A CANONICAL FORM FOR REAL MATRICES UNDER ORTHOGONAL TRANSFORMATIONS

BY F. D. MURNAGHAN AND A. WINTNER

DEPARTMENT OF MATHEMATICS, THE JOHNS HOPKINS UNIVERSITY

Communicated May 18, 1931

If $\mathbf{A}$ is a square matrix of order n with real or complex elements it is well known that it may be reduced by means of a unitary transformation $\mathbf{U}$ to a matrix of the same order all of whose elements below the leading diagonal are zero.[1] Even when the elements of $\mathbf{A}$ are real the elements of the transforming matrix $\mathbf{U}$ are complex if the characteristic numbers of $\mathbf{A}$ are not all real and it is desirable to give a canonical form which may be reached by the use of *real* unitary (i.e., orthogonal) matrices. The derivation of this canonical form differs only in detail from that given by Schur.

The characteristic numbers λ of the matrix $\mathbf{A}$ are determined by the equation $\det(\mathbf{A} - \lambda\mathbf{E}) = 0$, where $\mathbf{E}$ is the unit matrix, and they may be real or complex. If, as we suppose, the elements of $\mathbf{A}$ are real the complex roots will occur in conjugate imaginary pairs. If all the characteristic numbers are real the unitary transformations occurring in Schur's derivation will be real and the canonical form sought for is that given by Schur. On the other hand, let $\lambda_1 = \mu + i\nu$ and $\lambda_2 = \mu - i\nu$ be a pair of conjugate complex characteristic numbers of the matrix $\mathbf{A}$ (μ, ν real, $\nu \neq 0$); on denoting by $\mathbf{x}_1 = \mathbf{a} + i\mathbf{b}$, ($\mathbf{a}$, $\mathbf{b}$, real) a characteristic vector of $\mathbf{A}$ associated with the characteristic number λ_1 we have $\mathbf{A}\mathbf{x}_1 = \lambda_1\mathbf{x}_1$ which implies the two equations

$$\mathbf{A}\mathbf{a} = \mu\mathbf{a} - \nu\mathbf{b}; \quad \mathbf{A}\mathbf{b} = \mu\mathbf{b} + \nu\mathbf{a}. \tag{1}$$

It is clear that neither $\mathbf{a}$ nor $\mathbf{b}$ can be the zero vector; for if $\mathbf{a} = \mathbf{0}$ then $\mathbf{b} = \mathbf{0}$ from the first of the two equations just given and hence $\mathbf{x}_1 = \mathbf{0}$, which is at variance with the statement that $\mathbf{x}_1$ is a characteristic vector; similarly, if $\mathbf{b} = \mathbf{0}$, $\mathbf{a} = \mathbf{0}$ from the second of the two equations and $\mathbf{x}_1$ would again $= \mathbf{0}$. Furthermore $\mathbf{b}$ cannot be a multiple of $\mathbf{a}$, for then $\mathbf{x}_1' = \mathbf{a}$ would be a characteristic vector of $\mathbf{A}$ corresponding to λ_1 and we have just seen that this is impossible. Hence the two vectors $\mathbf{a}$ and $\mathbf{b}$ determine a plane and if $\mathbf{u}$ and $\mathbf{v}$ are two unit orthogonal vectors in this plane, $\mathbf{a}$ and $\mathbf{b}$ may be expressed in the form

$$\mathbf{a} = \alpha\mathbf{u} + \beta\mathbf{v}; \quad \mathbf{b} = \gamma\mathbf{u} + \delta\mathbf{v},$$

where we may, without lack of generality, assume $\alpha\delta - \beta\gamma = 1$ (since a characteristic vector is indeterminate to the extent of a scalar factor). On inserting these expressions in (1) and solving for $\mathbf{A}\mathbf{u}$ and $\mathbf{A}\mathbf{v}$ we obtain

$$\mathbf{Au} = [\mu - \nu(\alpha\beta + \gamma\delta)]\mathbf{u} - \nu(\beta^2 + \delta^2)\mathbf{v}; \ \mathbf{Av} = \nu(\alpha^2 + \gamma^2)\mathbf{u} + [\mu + \nu(\alpha\beta + \gamma\delta)]\mathbf{v}. \tag{2}$$

If now we denote by $\mathbf{O}$ an orthogonal matrix whose columns consist, in the order given, of the components of $\mathbf{u}$ and of $\mathbf{v}$ and of n-2 other real vectors forming with these a set of mutually orthogonal unit vectors, we have $\mathbf{u} = \mathbf{Oe}_1$, $\mathbf{v} = \mathbf{Oe}_2$, where $\mathbf{e}_1$ and $\mathbf{e}_2$ are the vectors (1, 0, 0 ... 0) and (0, 1, 0 ... 0), respectively. It follows from equations (2), on denoting the matrix $\mathbf{O}^{-1}\mathbf{AO}$ by $\mathbf{B}$ that

$$\mathbf{Be}_1 = [\mu - \nu(\alpha\beta + \gamma\delta)]\mathbf{e}_1 - \nu(\beta^2 + \delta^2)\mathbf{e}_2; \ \mathbf{Be}_2 = \nu(\alpha^2 + \gamma^2)\mathbf{e}_1 + [\mu + \nu(\alpha\beta + \gamma\delta)]\mathbf{e}_2, \tag{3}$$

and on considering the third, fourth, ... to last components of these vector equations we find $b_1^p = 0 = b_2^q$ (p, $q = 3, \ldots n$), where b_s^r denotes the element in the r^{th} row and s^{th} column of $\mathbf{B}$. Hence $\mathbf{B}$ has all elements in the first two columns, under the first two rows, zero. By properly choosing the unit orthogonal vectors ($\mathbf{u}$, $\mathbf{v}$) in the ($\mathbf{a}$, $\mathbf{b}$) plane the elements b_2^1 and b_1^2 may be made one the negative of the other. This requires $\alpha^2 + \gamma^2 = \beta^2 + \delta^2$ and if this equality is not satisfied a rotation of the ($\mathbf{u}$, $\mathbf{v}$) vectors through an angle

$$\theta = {}^1/_2 \operatorname{arc} \tan(\beta^2 + \delta^2 - \alpha^2 - \gamma^2)/2(\alpha\beta + \gamma\delta)$$

in their plane procures it. Proceeding in this way with the remaining pairs of conjugate complex characteristic numbers (if any) and by Schur's method with the real characteristic numbers we find as a canonical form, under orthogonal transformations, of an arbitrary real matrix $\mathbf{A}$ the form

$$\mathbf{C} = \begin{bmatrix} c_1^1 & c_2^1 \ldots & c_n^1 \\ c_1^2 & c_2^2 \ldots & c_n^2 \\ 0 & 0 & c_3^3 \ldots \\ 0 & 0 & c_3^4 \ldots \\ 0 & 0 & 0 \ldots \end{bmatrix}; c_1^2 = -c_2^1; \ c_3^4 = -c_4^3, \text{ etc.} \tag{4}$$

where when, instead of a pair of conjugate complex characteristic numbers, we have a real characteristic number the element such as c_1^2 or c_3^4 is also zero. Since the characteristic numbers of a matrix are invariant under transformations of the matrix it is clear that λ_1 and λ_2 are the characteristic numbers of the *two-rowed* matrix $\begin{pmatrix} c_1^1 & c_2^1 \\ c_1^2 & c_2^2 \end{pmatrix}$, $c_1^2 = -c_2^1$ and so on.

Normal Matrices.—A matrix $\mathbf{A}$ is said to be normal when it is commutable with its transposed, i.e., $\mathbf{AA}' = \mathbf{A}'\mathbf{A}$. This property is invariant under transformation by an orthogonal matrix. On expressing the fact that $\mathbf{C}$ is normal we find[4] that $c_3^1 = c_4^1 = \quad = c_n^1 = 0$; $c_1^1 = c_2^2$ and so on, so that the canonical form for normal matrices is

$$\mathbf{C}_N = \begin{bmatrix} c_1^1 & c_2^1 & 0 \ldots & & 0 \\ c_1^2 & c_1^1 & 0 \ldots & & 0 \\ 0 & 0 & c_3^3 & c_4^3 & 0 \ldots 0 \\ 0 & 0 & c_3^4 & c_3^3 & 0 \ldots 0 \end{bmatrix} ; c_1^2 = -c_2^1, \text{ etc.} \tag{5}$$

Since **A** is normal and remains so under orthogonal transformations, it is clear that its canonical form $\mathbf{C}_N$ consists of normal two-rowed matrices (if n is odd the canonical form also contains one "one-rowed" matrix). The transformation or invariant theory of real normal matrices with respect to the group of (real) orthogonal transformations is accordingly reduced to the case $n = 2$.

In order that a two-rowed matrix should be normal, it is clearly necessary and sufficient that it should be either symmetrical (containing three arbitrary elements) or of the form

$$\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \tag{6}$$

(containing only two arbitrary elements), the common boundary of the domains of these two classes being the set of diagonal matrices for which the two diagonal elements are equal, i.e., scalar matrices.

Returning to the case of an arbitrary n it is clear that the (real) symmetric, skew-symmetric and orthogonal matrices are normal and remain symmetric, skew-symmetric or orthogonal, respectively, under orthogonal transformations so that also the two-rowed matrices of their canonical forms must belong respectively to these three types. The skew-symmetric and orthogonal two-rowed matrices being characterized by the fact that they may be written in the form (6) where $\alpha = 0$ and $\alpha^2 + \beta^2 = 1$, respectively, we see that the above theorem on the canonical form of normal matrices yields, not only for the symmetric but also for the skew-symmetric and orthogonal matrices, the classical canonical forms which in the literature, for instance Weyl,[5] are not obtained by a common demonstration. We have, of course, canonical forms also for normal matrices which do not belong to these three special types. The characteristic numbers of the matrix (6) being $\alpha \pm i\beta$ it is clear from (5) that any n rowed real normal matrix is then and only then symmetric, skew-symmetric or orthogonal if all its characteristic numbers lie on the real axis, or the imaginary axis, or the boundary of the unit circle respectively (only the first part of the theorem is generally known, but from the polygon rule of Toeplitz the second part of the theorem concerning the sufficiency can be derived).

Any normal matrix **A** may be represented as the product of two commutable matrices **P** and **O** where **P** is symmetric and not-negative definite and **O** is orthogonal. Conversely any matrix **A** which may be represented in this form is clearly normal. This product representation[6] which corresponds to the polar representation $a = pt$, $p \geqslant$ zero and $|t| = 1$, of the ordinary complex numbers a may obviously be reduced to the case $n = 2$ and in this case a simple calculation verifies the theorem.

[1] Schur, I., *Math. Annalen*, **66**, 489–492 (1909).

[2] The corresponding question for those real matrices which can be transformed by a *not necessarily* unitary matrix into the diagonal form (i.e., those having only simple elementary divisors) but which cannot be transformed into a real diagonal form has been discussed by Weierstrass in his theory of small vibrations.

[3] Toeplitz, O., *Math. Zeit.*, **2**, 187–197 (1918).

[4] Cf. Toeplitz, loc. cit.

[5] Weyl, H., *Gruppentheorie und Quantenmechanik*, 20, 23, 274.

[6] Cf. Wintner, A., *Math. Zeit.*, **30**, 280–281 (1929).

WAVE MOTION AND THE EQUATION OF CONTINUITY

BY R. B. LINDSAY

DEPARTMENT OF PHYSICS, BROWN UNIVERSITY

Communicated May 27, 1931

It is well known that small disturbances from equilibrium in a compressible fluid medium are propagated in accordance with the wave equation

$$\nabla^2 \varphi = \ddot{\varphi}/c^2, \tag{1}$$

where φ is the velocity potential (i.e., the function of the coördinates such that $\dot{\xi}$, $\dot{\eta}$, $\dot{\rho}$, the component particle velocities of the medium in the x, y, z directions, are equal, respectively, to $\frac{\partial\varphi}{\partial x}, \frac{\partial\varphi}{\partial y}, \frac{\partial\varphi}{\partial z}$), and c is the velocity of propagation. Perhaps the simplest way to deduce this equation is to consider the two fundamental equations: (A), that which expresses the essential continuity of the medium and (B), the hydrodynamic equation of motion written under the approximation appropriate to the small changes assumed. If $\delta\rho$ is the variation from the equilibrium density ρ_0, we introduce the important auxiliary quantity s, the condensation, equal to $\delta\rho/\rho_0$. We may then write

(A) The equation of continuity

$$\nabla^2\varphi = -\dot{s}\ . \tag{2}$$

(B) The equation of motion[1]

$$\dot{\varphi} = -c^2 s. \tag{3}$$