right of equation (2.11). The signs agree by virtue of equation (2.12). Putting together $\sum_1$ and the $\sum_{2l}$, the induction is complete and equation (2.20) is proved for all almost self-adjoint a.

When a is self-adjoint, the proof of equation (2.20) is still valid. In this case, every $a(G_r)$ is self-adjoint, and we have already proved that each $C(a(G_r))$ is then a scalar independent of the choice of k in equation (2.14). Hence equation (2.20) defines $Q\det(a)$ as a scalar independent of the choice of k in equation (2.11). End of proof of Lemma 1.

For any matrix a over R, we define the adjoint matrix $a^\dagger$ by

$$(a^\dagger)_{ij} = (a_{ji})^\dagger. \tag{2.22}$$

If a is almost self-adjoint, $a^\dagger$ is also, and equation (2.15) implies

$$C(a^\dagger) = (C(a))^\dagger. \tag{2.23}$$

Hence, equation (2.20) gives

$$Q\det(a^\dagger) = (Q\det(a))^\dagger. \tag{2.24}$$

It follows from equation (2.24) that $Q\det(a)$ would have the same value if it were defined with an expansion by columns multiplying from the right instead of by rows multiplying from the left.

Another simple consequence of equation (2.20) is

Lemma 2. *If a is almost self-adjoint and has two identical rows (or columns) then $Q\det(a) = 0$.*

Proof of Lemma 2. Let the two identical rows have indices s and t. Consider the cycle-sum $C(a)$ defined by equation (2.15). Each cyclic permutation (2.16) defines a division G of $(1, \ldots, n)$ into two sets (G', G''). The division is made by cutting the cycle P at two points, one between s and its successor, the other between t and its successor.

We choose G' to be the set of the pair which contains k. If s belongs to G', a typical term in $C(a)$ has the form

$$a_{ki} \ldots a_{es} a_{sf} \ldots a_{gt} a_{th} \ldots a_{jk}. \tag{2.25}$$

If t belongs to G', the same argument will apply with s and t interchanged. Since

$$a_{sf} = a_{tf}, \quad a_{th} = a_{sh}, \tag{2.26}$$

the product (2.25) consists of a term of $C(a(G''))$ corresponding to the cyclic permutation $P'' = (t \to f \to \cdots \to g \to t)$, inserted in the middle of a term of $C(a(G'))$ corresponding to the cyclic permutation $P' = (k \to i \to \cdots \to e \to s \to h \to \cdots \to j \to k)$. If we keep G', G'' and P' fixed while summing over P'', all terms of $C(a(G''))$ appear in turn in equation (2.25). But $a(G'')$ is self-adjoint, and therefore $C(a(G''))$ is a scalar and commutes with the other factors in equation (2.25). After $C(a(G''))$ is extracted from the product, the remaining factors when summed over P' give $C(a(G'))$. We have thus proved the identity

$$C(a) = \sum_G C(a(G')) \, C(a(G'')). \tag{2.27}$$

Consider now the expansion (2.20) of $Q \det (a)$. Divide the sum $\sum_G$ into two parts $\sum_1$ and $\sum_2$, where $\sum_1$ contains the terms for which both s and t belong to the same subset G_u, and $\sum_2$ contains the terms with s and t in different subsets. Each term in $\sum_2$ is obtained from a unique term in $\sum_1$ by dividing G_u into two parts (G'_u, G''_u). Equation (2.27) applied to $a(G_u)$ shows that the term in $\sum_1$ is equal apart from sign to the sum of all the corresponding terms in $\sum_2$. The signs are opposite since m changes by one in going from $\sum_1$ to $\sum_2$. Therefore

$$\textstyle\sum_1 + \sum_2 = 0. \tag{2.28}$$

End of proof of Lemma 2.

III. Verification of Axioms

The definitions of Study and Dieudonné preserve the multiplicative property of a determinant (Axiom 2) but sacrifice the linearity property (Axiom 3). In contrast, Moore's definition (2.11) preserves Axiom 3 but sacrifices Axiom 2. It is clear that we cannot save Axiom 2, because equation (1.1) is usually meaningless. The product of two self-adjoint matrices a and b is in general neither self-adjoint nor almost self-adjoint, and so $Q \det (ab)$ is usually undefined. The fact that Axioms 1 and 3 are satisfied by Moore's definition of a determinant is expressed by [17]

Theorem 2. *Let R be a ring with commuting scalars and the scalar product property. Let a, b, c be almost self-adjoint matrices with elements in R, satisfying the conditions (1.2), (1.3), and (2.9) for some common row-index k. Then*

$$Q \det (a) + Q \det (b) = Q \det (c). \tag{3.1}$$

If further the ring S of scalars in R contains no zero-divisors of R, i.e., if

$$rs = 0, \quad r \in R, \quad s \in S \tag{3.2}$$

implies either $r = 0$ or $s = 0$, and if a is any self-adjoint matrix with elements in R, then

$$Q \det (a) = 0 \tag{3.3}$$

is a necessary and sufficient condition for the existence of a non-zero vector v with elements in R such that

$$av = 0. \tag{3.4}$$

This formulation of Axiom 1 does not require that R have no zero-divisors. We require only that there be no scalar zero-divisors. The ring of complex quaternions, for example, has zero-divisors but no scalar zero-divisors. The price we pay for this weakening of the conditions on R is that we can prove Axiom 1 only for self-adjoint matrices. For Axiom 1 to hold also for almost self-adjoint matrices, it is necessary that R have no zero-divisors. The statement of Axiom 1 for a (1×1) matrix, which is automatically almost self-adjoint, is just the statement that R has no zero-divisors. We do not know whether the absence of zero-divisors is a sufficient condition for Axiom 1 to hold for almost self-adjoint matrices. We shall return to this question in section IV.

Proof of Theorem 2. Equation (3.1) follows trivially from the definition (2.11) combined with equations (1.2), (1.3). We next assume that equation (3.4) holds and prove equation (3.3). Let k be an index for which

$$v_k \neq 0. \tag{3.5}$$

Write

$$d_{lk} = \epsilon_{kl} Q \det(a(k, l)), \tag{3.6}$$

so that equation (2.11) becomes

$$Q \det(a) = \sum_l a_{kl} d_{lk}. \tag{3.7}$$

For any $j \neq k$, the sum

$$\sum_l a_{jl} d_{lk} = Q \det(a'), \tag{3.8}$$

where a' is the almost self-adjoint matrix obtained from a by putting the elements of row k equal to the elements of row j. Lemma 2 states that $Q \det(a') = 0$. Now equation (3.4) with self-adjoint a implies

$$v^\dagger a = 0. \tag{3.9}$$

Thus, equations (3.7), (3.8), and (3.9) give

$$\sum_j \sum_l v_j^\dagger a_{jl} d_{lk} = v_k^\dagger Q \det(a) = 0. \tag{3.10}$$

Since $Q \det(a)$ is a scalar and cannot be a zero-divisor, equations (3.5) and (3.10) imply equation (3.3).

Next, we assume that equation (3.3) holds and find v to satisfy equation (3.4). If $a = 0$ there is nothing to prove. If $a \neq 0$ we can define an integer p in the range

$$2 \leqslant p \leqslant n \tag{3.11}$$

such that

$$Q \det(a(G_p)(k, l)) \neq 0 \tag{3.12}$$

for some G_p, k, l, while

$$Q \det(a(G_{p+1})(k, l)) = 0 \tag{3.13}$$

for all G_{p+1}, k, l. Here G_p means a subset containing p of the integers $(1, \ldots, n)$, $a(G_p)$ means the matrix $[a_{ij}]$ with i, j in G_p, and $a(G_p)(k,l)$ is defined like $a(k,l)$ in equation (2.11). If $p = n$, the condition (3.13) is empty. In any case we have for all G_p

$$Q \det (a(G_p)) = 0. \tag{3.14}$$

If $p = n$, equation (3.14) is just (3.3), and if $p < n$ equation (3.14) is included in equation (3.13). Now choose G_p, k, and l to satisfy equation (3.12), and write

$$v_j = \epsilon_{kj} Q \det (a(G_p)(k,j)), \quad j \in G_p, \tag{3.15}$$

with $v_j = 0$ for j not in G_p. Then $v \neq 0$. Consider the sum

$$s_i = \sum_j a_{ij} v_j. \tag{3.16}$$

For $i = k$, s_i is zero by equation (3.14). For i in G_p with $i \neq k$, s_i is a determinant of a matrix with two identical rows like equation (3.8) and is zero by Lemma 2. If i is not in G_p, then

$$s_i = Q \det (a(G_{p+1})(k, i)), \tag{3.17}$$

where G_{p+1} means the set obtained by adding i to G_p, and this vanishes by equation (3.13). Thus, $s_i = 0$ for all i, which means that equation (3.4) holds. End of proof of Theorem 2.

IV. Additional Results for Special Rings

We have been able to prove results stronger than Theorem 2 only for a restricted class of rings R. We say that R belongs to the *binary matrix class* if the elements of R are (2×2) matrices with conjugation defined by equation (2.8) and if the x, y, z, w (elements of elements of R) belong to a commutative ring S. The ring of scalars in R is either S or a subring of S. Every R in the binary matrix class is a ring with commuting scalars and the scalar product property. The three special rings mentioned in section I, the real and complex quaternions and the real (2×2) matrices, all belong to the binary matrix class. Jacobson's theorem [14] states that every non-commutative composition algebra belongs to the binary matrix class. For every ring R of the binary matrix class, and for every matrix a with elements in R, the Study determinant $\nabla(a)$ can be defined. A comparison of Study's definition [5] with equation (2.20) gives the relation

$$\nabla(a) = Q \det (a^\dagger a). \tag{4.1}$$

If R belongs to the binary matrix class, we may map each $(n \times n)$ matrix a with elements in R onto a $(2n \times 2n)$ matrix $A(a)$ with elements in S. $A(a)$ is obtained from a by letting each element $\begin{bmatrix} x & y \\ z & w \end{bmatrix}$ of a become a (2×2) block of elements in $A(a)$. The mapping $a \to A(a)$ preserves the operations of matrix addition and multiplication. The operation of conjugation is mapped according to

$$A(a^\dagger) = Y(A(a))^T Y^{-1}, \tag{4.2}$$

where T denotes transposition and Y is the matrix

$$Y = I_n \times \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \tag{4.3}$$

If we write

$$B(a) = -YA(a), \tag{4.4}$$

then

$$B(a^\dagger) = -(B(a))^T, \tag{4.5}$$

so that a is self-adjoint if and only if $B(a)$ is antisymmetric.

Since S is commutative, the ordinary determinant $\mathrm{Det}\,(A(a))$ exists and satisfies Axioms 1, 2, and 3 as an operation on matrices in S. Also the Pfaffian $\mathrm{Pf}(B(a))$ exists when a is self-adjoint [18]. These quantities are related to quaternion-determinants as follows.

$$Q\det(a^\dagger a) = Q\det(aa^\dagger) = \mathrm{Det}\,(A(a)) \tag{4.6}$$

for all a, and

$$Q\det(a) = \mathrm{Pf}(B(a)), \tag{4.7}$$

$$[Q\det(a)]^2 = \mathrm{Det}\,(A(a)), \tag{4.8}$$

for self-adjoint a. Equation (4.6) follows from equation (4.1) and was essentially proved by Study [5]. Equations (4.7), (4.8) were proved by Dyson [10]. The proofs assumed that S was the ring of complex numbers, but equations (4.6) to (4.8) are polynomial identities and therefore hold when S is any commutative ring.

As a consequence of equations (4.6) to (4.8) we can prove the following substitutes [19] for Axiom 2.

Theorem 3. *Let R be a ring of the binary matrix class. If c is any matrix over R and a any self-adjoint matrix, then*

$$Q\det(c^\dagger ac) = Q\det(c^\dagger c).\quad Q\det(a). \tag{4.9}$$

If a, b are two commuting self-adjoint matrices over R, then

$$Q\det(ab) = Q\det(a)\cdot Q\det(b). \tag{4.10}$$

Proof of Theorem 3. Since the ordinary determinant is multiplicative, equations (4.6) and (4.8) imply

$$\begin{aligned}[Q\det(c^\dagger ac)]^2 &= \mathrm{Det}\,(A(c^\dagger ac))\\ &= [\mathrm{Det}\,(A(c))]^2\,\mathrm{Det}\,(A(a)) = [Q\det(c^\dagger c)]^2[Q\det(a)]^2.\end{aligned} \tag{4.11}$$

Since both sides of equation (4.9) are polynomials in the elements of $A(a)$ and $A(c)$, which belong to the commutative ring S, equation (4.11) can only hold if equation (4.9) holds up to a sign. The sign is seen to be plus by taking $a = I_n$. Therefore equation (4.9) is proved. Similarly, if a and b are self-adjoint and commuting, ab is also self-adjoint, and equation (4.8) gives

$$[Q\det(ab)]^2 = \mathrm{Det}\,[A(ab)] = [Q\det(a)]^2[Q\det(b)]^2. \tag{4.12}$$

End of proof of Theorem 3.

It seems likely that equations (4.9) and (4.10) also hold for any ring with commuting scalars and the scalar product property, but we have not succeeded in proving them more generally. We had even less success with an extension of Axiom 2 to almost self-adjoint matrices, which we state here as

Conjecture 1. *Let R be a ring with commuting scalars and the scalar product property. Let a, b be almost self-adjoint matrices over R, satisfying equation (2.9) with the same value of k. Let a' be a self-adjoint matrix differing from a only in the kth row, and let b' be self-adjoint and differ from b only in the kth column. If a' and b' commute, then*

$$Q\det(ab) = Q\det(a)\cdot Q\det(b). \tag{4.13}$$

In particular, for any almost self-adjoint matrix a,

$$Q\det(a^\dagger a) = Q\det(a^\dagger)\cdot Q\det(a). \tag{4.14}$$

We did not succeed in proving equations (4.13) and (4.14) even for R in the binary matrix class. Equation (4.13) has been verified when a, b are (2×2) matrices, and for an extensive class of (3×3) matrices. If Conjecture 1 is taken as valid, then we have as a corollary an extension of Axiom 1 to determinants of almost self-adjoint matrices. We state this as

Conjecture 2. *Let R be a ring with commuting scalars and the scalar product property, and without zero-divisors. Let a be an almost self-adjoint matrix over R. Then*

$$Q\det(a) = 0 \tag{4.15}$$

if and only if a non-zero vector v exists with

$$av = 0. \tag{4.16}$$

Note the strengthening of the conditions on R as compared with Theorem 2, where we required only the absence of scalar zero-dividers.

Deduction of Conjecture 2 from Conjecture 1. We have been able to make this deduction only for R in the binary matrix class. Suppose that equation (4.15) holds. Then equations (4.6) and (4.14) imply

$$\mathrm{Det}\,(A(a) = 0. \tag{4.17}$$

Since Axiom 1 holds for ordinary determinants over S, there exists a non-zero $2n$-component vector V with elements in S such that

$$A(a)V = 0. \tag{4.18}$$

If the elements of v are defined by

$$v_j = \begin{bmatrix} V_{2j-1} & 0 \\ V_{2j} & 0 \end{bmatrix}, \tag{4.19}$$

v is non-zero and satisfies equation (4.16). Conversely, if equation (4.16) holds then

$$a^\dagger av = 0, \tag{4.20}$$

and Theorem 2 implies

$$Q\det(a^\dagger a) = 0. \tag{4.21}$$

Since we assume equation (4.14) to hold and R to have no zero-divisors, either $Q\det(a^\dagger)$ or $Q\det(a)$ must be zero. By equation (2.24), equation (4.15) holds in either case. End of deduction of Conjecture 2.

V. Conclusion

This paper has raised more questions than it has answered. The only substantial result proved with a satisfactory degree of generality is Theorem 2. We conclude with a list of the more interesting questions that remain open.

1. Is every ring with commuting scalars and the scalar product property equivalent to a ring in the binary matrix class? If not, can we construct a complete classification of such rings?

2. If R is a quaternion algebra which is also a division ring, the Dieudonné determinant $\det(a)$ is defined for all a, and Moore's determinant $Q\det(a)$ is defined for almost self-adjoint a. What is the relation between the two definitions?

3. The logical connections between the multiplicative properties of $Q\det(a)$ expressed in Theorem 3 and Conjecture 1 are not clear. Is there some more general multiplicative property from which equations (4.9), (4.10), (4.13) can be deduced as special cases? What is the widest class of ring R for which these properties hold?

4. Is there any natural way to extend the definition of $Q\det(a)$ to matrices which are not almost self-adjoint?

5. How generally does Conjecture 2 hold?

If these questions can be answered, the subject of quaternion-determinants will finally after 125 years have achieved the level of mathematical elegance and completeness that Hamilton would have wished for it. Even then, it is unlikely that the applications of quaternion-determinants to physics will ever be important. Hamilton's dream that the quaternion algebra would be the key to the understanding of the physical universe will remain a grand illusion. Hamilton's faith in the fundamental significance of quaternions was perhaps, like Kant's four categories and Jung's four psychological types, a manifestation of the ancient tradition of four-fold symbolism which Markus Fierz described in his essay *Die Vier Elemente*. Fierz showed how that tradition produced grotesque distortions of human thought which can still be traced in the writings of men of science from Aristotle to Jung. He might have added Hamilton and E. H. Moore to his list. Each of them was seized in his later years with an ambition to reconstruct the whole of mathematics in a grand design with the quaternions playing a central role. For each of them, the end result was, in Fierz's words, 'Das Symbol wird zum Schema und dieses zum furchtbaren Bette des Prokrustes, in das die Erscheinung nur verstümmelt gepresst werden kann'.

REFERENCES

[1] W. R. Hamilton, *Lectures on Quaternions* (Hodges and Smith, Dublin 1853).
[2] W. R. Hamilton, *Elements of Quaternions*, 2nd ed. (Ed. C. J. Joly; Longmans, London 1899).
[3] A. Cayley, Phil. Mag. *26*, 141 (1845).
[4] Ref. [2], Vol. 2, p. 361.
[5] E. Study, Acta Math. *42*, 1 (1920).
[6] E. H. Moore, Bull. Am. math. Soc. *28*, 161 (1922).
[7] E. H. Moore and R. W. Barnard, *General Analysis*, Part I (Memoirs of the American Philosophical Society, Philadelphia 1935).
[8] J. Dieudonné, Bull. Soc. math. Fr. *71*, 27 (1943). Dieudonné quotes the following earlier attempts to define generalized determinants: A. R. Richardson, Messenger Math. *55*, 145 (1926); A. R. Richardson, Proc. Lond. math. Soc. *28*, 395 (1928); A. Heyting, Math. Ann. *98*, 465 (1927); O. Ore, Ann. Math. *32*, 463 (1931).
[9] E. Artin, *Geometric Algebra*, Chap. 4 (Interscience, New York 1957).
[10] F. J. Dyson, Comm. Math. Phys. *19*, 235 (1970).
[11] M. L. Mehta, Comm. Math. Phys. *20*, 245 (1971).

[12] N. JACOBSON, Rend. Circ. mat. Palermo (2) *7*, 55 (1958).
[13] N. JACOBSON, *Structure and Representations of Jordan Algebras*, Am. math. Soc., Colloquium Pub., Vol. 39 (Providence, R.I. 1968).
[14] N. JACOBSON, Ref. [11], p. 164, Theorem 5.
[15] MOORE called it 'nearly Hermitian', see p. 131 of Ref. [7].
[16] Equation (2.20) is essentially MOORE's definition of the determinant, see p. 116 of Ref. [7].
[17] See p. 140 of Ref. [7].
[18] For a modern description of the properties of the Pfaffian see N. JACOBSON, Ref. [11], p. 230, Example C.
[19] Equations (4.9) and (4.10) appear on p. 135 and p. 147 of Ref. [7].