*Further Remarks on the Stress-Deformation Relations for Isotropic Materials**

R. S. RIVLIN

Brown University, Providence, R. I.

1. Introduction. Let us consider a body of some material to be undergoing deformation. If X_i denotes the position of a material particle of the body in the rectangular Cartesian coordinate system x_i at zero time, before the deformation has commenced, and x_i its position in the same coordinate system at a later time t , then we may describe the manner in which the body is deformed by the functional dependence of x_i on X_j and t , thus:

$$x_i = x_i(X_j, t). \quad (1.1)$$

At the instant of time t , the displacements u_i of the particles of the body are given by $x_i - X_i$, their velocities $v_i^{(1)}$ by Dx_i/Dt , their accelerations $v_i^{(2)}$ by D^2x_i/Dt^2 and their $(n - 1)^{\text{th}}$ accelerations $v_i^{(n)}$ by $D^n x_i/Dt^n$, where D/Dt denotes differentiation with respect to time keeping X_j constant. Let us assume that the components of stress t_{ij} in the coordinate system x_i at a point of the body may be expressed as polynomials in the gradients of the displacement, velocity, acceleration, $\dots$, $(n - 1)^{\text{th}}$ acceleration at that point; *i.e.* as polynomials in $\partial u_i/\partial X_j$, $\partial v_i^{(1)}/\partial x_j$, $\partial v_i^{(2)}/\partial x_j$, $\dots$, $\partial v_i^{(n)}/\partial x_j$. It has been shown (RIVLIN & ERICKSEN 1955) that the stress components must be expressible in terms of these gradients as polynomials in the components of $(n + 1)$ symmetric tensors C_{ij} , $A_{ij}^{(1)}$, $A_{ij}^{(2)}$, $\dots$, $A_{ij}^{(n)}$ defined by

$$C_{ij} = \frac{\partial x_i}{\partial X_k} \frac{\partial x_j}{\partial X_k}, \quad A_{ij}^{(1)} = \frac{\partial v_j^{(1)}}{\partial x_j} + \frac{\partial v_j^{(1)}}{\partial x_i}, \quad (1.2)$$

* The results presented in this paper were obtained in the course of research sponsored by the Office of Ordnance Research, Department of the Army, under contract No. DA-19-020-ORD-3487.

and the recurrence relation

$$A_{ij}^{(r)} = \frac{DA_{ij}^{(r-1)}}{Dt} + A_{mj}^{(r-1)} \frac{\partial v_m^{(1)}}{\partial x_i} + A_{im}^{(r-1)} \frac{\partial v_m^{(1)}}{\partial x_j} \quad (r = 2, \dots, n). \quad (1.3)$$

We define the stress matrix $\mathbf{T}$ and the kinematic matrices $\mathbf{C}$, $\mathbf{A}_1$, $\mathbf{A}_2$, $\dots$, $\mathbf{A}_n$ by

$$\mathbf{T} = \| t_{ij} \|, \quad \mathbf{C} = \| C_{ij} \| \quad \text{and} \quad \mathbf{A}_r = \| A_{ij}^{(r)} \|. \quad (1.4)$$

It has been shown (RIVLIN & ERICKSEN 1955) that if we assume the material of the body to be isotropic in its undeformed state, then the stress matrix $\mathbf{T}$ may be expressed as a matrix polynomial in the kinematic matrices $\mathbf{C}$, $\mathbf{A}_1$, $\mathbf{A}_2$, $\dots$, $\mathbf{A}_n$. The coefficients in this matrix polynomial are themselves polynomials in the elements of the kinematic matrices which are invariant under orthogonal transformations.

In the particular case when the stress components t_{ij} are assumed to be expressible as polynomials in terms of the velocity gradients only, the stress matrix $\mathbf{T}$ is expressible as a polynomial in the kinematic matrix $\mathbf{A}_1$ and the coefficients in this polynomial are expressible as polynomials in the invariants $\text{tr } \mathbf{A}_1$, $\text{tr } \mathbf{A}_1^2$ and $\text{tr } \mathbf{A}_1^3$. The Hamilton-Cayley theorem may then be used (as has been pointed out by REINER (1945) and in a slightly different context by PRAGER (1945)) to show that $\mathbf{T}$ must be expressible in the closed form

$$\mathbf{T} = \alpha_0 \mathbf{I} + \alpha_1 \mathbf{A}_1 + \alpha_2 \mathbf{A}_1^2, \quad (1.5)$$

where α_0 , α_1 and α_2 are polynomials in $\text{tr } \mathbf{A}_1$, $\text{tr } \mathbf{A}_1^2$ and $\text{tr } \mathbf{A}_1^3$.

Similarly, if the stress components are assumed to be expressible as polynomials in the displacement gradients only, the stress matrix $\mathbf{T}$ may be expressed in the form

$$\mathbf{T} = \alpha_0 \mathbf{I} + \alpha_1 \mathbf{C} + \alpha_2 \mathbf{C}^2, \quad (1.6)$$

where α_0 , α_1 and α_2 are polynomials in $\text{tr } \mathbf{C}$, $\text{tr } \mathbf{C}^2$ and $\text{tr } \mathbf{C}^3$.

In the present paper, we consider the more general case in which the stress matrix $\mathbf{T}$ is expressible as a polynomial in the $(n + 1)$ kinematic matrices $\mathbf{C}$, $\mathbf{A}_1$, $\mathbf{A}_2$, $\dots$, $\mathbf{A}_n$ defined by equations (1.2), (1.3) and (1.4). It is shown that if $\mathbf{T}$ is expressible as a polynomial in only two of the kinematic matrices, then the Hamilton-Cayley theorem can be used to reduce the expression for $\mathbf{T}$ as a matrix polynomial of arbitrary form in these two kinematic matrices to one of closed form. For example, it is shown that if any symmetric matrix $\mathbf{T}$ is expressible as a matrix polynomial in two symmetric 3×3 matrices $\mathbf{A}$ and $\mathbf{B}$, with coefficients which are polynomials in the elements of $\mathbf{A}$ and $\mathbf{B}$ invariant under orthogonal transformations, then $\mathbf{T}$ must be expressible in the form

$$\begin{aligned} \mathbf{T} = & \alpha_0 \mathbf{I} + \alpha_1 \mathbf{A} + \alpha_2 \mathbf{B} + \alpha_3 \mathbf{A}^2 + \alpha_4 \mathbf{B}^2 + \alpha_5 (\mathbf{AB} + \mathbf{BA}) + \alpha_6 (\mathbf{A}^2 \mathbf{B} + \mathbf{BA}^2) \\ & + \alpha_7 (\mathbf{AB}^2 + \mathbf{B}^2 \mathbf{A}) + \alpha_8 (\mathbf{A}^2 \mathbf{B}^2 + \mathbf{B}^2 \mathbf{A}^2), \end{aligned} \quad (1.7)$$

where the α 's are polynomials in $\text{tr } \mathbf{A}$, $\text{tr } \mathbf{B}$, $\text{tr } \mathbf{A}^2$, $\text{tr } \mathbf{B}^2$, $\text{tr } \mathbf{A}^3$, $\text{tr } \mathbf{B}^3$, $\text{tr } \mathbf{AB}$, $\text{tr } \mathbf{A}^2 \mathbf{B}$, $\text{tr } \mathbf{AB}^2$ and $\text{tr } \mathbf{A}^2 \mathbf{B}^2$. Replacing $\mathbf{A}$ and $\mathbf{B}$ in (1.7) by the kinematic matrices $\mathbf{C}$ and $\mathbf{A}_1$, defined by equations (1.2) and (1.4), we obtain an explicit expression, in closed form, for the stress matrix, if it is assumed that the stress components in the rectangular Cartesian coordinate system x_i are expressible as polynomials in the displacement gradients and velocity gradients only.

Again, replacing $\mathbf{A}$ and $\mathbf{B}$ in (1.7) by the kinematic matrices $\mathbf{A}_1$ and $\mathbf{A}_2$, defined by equations (1.2) to (1.4), we obtain an explicit expression, in closed form, for the stress matrix, if it is assumed that the stress components in the rectangular Cartesian coordinate system x_i are expressible as polynomials in the velocity gradients and acceleration gradients only.

Equation (1.7) can probably be derived from the result of the classical theory of invariants in which the integrity basis for an arbitrary number of ternary quadratic forms has been obtained (TODD, 1948). However, the specialised notation of the classical theory of invariants makes it desirable that the result be derived in a manner which will render it readily accessible to an appreciable number of applied mathematicians.

It is hoped in a later investigation to derive an expression for $\mathbf{T}$ analogous to (1.7) in the case when $\mathbf{T}$ is a symmetric 3×3 matrix expressible as a polynomial in R , symmetric 3×3 matrices. In the present paper, we consider the simpler problem in which $\mathbf{T}$ is a symmetric 2×2 matrix expressible as a polynomial in R , symmetric 2×2 matrices, $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_R$. It is shown that $\mathbf{T}$ may then be expressed in the form

$$\mathbf{T} = \varphi_0 \mathbf{I} + \sum_{r=1}^R \varphi_r \mathbf{A}_r, \quad (1.8)$$

where the φ 's are expressible as polynomials in $\text{tr } \mathbf{A}_r$ ($r = 1, 2, \dots, R$) and $\text{tr } \mathbf{A}_r \mathbf{A}_s$ ($r, s = 1, 2, \dots, R$).

We note that if the deformation described by (1.1) is planar so that $u_3 = 0$ and x_1 and x_2 are independent of x_3 , then

$$C_{i3} = \delta_{i3} \quad \text{and} \quad A_{ij}^{(r)} = 0. \quad (1.9)$$

If the stress matrix $\mathbf{T}$ is expressible as a matrix polynomial in the n kinematic matrices $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_n$ defined by (1.2) to (1.4), then it is apparent that for a planar deformation the components t_{i3} of the stress are zero and the remaining components are given by (1.8) in which $R = n$, $\mathbf{T}$ is the symmetric 2×2 matrix formed from the stress components t_{ij} ($i, j = 1, 2$) and $\mathbf{A}_r$ ($r = 1, 2, \dots, R$) are the symmetric 2×2 matrices formed from the components $A_{ij}^{(r)}$ ($i, j = 1, 2$) defined by (1.2) and (1.3).

In a previous paper (RIVLIN & ERICKSEN, 1955), a somewhat more general problem than that considered in the present paper was discussed. It was assumed that the components of stress t_{ij} are single-valued functions (not necessarily polynomials) of the gradients of displacement, velocity, acceleration, $\dots$, $(n-1)^{\text{th}}$ acceleration. It was shown that they must then be expressible as single-

valued functions of the elements of the kinematic matrices $\mathbf{C}$, $\mathbf{A}_1$, $\mathbf{A}_2$, $\dots$, $\mathbf{A}_n$ defined by equations (1.2), (1.3) and (1.4). It was further shown that the stress matrix $\mathbf{T}$ can then be expressed as a linear combination of a small number of simple linearly independent matrix polynomials in the kinematic matrices, the coefficients in this expression for $\mathbf{T}$ being invariants formed from the elements of the kinematic matrices. However, unlike those in the formulae for $\mathbf{T}$ developed in the present paper, these coefficients are not, in general, polynomials in the elements of the kinematic matrices and therefore may be indeterminate for particular values of these elements. This is generally the case even if it is assumed initially that t_{ij} is expressible as a polynomial in the gradients of the displacement, velocity, acceleration, $\dots$, $(n-1)^{\text{th}}$ acceleration and hence in the elements of the kinematic matrices.

For materials for which this assumption is valid, it is desirable to have expressions in closed form for the stress matrix, in which this artificial degeneracy is not introduced and it is this desirability which motivates the present paper.

2. Definitions. Let $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\dots$ be R , $n \times n$ matrices defined by

$$\mathbf{A} = \| a_{ij} \|, \quad \mathbf{B} = \| b_{ij} \| \quad \text{etc.} \quad (2.1)$$

Consider any product $\mathbf{\Pi}$ of these defined by

$$\mathbf{\Pi} = \mathbf{A}^{\alpha_1} \mathbf{B}^{\beta_1} \mathbf{C}^{\gamma_1} \dots \mathbf{A}^{\alpha_2} \mathbf{B}^{\beta_2} \mathbf{C}^{\gamma_2} \dots \mathbf{A}^{\alpha_r} \mathbf{B}^{\beta_r} \mathbf{C}^{\gamma_r} \dots, \quad (2.2)$$

and formed by a finite number of successive multiplications of the matrices in any order. It is considered that in the expression (2.2) for $\mathbf{\Pi}$ the indices α_r , β_r , γ_r , $\dots$ are positive integers or zero and that no two adjacent terms are powers of the same matrix. The degree in $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\dots$ of the product $\mathbf{\Pi}$ is

$$\sum_r (\alpha_r + \beta_r + \gamma_r). \quad (2.3)$$

Its *degree* in $\mathbf{A}$ is $\sum_r \alpha_r$, its *degree* in $\mathbf{B}$ is $\sum_r \beta_r$ and so on.

We define the *extension* of the matrix product $\mathbf{\Pi}$ as the number of non-zero terms in the sum (2.3). The unit matrix has, of course, zero extension.

A *polynomial* in the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\dots$ is defined as any sum of products of the type $\mathbf{\Pi}$, with scalar coefficients which are expressible as polynomials in the elements of the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\dots$. We define the extension of such a polynomial as that of its term of greatest extension, its degree in the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\dots$ as that of its term of highest degree and its degree in $\mathbf{A}$ as that of its term of highest degree in $\mathbf{A}$.

The replacement of a polynomial by an identically equal polynomial of lower extension will be called *contraction* of the polynomial.

We shall call the matrix product $\mathbf{\Pi}$ defined by (2.2) *symmetric* if $\mathbf{\Pi} = \mathbf{\Pi}^*$, where $\mathbf{\Pi}^*$ is the transpose of $\mathbf{\Pi}$. We note that if the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\dots$ are each symmetric matrices then

$$\mathbf{\Pi}^* = \dots \mathbf{C}^{\gamma_r} \mathbf{B}^{\beta_r} \mathbf{A}^{\alpha_r} \dots \mathbf{C}^{\gamma_2} \mathbf{B}^{\beta_2} \mathbf{A}^{\alpha_2} \dots \mathbf{C}^{\gamma_1} \mathbf{B}^{\beta_1} \mathbf{A}^{\alpha_1}. \quad (2.4)$$

A polynomial $\mathbf{P}$ in the matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$ will be described as a symmetric polynomial if $\mathbf{P} = \mathbf{P}^*$, where $\mathbf{P}^*$ is the transpose of $\mathbf{P}$. Again, we note that if the matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$ are each symmetric matrices then we can form the transpose of $\mathbf{P}$ by replacing each term $\mathbf{II}$ in the expression for $\mathbf{P}$ by its transpose $\mathbf{II}^*$ given by (2.4).

3. The reduction of degree of a matrix polynomial. We consider the matrix product $\mathbf{II}$, defined by (2.2), of a number of $n \times n$ matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$. We can employ the Hamilton-Cayley theorem

$$\mathbf{A}^n - \alpha_1 \mathbf{A}^{n-1} + \alpha_2 \mathbf{A}^{n-2} + \dots + (-1)^n \alpha_n \mathbf{I} = 0$$

for the matrix $\mathbf{A}$, where $\alpha_1, \alpha_2, \dots, \alpha_n$ are scalars expressible as polynomials in the elements of $\mathbf{A}$, to express $\mathbf{A}^{\alpha_r}$ with $\alpha_r \geq n$ as a polynomial of degree $n - 1$ in $\mathbf{A}$. Thus, if α_r has any positive integral value or zero, $\mathbf{A}^{\alpha_r}$ may be expressed as a polynomial of degree $n - 1$ or less in $\mathbf{A}$. Similarly, if $\beta_r, \gamma_r, \dots$ are positive integers or zero, $\mathbf{B}^{\beta_r}, \mathbf{C}^{\gamma_r}, \dots$ may be expressed as polynomials of degree $n - 1$ or less in $\mathbf{B}, \mathbf{C}, \dots$ respectively. It follows immediately that $\mathbf{II}$ may be expressed as a polynomial in $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$ in which each term is of the form (2.2) with $\alpha_r, \beta_r, \dots$ ($r = 1, 2, \dots$) each less than n , the coefficients of these terms being scalars which are expressible as polynomials in the elements of the matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$. It is thus apparent that any polynomial in $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$ may be expressed as a polynomial in which the matrix products take the form (2.2) with $\alpha_r, \beta_r, \dots$ ($r = 1, 2, \dots$) each less than n .

4. An extension of the Hamilton-Cayley theorem. Since

$$\begin{vmatrix} \delta_{i_1 j_1} & \delta_{i_1 j_2} & \dots & \delta_{i_1 j_{n+r}} \\ \delta_{i_2 j_1} & \delta_{i_2 j_2} & \dots & \delta_{i_2 j_{n+r}} \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \delta_{i_{n+r} j_1} & \delta_{i_{n+r} j_2} & \dots & \delta_{i_{n+r} j_{n+r}} \end{vmatrix} = 0, \quad (4.1)$$

for $r = 1, 2, \dots$, where the δ 's are Kronecker deltas in n dimensions, it follows immediately that, if $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots, \mathbf{N}$ are $n, n \times n$ matrices defined by (2.1) and we take $r = 1$ in (4.1),

$$\begin{vmatrix} \delta_{i_1 j_1} & \delta_{i_1 j_2} & \dots & \delta_{i_1 j_{n+1}} \\ \delta_{i_2 j_1} & \delta_{i_2 j_2} & \dots & \delta_{i_2 j_{n+1}} \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \delta_{i_{n+1} j_1} & \delta_{i_{n+1} j_2} & \dots & \delta_{i_{n+1} j_{n+1}} \end{vmatrix} a_{i_1 j_1} b_{i_2 j_2} c_{i_3 j_3} \dots n_{i_n j_n} = 0. \quad (4.2)$$

We may write

$$\begin{vmatrix} \delta_{i_1 j_1} & \delta_{i_2 j_2} & \cdots & \delta_{i_1 j_{n+1}} \\ \delta_{i_2 j_1} & \delta_{i_2 j_2} & \cdots & \delta_{i_2 j_{n+1}} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \delta_{i_{n+1} j_1} & \delta_{i_{n+1} j_2} & \cdots & \delta_{i_{n+1} j_{n+1}} \end{vmatrix} = \epsilon_{pq \cdots st} \delta_{i_1 j_p} \delta_{i_2 j_q} \cdots \delta_{i_n j_s} \delta_{i_{n+1} j_t}, \quad (4.3)$$

where

$$\begin{aligned} \epsilon_{pq \cdots st} &= 0 \text{ if } p, q, \cdots s, t \text{ are not all different} \\ &= 1 \text{ or } -1 \text{ accordingly as } p, q, \cdots s, t \text{ is an even or odd permutation of} \\ &\quad 1, 2, \cdots, n+1. \end{aligned}$$

In (4.3), the double suffix convention for summation is used for the indices $p, q, \cdots s, t$ even though they appear as subscripts to ϵ and subscripts to the subscripts i and j on δ .

Employing (4.3) in (4.2) we have

$$\begin{aligned} \epsilon_{pq \cdots st} \delta_{i_1 j_p} \delta_{i_2 j_q} \cdots \delta_{i_n j_s} \delta_{i_{n+1} j_t} a_{i_1 j_1} b_{i_2 j_2} \cdots n_{i_n j_n} \\ = \epsilon_{pq \cdots st} a_{j_p j_1} b_{j_q j_2} \cdots n_{j_s j_n} \delta_{i_{n+1} j_t} = 0. \end{aligned} \quad (4.4)$$

The Case of 2×2 Matrices. If **A** and **B** are 2×2 matrices, then (4.4) becomes

$$\epsilon_{pqr} a_{j_p j_1} b_{j_q j_2} \delta_{i_3 j_r} = 0. \quad (4.5)$$

This is equivalent to

$$\begin{aligned} (a_{j_1 j_1} b_{j_2 j_2} - a_{j_2 j_1} b_{j_1 j_2}) \delta_{i_3 j_3} - a_{j_1 j_1} b_{j_3 i_3} \\ - b_{j_2 j_2} a_{j_3 i_3} + b_{j_3 j_2} a_{j_2 i_3} + a_{j_3 j_1} b_{j_1 i_3} = 0. \end{aligned} \quad (4.6)$$

Equation (4.6) may be re-written as the matrix equation

$$\mathbf{AB} + \mathbf{BA} - \mathbf{A} \operatorname{tr} \mathbf{B} - \mathbf{B} \operatorname{tr} \mathbf{A} + (\operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{B} - \operatorname{tr} \mathbf{AB}) \mathbf{I} = 0. \quad (4.7)$$

Now, let us suppose that **A**, **B** and **C** are 2×2 matrices. Taking $n = 2$ and $r = 2$ in (4.1) and making use of a relation of the type (4.3), we have

$$\epsilon_{pqrs} \delta_{i_1 j_p} \delta_{i_2 j_q} \delta_{i_3 j_r} \delta_{i_4 j_s} = 0. \quad (4.8)$$

From (4.8), we obtain

$$\epsilon_{pqrs} \delta_{i_1 j_p} \delta_{i_2 j_q} \delta_{i_3 j_r} \delta_{i_4 j_s} a_{i_1 j_1} b_{i_2 j_2} c_{i_3 j_3} = 0. \quad (4.9)$$

Whence,

$$\epsilon_{pqrs} a_{j_p j_1} b_{j_q j_2} c_{j_r j_3} \delta_{i_4 j_s} = 0. \quad (4.10)$$

This may be re-written in matrix notation as

$$\begin{aligned} & \mathbf{ABC} + \mathbf{CBA} + \mathbf{BAC} + \mathbf{CAB} + \mathbf{ACB} + \mathbf{BCA} \\ &= (\mathbf{BC} + \mathbf{CB}) \operatorname{tr} \mathbf{A} + (\mathbf{CA} + \mathbf{AC}) \operatorname{tr} \mathbf{B} + (\mathbf{AB} + \mathbf{BA}) \operatorname{tr} \mathbf{C} \\ &+ \mathbf{A}(\operatorname{tr} \mathbf{BC} - \operatorname{tr} \mathbf{B} \operatorname{tr} \mathbf{C}) + \mathbf{B}(\operatorname{tr} \mathbf{CA} - \operatorname{tr} \mathbf{C} \operatorname{tr} \mathbf{A}) \\ &+ \mathbf{C}(\operatorname{tr} \mathbf{AB} - \operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{B}) + \mathbf{I}(\operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{B} \operatorname{tr} \mathbf{C} - \operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{BC} \\ &- \operatorname{tr} \mathbf{B} \operatorname{tr} \mathbf{CA} - \operatorname{tr} \mathbf{C} \operatorname{tr} \mathbf{AB} + \operatorname{tr} \mathbf{ABC} + \operatorname{tr} \mathbf{CBA}). \end{aligned} \quad (4.11)$$

Taking the traces of both sides of equation (4.11), we obtain

$$\begin{aligned} & \operatorname{tr} \mathbf{ABC} + \operatorname{tr} \mathbf{CBA} - \operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{BC} - \operatorname{tr} \mathbf{B} \operatorname{tr} \mathbf{CA} \\ & - \operatorname{tr} \mathbf{C} \operatorname{tr} \mathbf{AB} + \operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{B} \operatorname{tr} \mathbf{C} = 0. \end{aligned} \quad (4.12)$$

Introducing (4.12) into (4.11), we obtain

$$\begin{aligned} & \mathbf{ABC} + \mathbf{CBA} + \mathbf{BAC} + \mathbf{CAB} + \mathbf{ACB} + \mathbf{BCA} \\ &= (\mathbf{BC} + \mathbf{CB}) \operatorname{tr} \mathbf{A} + (\mathbf{CA} + \mathbf{AC}) \operatorname{tr} \mathbf{B} + (\mathbf{AB} + \mathbf{BA}) \operatorname{tr} \mathbf{C} \\ &+ \mathbf{A}(\operatorname{tr} \mathbf{BC} - \operatorname{tr} \mathbf{B} \operatorname{tr} \mathbf{C}) + \mathbf{B}(\operatorname{tr} \mathbf{CA} - \operatorname{tr} \mathbf{C} \operatorname{tr} \mathbf{A}) \\ &+ \mathbf{C}(\operatorname{tr} \mathbf{AB} - \operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{B}). \end{aligned} \quad (4.13)$$

This result may also be obtained directly from equation (4.7) and two similar equations obtained from it by replacing $\mathbf{A}$, $\mathbf{B}$ by $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{C}$, $\mathbf{A}$ respectively, as shown in §5.

The left-hand side of equation (4.13) may be written as

$$\mathbf{ABC} + \mathbf{CBA} + [\mathbf{B}(\mathbf{AC}) + (\mathbf{AC})\mathbf{B}] + [\mathbf{B}(\mathbf{CA}) + (\mathbf{CA})\mathbf{B}]. \quad (4.14)$$

Introducing relations of the type (4.7) into (4.13), we have

$$\begin{aligned} \mathbf{ABC} + \mathbf{CBA} &= \mathbf{A} \operatorname{tr} \mathbf{BC} + \mathbf{C} \operatorname{tr} \mathbf{AB} + \mathbf{B}(\operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{C} - \operatorname{tr} \mathbf{AC}) \\ &+ \mathbf{I} \operatorname{tr} \mathbf{B}(\operatorname{tr} \mathbf{AC} - \operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{C}). \end{aligned} \quad (4.15)$$

Multiplying both sides of equation (4.7) on the right by $\mathbf{C}$, we have

$$\mathbf{ABC} + \mathbf{BAC} = \mathbf{AC} \operatorname{tr} \mathbf{B} + \mathbf{BC} \operatorname{tr} \mathbf{A} - \mathbf{C}(\operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{B} - \operatorname{tr} \mathbf{AB}). \quad (4.16)$$

Thus,

$$\begin{aligned} \mathbf{ABC} - \mathbf{CBA} &= -(\mathbf{BA})\mathbf{C} - \mathbf{C}(\mathbf{BA}) + \mathbf{AC} \operatorname{tr} \mathbf{B} + \mathbf{BC} \operatorname{tr} \mathbf{A} \\ &- \mathbf{C}(\operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{B} - \operatorname{tr} \mathbf{AB}). \end{aligned} \quad (4.17)$$

Employing a relation of the type (4.7) to substitute for $(\mathbf{B}\mathbf{A})\mathbf{C} + \mathbf{C}(\mathbf{B}\mathbf{A})$ in (4.17), we obtain

$$\begin{aligned} \mathbf{ABC} - \mathbf{CBA} = & -\mathbf{BA} \operatorname{tr} \mathbf{C} + \mathbf{AC} \operatorname{tr} \mathbf{B} + \mathbf{BC} \operatorname{tr} \mathbf{A} \\ & - \mathbf{C} \operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{B} - \mathbf{I}(\operatorname{tr} \mathbf{CBA} - \operatorname{tr} \mathbf{BA} \operatorname{tr} \mathbf{C}). \end{aligned} \quad (4.18)$$

From (4.18) and (4.15), we obtain

$$\begin{aligned} 2\mathbf{ABC} = & -\mathbf{BA} \operatorname{tr} \mathbf{C} + \mathbf{AC} \operatorname{tr} \mathbf{B} + \mathbf{BC} \operatorname{tr} \mathbf{A} + \mathbf{A} \operatorname{tr} \mathbf{BC} + \mathbf{B}(\operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{C} \\ & - \operatorname{tr} \mathbf{AC}) + \mathbf{C}(\operatorname{tr} \mathbf{AB} - \operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{B}) + \mathbf{I}(\operatorname{tr} \mathbf{B} \operatorname{tr} \mathbf{AC} \\ & - \operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{B} \operatorname{tr} \mathbf{C} - \operatorname{tr} \mathbf{CBA} + \operatorname{tr} \mathbf{BA} \operatorname{tr} \mathbf{C}). \end{aligned} \quad (4.19)$$

The Case of 3×3 Matrices. If $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are 3×3 matrices, then (4.4) becomes

$$\epsilon_{pqrs} a_{jpj_1} b_{jqj_2} c_{jrj_3} \delta_{i_4j_4} = 0. \quad (4.20)$$

It can readily be shown in a manner similar to that adopted in the case of 2×2 matrices that this relation is equivalent to

$$\begin{aligned} (\mathbf{ABC} + \mathbf{ACB} + \mathbf{BCA} + \mathbf{BAC} + \mathbf{CAB} + \mathbf{CBA}) - \mathbf{A}(\operatorname{tr} \mathbf{BC} \\ - \operatorname{tr} \mathbf{B} \operatorname{tr} \mathbf{C}) - \mathbf{B}(\operatorname{tr} \mathbf{CA} - \operatorname{tr} \mathbf{C} \operatorname{tr} \mathbf{A}) - \mathbf{C}(\operatorname{tr} \mathbf{AB} - \operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{B}) \\ - (\mathbf{BC} + \mathbf{CB}) \operatorname{tr} \mathbf{A} - (\mathbf{CA} + \mathbf{AC}) \operatorname{tr} \mathbf{B} - (\mathbf{AB} + \mathbf{BA}) \operatorname{tr} \mathbf{C} \\ - \mathbf{I}(\operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{B} \operatorname{tr} \mathbf{C} - \operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{BC} - \operatorname{tr} \mathbf{B} \operatorname{tr} \mathbf{CA} \\ - \operatorname{tr} \mathbf{C} \operatorname{tr} \mathbf{AB} + \operatorname{tr} \mathbf{ABC} + \operatorname{tr} \mathbf{CBA}) = 0. \end{aligned} \quad (4.21)$$

If in (4.21) we write $\mathbf{C} = \mathbf{A}$, we obtain

$$\begin{aligned} \mathbf{ABA} + \mathbf{A}^2 \mathbf{B} + \mathbf{BA}^2 - \mathbf{A}[\operatorname{tr} \mathbf{AB} - \operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{B}] - \frac{1}{2} \mathbf{B}[\operatorname{tr} \mathbf{A}^2 - (\operatorname{tr} \mathbf{A})^2] \\ - (\mathbf{AB} + \mathbf{BA}) \operatorname{tr} \mathbf{A} - \mathbf{A}^2 \operatorname{tr} \mathbf{B} \\ - \mathbf{I}[\operatorname{tr} \mathbf{A}^2 \mathbf{B} - \operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{AB} + \frac{1}{2} \operatorname{tr} \mathbf{B}\{(\operatorname{tr} \mathbf{A})^2 - \operatorname{tr} \mathbf{A}^2\}] = 0. \end{aligned} \quad (4.22)$$

From equation (4.22), we have, multiplying throughout on the left and right by $\mathbf{A}$ and adding the equations so obtained,

$$\begin{aligned} 2(\mathbf{ABA}^2 + \mathbf{A}^2 \mathbf{BA}) = & -(\mathbf{A}^3 \mathbf{B} + \mathbf{BA}^3) + (\mathbf{A}^2 \mathbf{B} + \mathbf{BA}^2) \operatorname{tr} \mathbf{A} + 2\mathbf{ABA} \operatorname{tr} \mathbf{A} \\ & + 2(\mathbf{A}^3 \operatorname{tr} \mathbf{B} - 2\mathbf{A}^2(\operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{B} - \operatorname{tr} \mathbf{AB}) - \frac{1}{2}(\mathbf{AB} + \mathbf{BA})\{(\operatorname{tr} \mathbf{A})^2 - \operatorname{tr} \mathbf{A}^2\} \\ & + 2\mathbf{A}[\operatorname{tr} \mathbf{A}^2 \mathbf{B} - \operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{AB} + \frac{1}{2} \operatorname{tr} \mathbf{B}\{(\operatorname{tr} \mathbf{A})^2 - \operatorname{tr} \mathbf{A}^2\}]. \end{aligned} \quad (4.23)$$

Employing the Hamilton-Cayley theorem

$$\mathbf{A}^3 - \mathbf{A}^2 \operatorname{tr} \mathbf{A} + \frac{1}{2} \mathbf{A}[(\operatorname{tr} \mathbf{A})^2 - \operatorname{tr} \mathbf{A}^2] - \mathbf{I} \det \mathbf{A} = 0 \quad (4.24)$$

to substitute for $\mathbf{A}^3$ in equation (4.23), we obtain

$$\begin{aligned} \mathbf{A} \mathbf{B} \mathbf{A}^2 + \mathbf{A}^2 \mathbf{B} \mathbf{A} &= \mathbf{A} \mathbf{B} \mathbf{A} \operatorname{tr} \mathbf{A} + \mathbf{A}^2 \operatorname{tr} \mathbf{A} \mathbf{B} + \mathbf{A}(\operatorname{tr} \mathbf{A}^2 \mathbf{B} - \operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{A} \mathbf{B}) \\ &\quad - \mathbf{B} \det \mathbf{A} + \mathbf{I} \det \mathbf{A} \operatorname{tr} \mathbf{B}. \end{aligned} \quad (4.25)$$

Multiplying equation (4.25) on the left and right by $\mathbf{A}$ and employing the Hamilton-Cayley theorem (4.24), we obtain

$$\begin{aligned} \mathbf{A}^2 \mathbf{B} \mathbf{A}^2 &= \frac{1}{2} \mathbf{A} \mathbf{B} \mathbf{A}[(\operatorname{tr} \mathbf{A})^2 - \operatorname{tr} \mathbf{A}^2] - (\mathbf{A} \mathbf{B} + \mathbf{B} \mathbf{A}) \det \mathbf{A} + \mathbf{A}^2 \operatorname{tr} \mathbf{A}^2 \mathbf{B} \\ &\quad + \mathbf{A}[\det \mathbf{A} \operatorname{tr} \mathbf{B} - \frac{1}{2}\{(\operatorname{tr} \mathbf{A})^2 - \operatorname{tr} \mathbf{A}^2\} \operatorname{tr} \mathbf{A} \mathbf{B}] + \mathbf{I} \det \mathbf{A} \operatorname{tr} \mathbf{A} \mathbf{B}. \end{aligned} \quad (4.26)$$

It can readily be seen that (4.7) and (4.21) are generalizations of the Hamilton-Cayley theorem for 2×2 and 3×3 matrices respectively. Writing $\mathbf{A} = \mathbf{B}$ in (4.7), we obtain

$$2\mathbf{A}^2 - 2\mathbf{A} \operatorname{tr} \mathbf{A} + [(\operatorname{tr} \mathbf{A})^2 - \operatorname{tr} \mathbf{A}^2]\mathbf{I} = 0, \quad (4.27)$$

which is the Hamilton-Cayley theorem for a 2×2 matrix. Again, writing $\mathbf{A} = \mathbf{B} = \mathbf{C}$ in (4.21), we obtain equation (4.24), which expresses the Hamilton-Cayley theorem for a 3×3 matrix.

It is thus apparent that the relations (4.7) and (4.21) express generalizations of the Hamilton-Cayley theorem for 2×2 and 3×3 matrices respectively. The corresponding generalization of the Hamilton-Cayley theorem for $n \times n$ matrices could be obtained in an analogous manner to (4.7) and (4.21) from the relation (4.2).

5. An alternative derivation of the relations in §4 for 2×2 matrices. In §4, the relations (4.7) and (4.13) for 2×2 matrices have been obtained from first principles and it has been shown that the Hamilton-Cayley theorem for a 2×2 matrix may be obtained as a particular case of (4.7). It can also be shown that equations (4.7) and (4.13) are obtainable from the Hamilton-Cayley theorem for a 2×2 matrix. Thus, for a 2×2 matrix $\mathbf{D}$, the Hamilton-Cayley theorem may be written as

$$\mathbf{D}^2 - (\operatorname{tr} \mathbf{D})\mathbf{D} + \frac{1}{2}[(\operatorname{tr} \mathbf{D})^2 - \operatorname{tr} \mathbf{D}^2]\mathbf{I} = 0. \quad (5.1)$$

We consider two 2×2 matrices $\mathbf{A}$ and $\mathbf{B}$ and we write $\mathbf{D} = \mathbf{A} + \mathbf{B}$ in equation (5.1). We obtain

$$\begin{aligned} \mathbf{A} \mathbf{B} + \mathbf{B} \mathbf{A} + \mathbf{A}^2 + \mathbf{B}^2 - (\operatorname{tr} \mathbf{A} + \operatorname{tr} \mathbf{B})(\mathbf{A} + \mathbf{B}) \\ + \frac{1}{2}[(\operatorname{tr} \mathbf{A} + \operatorname{tr} \mathbf{B})^2 - \operatorname{tr} \mathbf{A}^2 - \operatorname{tr} \mathbf{B}^2 - 2 \operatorname{tr} \mathbf{A} \mathbf{B}]\mathbf{I} = 0. \end{aligned} \quad (5.2)$$

Again, if we take $\mathbf{D} = \mathbf{A} - \mathbf{B}$ in (5.1), we obtain

$$\begin{aligned}
 -\mathbf{AB} - \mathbf{BA} + \mathbf{A}^2 + \mathbf{B}^2 - (\text{tr } \mathbf{A} - \text{tr } \mathbf{B})(\mathbf{A} - \mathbf{B}) \\
 + \frac{1}{2}[(\text{tr } \mathbf{A} - \text{tr } \mathbf{B})^2 - \text{tr } \mathbf{A}^2 - \text{tr } \mathbf{B}^2 + 2 \text{tr } \mathbf{AB}]\mathbf{I} = 0.
 \end{aligned} \tag{5.3}$$

Subtracting (5.3) from (5.2), we obtain equation (4.7).

If $\mathbf{A}, \mathbf{B}, \mathbf{C}$ are three 2×2 matrices, we have, from (4.7),

$$\begin{aligned}
 \mathbf{AB} + \mathbf{BA} - \mathbf{A} \text{tr } \mathbf{B} - \mathbf{B} \text{tr } \mathbf{A} + (\text{tr } \mathbf{A} \text{tr } \mathbf{B} - \text{tr } \mathbf{AB})\mathbf{I} &= 0, \\
 \mathbf{BC} + \mathbf{CB} - \mathbf{B} \text{tr } \mathbf{C} - \mathbf{C} \text{tr } \mathbf{B} + (\text{tr } \mathbf{B} \text{tr } \mathbf{C} - \text{tr } \mathbf{BC})\mathbf{I} &= 0
 \end{aligned} \tag{5.4}$$

and

$$\mathbf{CA} + \mathbf{AC} - \mathbf{C} \text{tr } \mathbf{A} - \mathbf{A} \text{tr } \mathbf{C} + (\text{tr } \mathbf{C} \text{tr } \mathbf{A} - \text{tr } \mathbf{CA})\mathbf{I} = 0.$$

Multiplying the first, second and third of equations (5.4) on the left by $\mathbf{C}$, $\mathbf{A}$ and $\mathbf{B}$ respectively, and adding the equations so obtained we obtain equation (4.13). We may then proceed as in §4 to obtain equation (4.19).

6. An alternative derivation of the relations in §4 for 3×3 matrices. In §4, the relation (4.21) is obtained from first principles and it has been pointed out that the Hamilton-Cayley theorem for a 3×3 matrix is a particular case of (4.21). Equation (4.21) can also be derived directly from the Hamilton-Cayley theorem for a 3×3 matrix. Thus, for a 3×3 matrix $\mathbf{D}$, the Hamilton-Cayley theorem may be written as

$$\begin{aligned}
 \mathbf{D}^3 - (\text{tr } \mathbf{D})\mathbf{D}^2 + \frac{1}{2}[(\text{tr } \mathbf{D})^2 - \text{tr } \mathbf{D}^2]\mathbf{D} \\
 - [\frac{1}{3} \text{tr } \mathbf{D}^3 - \frac{1}{2} \text{tr } \mathbf{D} \text{tr } \mathbf{D}^2 + \frac{1}{6}(\text{tr } \mathbf{D})^3]\mathbf{I} = 0.
 \end{aligned} \tag{6.1}$$

Taking $\mathbf{D} = \mathbf{A} + \mathbf{B}$ and $\mathbf{D} = \mathbf{A} - \mathbf{B}$ respectively in (6.1) and subtracting the resulting equations, we obtain

$$\begin{aligned}
 2(\mathbf{ABA} + \mathbf{A}^2\mathbf{B} + \mathbf{BA}^2 + \mathbf{B}^3) - 2(\mathbf{AB} + \mathbf{BA})\text{tr } \mathbf{A} - 2(\mathbf{A}^2 + \mathbf{B}^2)\text{tr } \mathbf{B} \\
 + 2\mathbf{A}(\text{tr } \mathbf{A} \text{tr } \mathbf{B} - \text{tr } \mathbf{AB}) + \mathbf{B}[(\text{tr } \mathbf{A})^2 + (\text{tr } \mathbf{B})^2 - \text{tr } \mathbf{A}^2 - \text{tr } \mathbf{B}^2] \\
 - [\frac{2}{3}(\text{tr } \mathbf{B}^3 + \text{tr } \mathbf{ABA} + \text{tr } \mathbf{BA}^2 + \text{tr } \mathbf{A}^2\mathbf{B}) + \frac{1}{3}\{3(\text{tr } \mathbf{A})^2 + (\text{tr } \mathbf{B})^2\}\text{tr } \mathbf{B} \\
 - \text{tr } \mathbf{B}(\text{tr } \mathbf{A}^2 + \text{tr } \mathbf{B}^2) - 2 \text{tr } \mathbf{A} \text{tr } \mathbf{AB}]\mathbf{I} = 0.
 \end{aligned} \tag{6.2}$$

Now, since $\text{tr } \mathbf{ABA} = \text{tr } \mathbf{A}^2\mathbf{B}$, we obtain from (6.2), together with the Hamilton-Cayley theorem for the matrix $\mathbf{B}$, the relation (4.22). Then, taking $\mathbf{D} = \mathbf{A} + \mathbf{B} + \mathbf{C}$ in (6.1) and employing relations of the types (4.22) and (5.1), we obtain the relation (4.21).

7. Further relations for 3×3 matrices. If $\mathbf{A}$ is a 3×3 matrix, then the determinant

$$\begin{vmatrix} a_{i_1 j_1} & a_{i_2 j_1} & a_{i_3 j_1} \\ a_{i_1 j_2} & a_{i_2 j_2} & a_{i_3 j_2} \\ a_{i_1 j_3} & a_{i_2 j_3} & a_{i_3 j_3} \end{vmatrix} \quad (7.1)$$

is readily seen to be antisymmetric in the i 's and in the j 's. We thus see that

$$\begin{vmatrix} a_{i_1 j_1} & a_{i_2 j_1} & a_{i_3 j_1} \\ a_{i_1 j_2} & a_{i_2 j_2} & a_{i_3 j_2} \\ a_{i_1 j_3} & a_{i_2 j_3} & a_{i_3 j_3} \end{vmatrix} = \lambda \begin{vmatrix} \delta_{i_1 j_1} & \delta_{i_2 j_1} & \delta_{i_3 j_1} \\ \delta_{i_1 j_2} & \delta_{i_2 j_2} & \delta_{i_3 j_2} \\ \delta_{i_1 j_3} & \delta_{i_2 j_3} & \delta_{i_3 j_3} \end{vmatrix}, \quad (7.2)$$

where λ is a scalar. From (7.2) it follows that

$$\begin{aligned} & \begin{vmatrix} a_{i_1 j_1} & a_{i_2 j_1} & a_{i_3 j_1} \\ a_{i_1 j_2} & a_{i_2 j_2} & a_{i_3 j_2} \\ a_{i_1 j_3} & a_{i_2 j_3} & a_{i_3 j_3} \end{vmatrix} \times \begin{vmatrix} \delta_{i_1 j_1} & \delta_{i_1 j_2} & \delta_{i_1 j_3} \\ \delta_{i_2 j_1} & \delta_{i_2 j_2} & \delta_{i_2 j_3} \\ \delta_{i_3 j_1} & \delta_{i_3 j_2} & \delta_{i_3 j_3} \end{vmatrix} \\ &= \lambda \begin{vmatrix} \delta_{i_1 j_1} & \delta_{i_2 j_1} & \delta_{i_3 j_1} \\ \delta_{i_1 j_2} & \delta_{i_2 j_2} & \delta_{i_3 j_2} \\ \delta_{i_1 j_3} & \delta_{i_2 j_3} & \delta_{i_3 j_3} \end{vmatrix} \times \begin{vmatrix} \delta_{i_1 j_1} & \delta_{i_1 j_2} & \delta_{i_1 j_3} \\ \delta_{i_2 j_1} & \delta_{i_2 j_2} & \delta_{i_2 j_3} \\ \delta_{i_3 j_1} & \delta_{i_3 j_2} & \delta_{i_3 j_3} \end{vmatrix}, \end{aligned} \quad (7.3)$$

whence

$$\begin{vmatrix} \text{tr } \mathbf{A} & a_{j_2 j_1} & a_{j_3 j_1} \\ a_{j_1 j_2} & \text{tr } \mathbf{A} & a_{j_3 j_2} \\ a_{j_1 j_3} & a_{j_2 j_3} & \text{tr } \mathbf{A} \end{vmatrix} = \lambda \begin{vmatrix} 3 & \delta_{j_2 j_1} & \delta_{j_3 j_1} \\ \delta_{j_1 j_2} & 3 & \delta_{j_3 j_2} \\ \delta_{j_1 j_3} & \delta_{j_2 j_3} & 3 \end{vmatrix};$$

i.e.

$$\lambda = \frac{1}{6}[(\text{tr } \mathbf{A})^3 - 3 \text{tr } \mathbf{A} \text{tr } \mathbf{A}^2 + 2 \text{tr } \mathbf{A}^3]. \quad (7.4)$$

If $\mathbf{B}$ and $\mathbf{C}$ are two 3×3 matrices, we have, from (7.2),

$$\begin{vmatrix} a_{i_1 j_1} & a_{i_2 j_1} & a_{i_3 j_1} \\ a_{i_1 j_2} & a_{i_2 j_2} & a_{i_3 j_2} \\ a_{i_1 j_3} & a_{i_2 j_3} & a_{i_3 j_3} \end{vmatrix} b_{j_1 i_2} c_{j_2 i_3} = \lambda \begin{vmatrix} \delta_{i_1 j_1} & \delta_{i_2 j_1} & \delta_{i_3 j_1} \\ \delta_{i_1 j_2} & \delta_{i_2 j_2} & \delta_{i_3 j_2} \\ \delta_{i_1 j_3} & \delta_{i_2 j_3} & \delta_{i_3 j_3} \end{vmatrix} b_{j_1 i_2} c_{j_2 i_3}, \quad (7.5)$$

where λ is given by (7.4). Equation (7.5) may be re-written as

$$\begin{aligned} & \mathbf{ABACA} + \mathbf{ACABA} + (\text{tr } \mathbf{AB})(\text{tr } \mathbf{AC})\mathbf{A} - (\text{tr } \mathbf{AB})\mathbf{ACA} - (\text{tr } \mathbf{AC})\mathbf{ABA} \\ & - (\text{tr } \mathbf{ABAC})\mathbf{A} = \frac{1}{6}[(\text{tr } \mathbf{A})^3 - 3 \text{tr } \mathbf{A} \text{tr } \mathbf{A}^2 + 2 \text{tr } \mathbf{A}^3] \\ & \cdot [\mathbf{BC} + \mathbf{CB} - \mathbf{B} \text{tr } \mathbf{C} - \mathbf{C} \text{tr } \mathbf{B} + (\text{tr } \mathbf{B} \text{tr } \mathbf{C} - \text{tr } \mathbf{BC})\mathbf{I}]. \end{aligned} \quad (7.6)$$

Substituting $\mathbf{B}^2$ for $\mathbf{C}$ in (7.6), we obtain

$$\begin{aligned} \mathbf{ABAB}^2\mathbf{A} + \mathbf{AB}^2\mathbf{ABA} &= -(\operatorname{tr} \mathbf{AB})(\operatorname{tr} \mathbf{AB}^2)\mathbf{A} + (\operatorname{tr} \mathbf{AB})\mathbf{AB}^2\mathbf{A} \\ &+ (\operatorname{tr} \mathbf{AB}^2)\mathbf{ABA} + (\operatorname{tr} \mathbf{ABAB}^2)\mathbf{A} + \frac{1}{6}[(\operatorname{tr} \mathbf{A})^3 - 3 \operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{A}^2 \\ &+ 2 \operatorname{tr} \mathbf{A}^3][2\mathbf{B}^3 - \mathbf{B}^2 \operatorname{tr} \mathbf{B} - \mathbf{B} \operatorname{tr} \mathbf{B}^2 + (\operatorname{tr} \mathbf{B} \operatorname{tr} \mathbf{B}^2 - \operatorname{tr} \mathbf{B}^3)\mathbf{I}]. \end{aligned} \quad (7.7)$$

8. The contraction of polynomials in two 2×2 matrices. From the result given in §3, it follows that any polynomial in two 2×2 matrices $\mathbf{A}$ and $\mathbf{B}$ may be expressed as a polynomial in $\mathbf{A}$ and $\mathbf{B}$ which is the sum of terms of the type $\mathbf{A}^{\alpha_1} \mathbf{BA} \cdots \mathbf{B}^{\beta_n}$, where α_1 and β_n may be either 1 or 0, with scalar coefficients expressible as polynomials in the elements of $\mathbf{A}$ and $\mathbf{B}$. Now, employing the Hamilton-Cayley theorem we can express $(\mathbf{BA})^{n-1}$ in the form

$$(\mathbf{BA})^{n-1} = f_{n-1} \mathbf{BA} + f'_{n-1} \mathbf{I}, \quad (8.1)$$

where f_{n-1} and f'_{n-1} are scalars expressible as polynomials in the elements of $\mathbf{BA}$. Thus,

$$\mathbf{A}^{\alpha_1} \mathbf{BAB} \cdots \mathbf{B}^{\beta_n} = f_{n-1} \mathbf{A}^{\alpha_1} \mathbf{BAB}^{\beta_n} + f'_{n-1} \mathbf{A}^{\alpha_1} \mathbf{B}^{\beta_n}. \quad (8.2)$$

If $\alpha_1 = \beta_n = 0$, then

$$\mathbf{A}^{\alpha_1} \mathbf{BAB} \cdots \mathbf{B}^{\beta_n} = (\mathbf{BA})^{n-1} = f_{n-1} \mathbf{BA} + f'_{n-1} \mathbf{I}. \quad (8.3)$$

If $\alpha_1 = \beta_1 = 1$, then

$$\mathbf{A}^{\alpha_1} \mathbf{BAB} \cdots \mathbf{B}^{\beta_n} = (\mathbf{AB})^n = f_n \mathbf{AB} + f'_n \mathbf{I}. \quad (8.4)$$

If $\alpha_1 = 1$ and $\beta_n = 0$, then

$$\mathbf{A}^{\alpha_1} \mathbf{BAB} \cdots \mathbf{B}^{\beta_n} = f_{n-1} \mathbf{ABA} + f'_{n-1} \mathbf{A}. \quad (8.5)$$

If $\alpha_1 = 0$ and $\beta_n = 1$, then

$$\mathbf{A}^{\alpha_1} \mathbf{BAB} \cdots \mathbf{B}^{\beta_n} = f_{n-1} \mathbf{BAB} + f'_{n-1} \mathbf{B}. \quad (8.6)$$

Employing the relations (8.3) to (8.6), it is clear that any polynomial in $\mathbf{A}$ and $\mathbf{B}$ may be expressed as a polynomial of extension ≤ 3 of the form

$$\varphi_0 \mathbf{I} + \varphi_1 \mathbf{A} + \varphi_2 \mathbf{B} + \varphi_3 \mathbf{AB} + \varphi_4 \mathbf{BA} + \varphi_5 \mathbf{ABA} + \varphi_6 \mathbf{BAB}, \quad (8.7)$$

where the φ 's are scalars expressible as polynomials in the elements of $\mathbf{A}$ and $\mathbf{B}$. From (4.7), we have

$$\mathbf{ABA} = -\mathbf{BA}^2 + \mathbf{A}^2 \operatorname{tr} \mathbf{B} + \mathbf{BA} \operatorname{tr} \mathbf{A} - (\operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{B} - \operatorname{tr} \mathbf{AB})\mathbf{A}$$

and

$$\mathbf{BAB} = -\mathbf{B}^2 \mathbf{A} + \mathbf{BA} \operatorname{tr} \mathbf{B} + \mathbf{B}^2 \operatorname{tr} \mathbf{A} - (\operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{B} - \operatorname{tr} \mathbf{AB})\mathbf{B}. \quad (8.8)$$

Employing the Hamilton-Cayley theorem in the forms

$$\mathbf{A}^2 - \mathbf{A} \operatorname{tr} \mathbf{A} + \mathbf{I} \det \mathbf{A} = 0 \quad \text{and} \quad \mathbf{B}^2 - \mathbf{B} \operatorname{tr} \mathbf{B} + \mathbf{I} \det \mathbf{B} = 0 \quad (8.9)$$

to substitute for $\mathbf{A}^2$ and $\mathbf{B}^2$ in (8.8), and using the resulting relations to substitute for $\mathbf{ABA}$ and $\mathbf{BAB}$ in (8.7), we obtain

Lemma 1. *Any polynomial in two 2×2 matrices may be expressed as a polynomial of extension 2 in these matrices and of first degree in each of them.*

Consider the polynomial $\mathbf{P}$ in $\mathbf{A}$ and $\mathbf{B}$. From Lemma 1, it follows that we may write $\mathbf{P}$ in the form

$$\mathbf{P} = \varphi_0 \mathbf{I} + \varphi_1 \mathbf{A} + \varphi_2 \mathbf{B} + \varphi_3 \mathbf{AB} + \varphi_4 \mathbf{BA}, \quad (8.10)$$

where the φ 's are scalars expressible as polynomials in the elements of $\mathbf{A}$ and $\mathbf{B}$. Then, if $\mathbf{A}$ and $\mathbf{B}$ are symmetric matrices, the transpose $\mathbf{P}^*$ of $\mathbf{P}$ is given by

$$\mathbf{P}^* = \varphi_0 \mathbf{I} + \varphi_1 \mathbf{A} + \varphi_2 \mathbf{B} + \varphi_3 \mathbf{BA} + \varphi_4 \mathbf{AB}. \quad (8.11)$$

If $\mathbf{P}$ is symmetric, then by definition $\mathbf{P} = \mathbf{P}^*$, so that

$$(\varphi_3 - \varphi_4)\mathbf{AB} = (\varphi_3 - \varphi_4)\mathbf{BA}.$$

Whence we obtain

$$\varphi_3 = \varphi_4 \quad \text{or} \quad \mathbf{AB} = \mathbf{BA}. \quad (8.12)$$

In either case, $\mathbf{P}$ may be written

$$\mathbf{P} = \varphi_0 \mathbf{I} + \varphi_1 \mathbf{A} + \varphi_2 \mathbf{B} + \frac{1}{2}(\varphi_3 + \varphi_4)(\mathbf{AB} + \mathbf{BA}). \quad (8.13)$$

Employing the relation (4.7), we obtain

Lemma 2. *Any symmetric polynomial in two symmetric 2×2 matrices may be expressed as a polynomial in the matrices of extension 1 and of first degree.*

9. The contraction of polynomials in $R, 2 \times 2$ matrices. We shall now consider polynomials in $R, 2 \times 2$ matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$. From §3 it follows that any polynomial in these matrices can be expressed as a polynomial in the matrices which is the sum of terms of the type

$$\mathbf{A}^{\alpha_1} \mathbf{B}^{\beta_1} \mathbf{C}^{\gamma_1} \dots \mathbf{A}^{\alpha_2} \mathbf{B}^{\beta_2} \mathbf{C}^{\gamma_2} \dots, \quad (9.1)$$

with scalar coefficients which are expressible as polynomials in the elements of $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$, where the α 's, β 's, γ 's etc. take the values 0 or 1, but no two adjacent terms in the product involve the same matrix.

Let us consider a term of the type (9.1) which has extension $n > R$. Then, at least one of the matrices $\mathbf{A}$ (say) must be repeated in this product. We can then write the product in the form

$$\mathbf{YAXAZ}, \quad (9.2)$$

where $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ are themselves products of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\dots$ of the type (9.1). The extension of $\mathbf{X}$ is ≥ 1 and $\leq n - 2$, while the extensions of $\mathbf{Y}$ and $\mathbf{Z}$ are $\leq n - 3$. Let n_1 , n_2 and n_3 be the extensions of $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ respectively in the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\dots$. Then,

$$n_1 + n_2 + n_3 = n - 2. \quad (9.3)$$

From equation (4.15), we have

$$\mathbf{A}\mathbf{X}\mathbf{A} = \varphi_0 \mathbf{I} + \varphi_1 \mathbf{A} + \varphi_2 \mathbf{X}, \quad (9.4)$$

where the φ 's are scalars expressible as polynomials in the elements of $\mathbf{A}$ and $\mathbf{X}$.

Introducing (9.4) into (9.2), we see that

$$\mathbf{Y}\mathbf{A}\mathbf{X}\mathbf{A}\mathbf{Z} = \varphi_0 \mathbf{Y}\mathbf{Z} + \varphi_1 \mathbf{Y}\mathbf{A}\mathbf{Z} + \varphi_2 \mathbf{Y}\mathbf{X}\mathbf{Z}. \quad (9.5)$$

Now, the expression on the right hand side of (9.5), considered as a polynomial in the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\dots$, has extension $\leq n_1 + n_2 + n_3$, i.e. $\leq n - 2$. Thus, any polynomial in the R matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\dots$ of extension n ($> R$) may be expressed as a polynomial of extension $\leq n - 2$ and, by a repetition of this process, we obtain

Lemma 3. *Any polynomial in R , 2×2 matrices may be expressed as a polynomial of extension $\leq R$ and of first degree in each of the matrices.*

It follows from equation (4.19) that any product of three 2×2 matrices of degree 1 in each of them may be expressed as a polynomial of extension 2 in the three matrices and of degree 1 in each of them. With Lemma 3, we obtain immediately:

Theorem 1. *Any polynomial in any number of 2×2 matrices may be expressed as a polynomial of extension ≤ 2 in the matrices and of degree < 2 in each of them.*

If the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\dots$ are symmetric matrices, then we can proceed in a manner similar to that adopted in proving lemma 2 to prove the

Corollary. *Any symmetric polynomial in any number of 2×2 matrices may be expressed as a polynomial of extension ≤ 1 in the matrices and of degree less than two in each of them.*

10. The contraction of polynomials in two 3×3 matrices. If $\mathbf{A}$ ($= || a_{ij} ||$) and $\mathbf{B}$ ($= || b_{ij} ||$) are two 3×3 matrices, it is seen from §3 that the most general polynomial in $\mathbf{A}$ and $\mathbf{B}$ may be expressed as a polynomial in these matrices in which each term is of the form

$$\mathbf{A}^{\alpha_1} \mathbf{B}^{\beta_1} \mathbf{A}^{\alpha_2} \mathbf{B}^{\beta_2} \dots \mathbf{A}^{\alpha_r} \mathbf{B}^{\beta_r}, \quad (10.1)$$

where $\alpha_2, \dots, \alpha_r$ and $\beta_1, \dots, \beta_r$ may take the values 1 or 2 and α_1 and β_r may take the values 0, 1 or 2. Each such term in the polynomial has a coefficient which is a scalar expressible as a polynomial in the elements of $\mathbf{A}$ and $\mathbf{B}$.

Let us consider a matrix product $\mathbf{\Pi}$ of the form (10.1). If the extension of this product is greater than or equal to 5 then at least two of the indices α_1 ,

$\alpha_2, \dots, \alpha_r$ or two of the indices $\beta_1, \beta_2, \dots, \beta_r$ are equal. Suppose, for example, $\alpha_m = \alpha_n$. Then, we may write

$$\mathbf{\Pi} = \mathbf{A}^{\alpha_1} \mathbf{B}^{\beta_1} \dots \mathbf{A}^{\alpha_{m-1}} \mathbf{B}^{\beta_{m-1}} \mathbf{A}^{\alpha_m} (\mathbf{B}^{\beta_m} \mathbf{A}^{\alpha_{m+1}} \dots \mathbf{A}^{\alpha_{n-1}} \mathbf{B}^{\beta_{n-1}}) \cdot \mathbf{A}^{\alpha_n} \mathbf{B}^{\beta_n} \dots \mathbf{A}^{\alpha_r} \mathbf{B}^{\beta_r}, \quad (10.2)$$

with $\alpha_m = \alpha_n = 1$ or 2 . We may then employ a formula of the type (4.22) to express $\mathbf{A}^{\alpha_m} (\mathbf{B}^{\beta_m} \mathbf{A}^{\alpha_{m+1}} \dots \mathbf{A}^{\alpha_{n-1}} \mathbf{B}^{\beta_{n-1}}) \mathbf{A}^{\alpha_n}$, which has extension $2(n - m) + 1$, as a polynomial of extension $2(n - m)$ in $\mathbf{A}$ and $\mathbf{B}$. By repetition of this process, we obtain

Lemma 4. *Any matrix product in two 3×3 matrices may be expressed as a polynomial in these matrices of extension 4 or less.*

We have immediately the

Corollary. *Any polynomial in two 3×3 matrices may be expressed as a polynomial in these matrices of extension 4 or less.*

We shall now consider matrix products of $\mathbf{A}$ and $\mathbf{B}$ of extension 4 in which the indices take the values 1 or 2. A product $\mathbf{\Pi}$ of this type may be written either as

$$\mathbf{\Pi} = \mathbf{A}^{\alpha_1} \mathbf{B}^{\beta_1} \mathbf{A}^{\alpha_2} \mathbf{B}^{\beta_2} \quad \text{or} \quad \mathbf{B}^{\beta_1} \mathbf{A}^{\alpha_1} \mathbf{B}^{\beta_2} \mathbf{A}^{\alpha_2}, \quad (10.3)$$

where $\alpha_1, \beta_1, \alpha_2$ and β_2 may take the values 1 or 2. If $\alpha_1 = \alpha_2$ or $\beta_1 = \beta_2$, then we may use a formula of the type (4.22) to express $\mathbf{\Pi}$ as a polynomial of extension 3 in $\mathbf{A}$ and $\mathbf{B}$. If $\alpha_1 \neq \alpha_2$ and $\beta_1 \neq \beta_2$, then

$$\begin{aligned} \mathbf{\Pi} = & \mathbf{ABA}^2 \mathbf{B}^2, \mathbf{BAB}^2 \mathbf{A}^2, \mathbf{A}^2 \mathbf{BAB}^2, \mathbf{B}^2 \mathbf{ABA}^2, \mathbf{AB}^2 \mathbf{A}^2 \mathbf{B}, \\ & \mathbf{BA}^2 \mathbf{B}^2 \mathbf{A}, \mathbf{A}^2 \mathbf{B}^2 \mathbf{AB} \quad \text{or} \quad \mathbf{B}^2 \mathbf{A}^2 \mathbf{BA}. \end{aligned} \quad (10.4)$$

Thus, if we consider any polynomial $\mathbf{P}$ in the two 3×3 matrices $\mathbf{A}$ and $\mathbf{B}$, it can be expressed in the form

$$\begin{aligned} \mathbf{P} = & \mathbf{P}' + \varphi'_1 \mathbf{ABA}^2 \mathbf{B}^2 + \varphi'_2 \mathbf{BAB}^2 \mathbf{A}^2 + \varphi'_3 \mathbf{A}^2 \mathbf{BAB}^2 + \varphi'_4 \mathbf{B}^2 \mathbf{ABA}^2 \\ & + \varphi'_5 \mathbf{AB}^2 \mathbf{A}^2 \mathbf{B} + \varphi'_6 \mathbf{BA}^2 \mathbf{B}^2 \mathbf{A} + \varphi'_7 \mathbf{A}^2 \mathbf{B}^2 \mathbf{AB} + \varphi'_8 \mathbf{B}^2 \mathbf{A}^2 \mathbf{BA}, \end{aligned} \quad (10.5)$$

where $\mathbf{P}'$ is a polynomial of extension 3 or less in the matrices $\mathbf{A}$ and $\mathbf{B}$ and $\varphi'_1, \varphi'_2, \dots, \varphi'_8$ are scalar invariants of $\mathbf{A}$ and $\mathbf{B}$ expressible as polynomials in the elements of $\mathbf{A}$ and $\mathbf{B}$. Equations (4.23) and (4.25) can be used to express $\mathbf{A}^2 \mathbf{BA}$ as a linear combination of $\mathbf{ABA}^2$ and a polynomial of extension 2 in $\mathbf{A}$ and $\mathbf{B}$. Analogous formulae can be used to express $\mathbf{B}^2 \mathbf{AB}$, $\mathbf{A}^2 \mathbf{B}^2 \mathbf{A}$, $\mathbf{B}^2 \mathbf{A}^2 \mathbf{B}$ as linear combinations of $\mathbf{BAB}^2$, $\mathbf{AB}^2 \mathbf{A}^2$, $\mathbf{BA}^2 \mathbf{B}^2$ respectively and polynomials of extension 2 in $\mathbf{A}$ and $\mathbf{B}$. We thus obtain, from (10.5),

$$\mathbf{P} = \mathbf{P}'' + \varphi''_1 \mathbf{ABA}^2 \mathbf{B}^2 + \varphi''_2 \mathbf{BAB}^2 \mathbf{A}^2, \quad (10.6)$$

where $\mathbf{P}''$ is a polynomial of extension 3 or less in the matrices $\mathbf{A}$ and $\mathbf{B}$ and

φ_1'' and φ_2'' are scalar invariants of $\mathbf{A}$ and $\mathbf{B}$ expressible as polynomials in the elements of $\mathbf{A}$ and $\mathbf{B}$.

Again using formulae analogous to (4.23) and (4.25), we can replace $\mathbf{BAB}^2$ and $\mathbf{ABA}^2$ by linear combinations of $\mathbf{B}^2\mathbf{AB}$ and $\mathbf{A}^2\mathbf{BA}$ respectively with polynomials of extension 2 in $\mathbf{A}$ and $\mathbf{B}$. Employing such relations in the term $\mathbf{BAB}^2\mathbf{A}^2$ of (10.6), we obtain

$$\mathbf{P} = \mathbf{Q} + \varphi_1 \mathbf{ABA}^2\mathbf{B}^2 + \varphi_2 \mathbf{B}^2\mathbf{A}^2\mathbf{BA}, \quad (10.7)$$

where $\mathbf{Q}$ is a polynomial of extension 3 or less in $\mathbf{A}$ and $\mathbf{B}$ and φ_1 and φ_2 are invariants of $\mathbf{A}$ and $\mathbf{B}$ expressible as polynomials in the elements of $\mathbf{A}$ and $\mathbf{B}$.

If $\mathbf{P}$, $\mathbf{A}$ and $\mathbf{B}$ are symmetric matrices, then $\mathbf{P} = \mathbf{P}^*$, where $\mathbf{P}^*$ denotes the transpose of $\mathbf{P}$. Now,

$$\mathbf{P}^* = \mathbf{Q}^* + \varphi_1 \mathbf{B}^2\mathbf{A}^2\mathbf{BA} + \varphi_2 \mathbf{ABA}^2\mathbf{B}^2, \quad (10.8)$$

since

$$\mathbf{B}^2\mathbf{A}^2\mathbf{BA} = (\mathbf{ABA}^2\mathbf{B}^2)^*.$$

Thus, from (10.7), (10.8) and $\mathbf{P} = \mathbf{P}^*$,

$$\mathbf{P} = \frac{1}{2}(\mathbf{P} + \mathbf{P}^*) = \frac{1}{2}(\mathbf{Q} + \mathbf{Q}^*) + \frac{1}{2}(\varphi_1 + \varphi_2)(\mathbf{ABA}^2\mathbf{B}^2 + \mathbf{B}^2\mathbf{A}^2\mathbf{BA}). \quad (10.9)$$

$\frac{1}{2}(\mathbf{Q} + \mathbf{Q}^*)$ is a polynomial of extension 3 or less in $\mathbf{A}$ and $\mathbf{B}$. It is, of course, a symmetric matrix. It can readily be shown that if $\mathbf{R}$ is a symmetric matrix polynomial of extension 3 in the symmetric 3×3 matrices $\mathbf{A}$ and $\mathbf{B}$, it can be expressed as a symmetric matrix polynomial of extension 2 in $\mathbf{A}$ and $\mathbf{B}$.

Any matrix polynomial $\mathbf{R}$ of extension 3 in the matrices $\mathbf{A}$ and $\mathbf{B}$ can be expressed as

$$\begin{aligned} \mathbf{R} = & \mathbf{T} + \theta_1 \mathbf{ABA} + \theta_2 \mathbf{BAB} + \theta_3 \mathbf{AB}^2\mathbf{A} + \theta_4 \mathbf{BA}^2\mathbf{B} + \theta_5 \mathbf{A}^2\mathbf{BA}^2 \\ & + \theta_6 \mathbf{B}^2\mathbf{AB}^2 + \theta_7 \mathbf{A}^2\mathbf{B}^2\mathbf{A}^2 + \theta_8 \mathbf{B}^2\mathbf{A}^2\mathbf{B}^2 + \theta_9 \mathbf{ABA}^2 + \theta_{10} \mathbf{A}^2\mathbf{BA} \\ & + \theta_{11} \mathbf{BAB}^2 + \theta_{12} \mathbf{B}^2\mathbf{AB} + \theta_{13} \mathbf{AB}^2\mathbf{A}^2 + \theta_{14} \mathbf{A}^2\mathbf{B}^2\mathbf{A} \\ & + \theta_{15} \mathbf{BA}^2\mathbf{B}^2 + \theta_{16} \mathbf{B}^2\mathbf{A}^2\mathbf{B}, \end{aligned} \quad (10.10)$$

where the θ 's are invariants of the matrices $\mathbf{A}$ and $\mathbf{B}$ which can be expressed as polynomials in their elements and $\mathbf{T}$ is a polynomial in $\mathbf{A}$ and $\mathbf{B}$ of extension 2. Employing the relation (4.22), the relation derived from it by interchanging $\mathbf{A}$ and $\mathbf{B}$ and the relations derived from these two relations by replacing $\mathbf{A}$ by $\mathbf{A}^2$ or $\mathbf{B}$ by $\mathbf{B}^2$ or $\mathbf{A}$ and $\mathbf{B}$ by $\mathbf{A}^2$ and $\mathbf{B}^2$ respectively, we obtain from (10.10),

$$\begin{aligned} \mathbf{R} = & \mathbf{U} + \theta_9 \mathbf{ABA}^2 + \theta_{10} \mathbf{A}^2\mathbf{BA} + \theta_{11} \mathbf{BAB}^2 + \theta_{12} \mathbf{B}^2\mathbf{AB} \\ & + \theta_{13} \mathbf{AB}^2\mathbf{A}^2 + \theta_{14} \mathbf{A}^2\mathbf{B}^2\mathbf{A} + \theta_{15} \mathbf{BA}^2\mathbf{B}^2 + \theta_{16} \mathbf{B}^2\mathbf{A}^2\mathbf{B}, \end{aligned} \quad (10.11)$$

where $\mathbf{U}$ is a polynomial in $\mathbf{A}$ and $\mathbf{B}$ of extension 2.

If $\mathbf{R}$, $\mathbf{A}$ and $\mathbf{B}$ are symmetric, then

$$\begin{aligned}\mathbf{R} = \frac{1}{2}(\mathbf{R} + \mathbf{R}^*) &= \frac{1}{2}(\mathbf{U} + \mathbf{U}^*) + \frac{1}{2}[(\theta_9 + \theta_{10})(\mathbf{A}\mathbf{B}\mathbf{A}^2 + \mathbf{A}^2\mathbf{B}\mathbf{A}) \\ &+ (\theta_{11} + \theta_{12})(\mathbf{B}\mathbf{A}\mathbf{B}^2 + \mathbf{B}^2\mathbf{A}\mathbf{B}) + (\theta_{13} + \theta_{14})(\mathbf{A}\mathbf{B}^2\mathbf{A}^2 \\ &+ \mathbf{A}^2\mathbf{B}^2\mathbf{A}) + (\theta_{15} + \theta_{16})(\mathbf{B}\mathbf{A}^2\mathbf{B}^2 + \mathbf{B}^2\mathbf{A}^2\mathbf{B})].\end{aligned}\quad (10.12)$$

Now relations of the type (4.22) and (4.25), can be used to substitute polynomials of extension 2 in $\mathbf{A}$ and $\mathbf{B}$ for the terms of extension 3 in (10.12). We thus obtain

Lemma 5. *Any symmetric matrix polynomial of extension 3 in two, symmetric, 3×3 matrices can be expressed as a matrix polynomial of extension 2.*

It follows from (10.9) that the symmetric matrix polynomial $\mathbf{P}$ of arbitrary extension in the symmetric matrices $\mathbf{A}$ and $\mathbf{B}$ may be expressed as

$$\mathbf{P} = \mathbf{Q} + \varphi(\mathbf{A}\mathbf{B}\mathbf{A}^2\mathbf{B}^2 + \mathbf{B}^2\mathbf{A}^2\mathbf{B}\mathbf{A}), \quad (10.13)$$

where $\mathbf{Q}$ is a polynomial of extension 2 and φ is a scalar invariant of $\mathbf{A}$ and $\mathbf{B}$ expressible as a polynomial in their elements. $\mathbf{Q}$ must, of course, be expressible in the form

$$\begin{aligned}\mathbf{Q} = \psi_0 \mathbf{I} + \psi_1 \mathbf{A} + \psi_2 \mathbf{B} + \psi_3 \mathbf{A}^2 + \psi_4 \mathbf{B}^2 + \psi_5(\mathbf{A}\mathbf{B} + \mathbf{B}\mathbf{A}) \\ + \psi_6(\mathbf{A}^2\mathbf{B} + \mathbf{B}\mathbf{A}^2) + \psi_7(\mathbf{B}^2\mathbf{A} + \mathbf{A}\mathbf{B}^2) + \psi_8(\mathbf{A}^2\mathbf{B}^2 + \mathbf{B}^2\mathbf{A}^2),\end{aligned}\quad (10.14)$$

where the ψ 's are scalar invariants of $\mathbf{A}$ and $\mathbf{B}$ expressible as polynomials in their elements.

Now, replacing $\mathbf{B}$ by $\mathbf{B}^2$ in (4.22), we obtain

$$\begin{aligned}\mathbf{A}\mathbf{B}^2\mathbf{A} + \mathbf{A}^2\mathbf{B}^2 + \mathbf{B}^2\mathbf{A}^2 - \mathbf{A}[\text{tr } \mathbf{A}\mathbf{B}^2 - \text{tr } \mathbf{A} \text{ tr } \mathbf{B}^2] \\ - \frac{1}{2}\mathbf{B}^2[\text{tr } \mathbf{A}^2 - (\text{tr } \mathbf{A})^2] - (\mathbf{A}\mathbf{B}^2 + \mathbf{B}^2\mathbf{A})\text{tr } \mathbf{A} - \mathbf{A}^2 \text{tr } \mathbf{B}^2 \\ - \mathbf{I}[\text{tr } \mathbf{A}^2\mathbf{B}^2 - \text{tr } \mathbf{A} \text{ tr } \mathbf{A}\mathbf{B}^2 + \frac{1}{2} \text{tr } \mathbf{B}^2 \{(\text{tr } \mathbf{A})^2 - \text{tr } \mathbf{A}^2\}] = 0.\end{aligned}\quad (10.15)$$

Multiplying equation (10.15) on the left by $\mathbf{A}\mathbf{B}$ and on the right by $\mathbf{B}\mathbf{A}$ and adding the equations so obtained, we obtain

$$\begin{aligned}\mathbf{A}\mathbf{B}\mathbf{A}^2\mathbf{B}^2 + \mathbf{B}^2\mathbf{A}^2\mathbf{B}\mathbf{A} + \mathbf{A}\mathbf{B}^2\mathbf{A}\mathbf{B}\mathbf{A} + \mathbf{A}\mathbf{B}\mathbf{A}\mathbf{B}^2\mathbf{A} + \mathbf{A}\mathbf{B}^3\mathbf{A}^2 + \mathbf{A}^2\mathbf{B}^3\mathbf{A} \\ - 2\mathbf{A}\mathbf{B}\mathbf{A}[\text{tr } \mathbf{A}\mathbf{B}^2 - \text{tr } \mathbf{A} \text{ tr } \mathbf{B}^2] - \frac{1}{2}(\mathbf{A}\mathbf{B}^3 + \mathbf{B}^3\mathbf{A})[\text{tr } \mathbf{A}^2 \\ - (\text{tr } \mathbf{A})^2] - (\mathbf{A}\mathbf{B}\mathbf{A}\mathbf{B}^2 + \mathbf{B}^2\mathbf{A}\mathbf{B}\mathbf{A})\text{tr } \mathbf{A} - 2\mathbf{A}\mathbf{B}^3\mathbf{A} \text{tr } \mathbf{A} \\ - (\mathbf{A}\mathbf{B}\mathbf{A}^2 + \mathbf{A}^2\mathbf{B}\mathbf{A})\text{tr } \mathbf{B}^2 - (\mathbf{A}\mathbf{B} + \mathbf{B}\mathbf{A})[\text{tr } \mathbf{A}^2\mathbf{B}^2 - \text{tr } \mathbf{A} \text{ tr } \mathbf{A}\mathbf{B}^2] \\ + \frac{1}{2} \text{tr } \mathbf{B}^2 \{(\text{tr } \mathbf{A})^2 - \text{tr } \mathbf{A}^2\} = 0.\end{aligned}\quad (10.16)$$

It is seen from (7.7) and (4.22) that $\mathbf{ABAB}^2\mathbf{A} + \mathbf{AB}^2\mathbf{ABA}$ can be expressed as a polynomial of extension 2. Also, since $\mathbf{ABA}$ can be expressed as a polynomial of extension 2, $\mathbf{ABAB}^2 + \mathbf{B}^2\mathbf{ABA}$ can be expressed as a symmetric matrix polynomial of extension 3. From equation (10.16) and lemma 5, we thus find that $\mathbf{ABA}^2\mathbf{B}^2 + \mathbf{B}^2\mathbf{A}^2\mathbf{BA}$ can be expressed as a symmetric polynomial of extension 2 in $\mathbf{A}$ and $\mathbf{B}$. Thus, from (10.13) and (10.14), we obtain

$$\begin{aligned} \mathbf{P} = & \chi_0 \mathbf{I} + \chi_1 \mathbf{A} + \chi_2 \mathbf{B} + \chi_3 \mathbf{A}^2 + \chi_4 \mathbf{B}^2 + \chi_5(\mathbf{AB} + \mathbf{BA}) \\ & + \chi_6(\mathbf{A}^2\mathbf{B} + \mathbf{BA}^2) + \chi_7(\mathbf{B}^2\mathbf{A} + \mathbf{AB}^2) + \chi_8(\mathbf{A}^2\mathbf{B}^2 + \mathbf{B}^2\mathbf{A}^2), \end{aligned} \quad (10.17)$$

where the χ 's are scalar invariants of $\mathbf{A}$ and $\mathbf{B}$ expressible as polynomials in their elements. This result may be expressed as

Lemma 6. *Any symmetric matrix polynomial in two, symmetric, 3×3 matrices can be expressed as a symmetric matrix polynomial of extension 2.*

11. Scalar invariants of R , 3×3 matrices. When we use the term scalar invariant in this paper, we imply invariance under the orthogonal group and we consider that the invariant is expressible as a polynomial in the elements of the matrices. Let J be a scalar invariant formed from the elements of the R , 3×3 matrices $\mathbf{A}, \mathbf{B}, \dots, \mathbf{R}$. We may write J in the form

$$\begin{aligned} J = \sum_{r,s,\dots,t} & \alpha_{k_1 l_1 k_2 l_2 \dots k_r l_r m_1 n_1 m_2 n_2 \dots m_s n_s \dots p_1 q_1 p_2 q_2 \dots p_t q_t} \\ & \cdot a_{k_1 l_1} a_{k_2 l_2} \dots a_{k_r l_r} b_{m_1 n_1} b_{m_2 n_2} \dots b_{m_s n_s} \dots r_{p_1 q_1} r_{p_2 q_2} \dots r_{p_t q_t}, \end{aligned} \quad (11.1)$$

where the α 's are constant coefficients.

Since J is invariant under an arbitrary orthogonal transformation $\mathbf{S}$ ($= \|s_{ij}\|$) say, the α 's in (11.1) must be such that

$$\begin{aligned} & \alpha_{k_1 l_1 k_2 l_2 \dots k_r l_r m_1 n_1 m_2 n_2 \dots m_s n_s \dots p_1 q_1 p_2 q_2 \dots p_t q_t} a_{k_1 l_1} a_{k_2 l_2} \dots a_{k_r l_r} b_{m_1 n_1} b_{m_2 n_2} \dots \\ & b_{m_s n_s} \dots r_{p_1 q_1} r_{p_2 q_2} \dots r_{p_t q_t} = \alpha_{e_1 f_1 e_2 f_2 \dots e_r f_r g_1 h_1 g_2 h_2 \dots g_s h_s \dots i_1 j_1 i_2 j_2 \dots i_t j_t} \\ & \cdot \bar{a}_{e_1 f_1} \bar{a}_{e_2 f_2} \dots \bar{a}_{e_r f_r} \bar{b}_{g_1 h_1} \bar{b}_{g_2 h_2} \dots \bar{b}_{g_s h_s} \dots \bar{r}_{i_1 j_1} \bar{r}_{i_2 j_2} \dots \bar{r}_{i_t j_t}, \end{aligned} \quad (11.2)$$

where

$$\begin{aligned} \bar{a}_{ef} &= s_{ek} s_{fl} a_{kl}, \\ \bar{b}_{gh} &= s_{gm} s_{hn} b_{mn}, \\ & \dots \dots \dots \\ \bar{r}_{ij} &= s_{ip} s_{jq} r_{pq}. \end{aligned} \quad (11.3)$$

Substituting from (11.3) in (11.2), we obtain

$$\begin{aligned}
 & \alpha_{k_1 l_1 k_2 l_2 \dots k_r l_r m_1 n_1 m_2 n_2 \dots m_s n_s \dots p_1 q_1 p_2 q_2 \dots p_t q_t} \\
 & \quad \cdot a_{k_1 l_1} a_{k_2 l_2} \dots a_{k_r l_r} b_{m_1 n_1} b_{m_2 n_2} \dots b_{m_s n_s} \dots r_{p_1 q_1} r_{p_2 q_2} \dots r_{p_t q_t} \\
 & \quad = \alpha_{e_1 f_1 e_2 f_2 \dots e_r f_r g_1 h_1 g_2 h_2 \dots g_s h_s \dots i_1 j_1 i_2 j_2 \dots i_t j_t} \\
 & \quad \quad \cdot s_{e_1 k_1} s_{f_1 l_1} s_{e_2 k_2} s_{f_2 l_2} \dots s_{e_r k_r} s_{f_r l_r} \\
 & \quad \quad \cdot s_{g_1 m_1} s_{h_1 n_1} s_{g_2 m_2} s_{h_2 n_2} \dots s_{g_s m_s} s_{h_s n_s} \\
 & \quad \quad \cdot s_{i_1 p_1} s_{j_1 q_1} s_{i_2 p_2} s_{j_2 q_2} \dots s_{i_t p_t} s_{j_t q_t} \\
 & \quad \cdot a_{k_1 l_1} a_{k_2 l_2} \dots a_{k_r l_r} b_{m_1 n_1} b_{m_2 n_2} \dots b_{m_s n_s} \dots r_{p_1 q_1} r_{p_2 q_2} \dots r_{p_t q_t}.
 \end{aligned} \tag{11.4}$$

From (11.4), we have

$$\begin{aligned}
 & \alpha_{k_1 l_1 k_2 l_2 \dots k_r l_r m_1 n_1 m_2 n_2 \dots m_s n_s \dots p_1 q_1 p_2 q_2 \dots p_t q_t} \\
 & \quad = \alpha_{e_1 f_1 e_2 f_2 \dots e_r f_r g_1 h_1 g_2 h_2 \dots g_s h_s \dots i_1 j_1 i_2 j_2 \dots i_t j_t} \\
 & \quad \quad \cdot s_{e_1 k_1} s_{f_1 l_1} s_{e_2 k_2} s_{f_2 l_2} \dots s_{e_r k_r} s_{f_r l_r} \\
 & \quad \quad \cdot s_{g_1 m_1} s_{h_1 n_1} s_{g_2 m_2} s_{h_2 n_2} \dots s_{g_s m_s} s_{h_s n_s} \\
 & \quad \quad \cdot s_{i_1 p_1} s_{j_1 q_1} s_{i_2 p_2} s_{j_2 q_2} \dots s_{i_t p_t} s_{j_t q_t}.
 \end{aligned} \tag{11.5}$$

Thus, $\alpha_{k_1 l_1 k_2 l_2 \dots k_r l_r m_1 n_1 m_2 n_2 \dots m_s n_s \dots p_1 q_1 p_2 q_2 \dots p_t q_t}$ are the components of an isotropic tensor of rank $2(r + s + \dots + t)$. Since they are not functions of the elements of the matrices $\mathbf{A}, \mathbf{B}, \dots, \mathbf{R}$, they must be expressible as a linear combination, with constant coefficients, of the outer products of $r + s + \dots + t$ Kronecker δ 's. These outer products are

$$\delta_{k_1 l_1} \delta_{k_2 l_2} \dots \delta_{k_r l_r} \delta_{m_1 n_1} \delta_{m_2 n_2} \dots \delta_{m_s n_s} \dots \delta_{p_1 q_1} \delta_{p_2 q_2} \dots \delta_{p_t q_t} \tag{11.6}$$

and the tensors obtained by all possible permutations of the indices in (11.6). Employing such an expression for $\alpha_{k_1 l_1 \dots p_t q_t}$ in (11.1), we see that J may be expressed as a polynomial in traces of products formed from the matrices $\mathbf{A}, \mathbf{B}, \dots, \mathbf{R}$.

It is evident that the argument employed above remains valid if $\mathbf{A}, \mathbf{B}, \dots, \mathbf{R}$ are $R, n \times n$ matrices.

12. Scalar invariants of two symmetric 3×3 matrices. Let us suppose that $\mathbf{\Pi}$ is any product, formed from two symmetric 3×3 matrices $\mathbf{A}$ and $\mathbf{B}$, of degree m in $\mathbf{A}$ and n in $\mathbf{B}$. Then, $\mathbf{\Pi}$ may be expressed either as $\mathbf{A}\mathbf{\Pi}'$ or $\mathbf{B}\mathbf{\Pi}'$, where $\mathbf{\Pi}'$ is also a product formed from $\mathbf{A}$ and $\mathbf{B}$. Let us suppose that

$$\mathbf{\Pi} = \mathbf{A}\mathbf{\Pi}'. \tag{12.1}$$

Then, Π' is of degree $m - 1$ in $\mathbf{A}$ and of degree n in $\mathbf{B}$. $\Pi' + \Pi'^*$ is a symmetric matrix polynomial in $\mathbf{A}$ and $\mathbf{B}$. From equation (10.17), and the result of §11, $\Pi' + \Pi'^*$ may be expressed as

$$\begin{aligned} \Pi' + \Pi'^* = & \chi_0 \mathbf{I} + \chi_1 \mathbf{A} + \chi_2 \mathbf{B} + \chi_3 \mathbf{A}^2 + \chi_4 \mathbf{B}^2 + \chi_5(\mathbf{AB} + \mathbf{BA}) \\ & + \chi_6(\mathbf{A}^2 \mathbf{B} + \mathbf{BA}^2) + \chi_7(\mathbf{B}^2 \mathbf{A} + \mathbf{AB}^2) + \chi_8(\mathbf{A}^2 \mathbf{B}^2 + \mathbf{B}^2 \mathbf{A}^2), \end{aligned} \quad (12.2)$$

where $\chi_0, \chi_1, \dots, \chi_8$ are expressible as polynomials in traces of products formed from the matrices $\mathbf{A}$ and $\mathbf{B}$. We note that χ_0 must be of degree $m - 1$ in the elements of $\mathbf{A}$ and of degree n in the elements of $\mathbf{B}$. $\chi_1, \chi_2, \dots, \chi_8$ are of lower or equal degrees in the elements of $\mathbf{A}$ and in the elements of $\mathbf{B}$ than is χ_0 .

Now, from (12.2),

$$\begin{aligned} \mathbf{A}(\Pi' + \Pi'^*) = & \chi_0 \mathbf{A} + \chi_1 \mathbf{A}^2 + \chi_2 \mathbf{AB} + \chi_3 \mathbf{A}^3 + \chi_4 \mathbf{AB}^2 \\ & + \chi_5(\mathbf{A}^2 \mathbf{B} + \mathbf{ABA}) + \chi_6(\mathbf{A}^3 \mathbf{B} + \mathbf{ABA}^2) + \chi_7(\mathbf{AB}^2 \mathbf{A} + \mathbf{A}^2 \mathbf{B}^2) \\ & + \chi_8(\mathbf{A}^3 \mathbf{B}^2 + \mathbf{AB}^2 \mathbf{A}^2). \end{aligned} \quad (12.3)$$

Since,

$$\text{tr } \mathbf{A}(\Pi' + \Pi'^*) = \text{tr } \mathbf{A}\Pi' + \text{tr } \mathbf{A}\Pi'^* = \text{tr } \mathbf{A}\Pi' + \text{tr } \Pi'^* \mathbf{A} = 2 \text{tr } \Pi, \quad (12.4)$$

we have

$$\begin{aligned} 2 \text{tr } \Pi = & \chi_0 \text{tr } \mathbf{A} + \chi_1 \text{tr } \mathbf{A}^2 + \chi_2 \text{tr } \mathbf{AB} + \chi_3 \text{tr } \mathbf{A}^3 + \chi_4 \text{tr } \mathbf{AB}^2 \\ & + 2\chi_5 \text{tr } \mathbf{A}^2 \mathbf{B} + 2\chi_6 \text{tr } \mathbf{A}^3 \mathbf{B} + 2\chi_7 \text{tr } \mathbf{A}^2 \mathbf{B}^2 + 2\chi_8 \text{tr } \mathbf{A}^3 \mathbf{B}^2. \end{aligned} \quad (12.5)$$

Now, employing the Hamilton-Cayley theorem to substitute for $\mathbf{A}^3$ in $\mathbf{A}^3 \mathbf{B}$ and $\mathbf{A}^3 \mathbf{B}^2$, we obtain

$$\begin{aligned} \text{tr } \mathbf{A}^3 \mathbf{B} = & \text{tr } \mathbf{A} \text{tr } \mathbf{A}^2 \mathbf{B} - \frac{1}{2}[(\text{tr } \mathbf{A})^2 - \text{tr } \mathbf{A}^2] \text{tr } \mathbf{AB} \\ & + [\frac{1}{3} \text{tr } \mathbf{A}^3 - \frac{1}{2} \text{tr } \mathbf{A} \text{tr } \mathbf{A}^2 + \frac{1}{6}(\text{tr } \mathbf{A})^3] \text{tr } \mathbf{B} \end{aligned} \quad (12.6)$$

and

$$\begin{aligned} \text{tr } \mathbf{A}^3 \mathbf{B}^2 = & \text{tr } \mathbf{A} \text{tr } \mathbf{A}^2 \mathbf{B}^2 - \frac{1}{2}[(\text{tr } \mathbf{A})^2 - \text{tr } \mathbf{A}^2] \text{tr } \mathbf{AB}^2 \\ & + [\frac{1}{3} \text{tr } \mathbf{A}^3 - \frac{1}{2} \text{tr } \mathbf{A} \text{tr } \mathbf{A}^2 + \frac{1}{6}(\text{tr } \mathbf{A})^3] \text{tr } \mathbf{B}^2. \end{aligned} \quad (12.7)$$

Substituting from (12.6) and (12.7) in (12.5), we see that $\text{tr } \Pi$ can be expressed as a polynomial in $\text{tr } \mathbf{A}$, $\text{tr } \mathbf{B}$, $\text{tr } \mathbf{A}^2$, $\text{tr } \mathbf{B}^2$, $\text{tr } \mathbf{A}^3$, $\text{tr } \mathbf{AB}$, $\text{tr } \mathbf{A}^2 \mathbf{B}$, $\text{tr } \mathbf{AB}^2$, $\text{tr } \mathbf{A}^2 \mathbf{B}^2$ and traces of products of $\mathbf{A}$ and $\mathbf{B}$ of degree $m - 1$ or less in $\mathbf{A}$ and n or less in $\mathbf{B}$. Similarly, if Π can be expressed as $\mathbf{B}\Pi'$, then $\text{tr } \Pi$ can be expressed as a polynomial in $\text{tr } \mathbf{A}$, $\text{tr } \mathbf{B}$, $\text{tr } \mathbf{A}^2$, $\text{tr } \mathbf{B}^2$, $\text{tr } \mathbf{B}^3$, $\text{tr } \mathbf{AB}$, $\text{tr } \mathbf{A}^2 \mathbf{B}$, $\text{tr } \mathbf{AB}^2$, $\text{tr } \mathbf{A}^2 \mathbf{B}^2$ and traces of products of $\mathbf{A}$ and $\mathbf{B}$ of degree m or less in $\mathbf{A}$ and $n - 1$ or less in $\mathbf{B}$. It follows,

by repeated application of these results, that $\text{tr } \Pi$ and hence any scalar invariant can be expressed as a polynomial in the scalar invariants

$$\text{tr } \mathbf{A}, \text{tr } \mathbf{B}, \text{tr } \mathbf{A}^2, \text{tr } \mathbf{B}^2, \text{tr } \mathbf{A}^3, \text{tr } \mathbf{B}^3, \text{tr } \mathbf{AB}, \text{tr } \mathbf{A}^2 \mathbf{B}, \text{tr } \mathbf{AB}^2 \quad \text{and} \quad \text{tr } \mathbf{A}^2 \mathbf{B}^2. \quad (12.8)$$

13. Scalar invariants of R , symmetric 2×2 matrices. From §11, it is apparent that any scalar invariant of R , symmetric 2×2 matrices is expressible as a polynomial in traces of products of these matrices. Now, let us suppose that Π is a product formed from the R , symmetric 2×2 matrices $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_R$ of degree $n_1, n_2, \dots, n_R$ in $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_R$ respectively. Let us suppose that Π may be expressed in the form $\mathbf{A}_p \Pi'$, where Π' is also a product formed from the matrices $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_R$. Π' is, of course, of degree $n_1, n_2, \dots, n_p - 1, \dots, n_R$ in $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_p, \dots, \mathbf{A}_R$ respectively. Then, from lemma 2, $\Pi' + \Pi'^*$ is expressible in the form

$$\Pi' + \Pi'^* = \alpha_0 \mathbf{I} + \sum_{r=1}^R \alpha_r \mathbf{A}_r, \quad (13.1)$$

where the α 's are scalar invariants of $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_R$.

From (13.1), we have

$$\mathbf{A}_p(\Pi' + \Pi'^*) = \alpha_0 \mathbf{A}_p + \sum_{r=1}^R \alpha_r \mathbf{A}_p \mathbf{A}_r. \quad (13.2)$$

Thus,

$$\text{tr } \mathbf{A}_p \Pi' + \text{tr } \mathbf{A}_p \Pi'^* = 2 \text{tr } \Pi = \alpha_0 \text{tr } \mathbf{A}_p + \sum_{r=1}^R \alpha_r \text{tr } \mathbf{A}_p \mathbf{A}_r. \quad (13.3)$$

Unless α_0 is zero, it is of degree $n_1, n_2, \dots, n_p - 1, \dots, n_R$ in the elements of $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_p, \dots, \mathbf{A}_R$ respectively. Also, unless α_r ($r \neq p$) is zero it is of degree $n_1, n_2, \dots, n_p - 1, \dots, n_r - 1, \dots, n_R$ in the elements of $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_p, \dots, \mathbf{A}_r, \dots, \mathbf{A}_R$ respectively. Unless α_p is zero, it is, of course, of degree $n_1, n_2, \dots, n_p - 2, \dots, n_R$ in the elements of $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_p, \dots, \mathbf{A}_R$ respectively. Thus, equation (13.3) expresses $\text{tr } \Pi$ as a polynomial in $\text{tr } \mathbf{A}_p$, $\text{tr } \mathbf{A}_p \mathbf{A}_r$ ($r = 1, 2, \dots, R$) and traces of products of $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_R$ of total degree less than $n_1 + n_2 + \dots + n_R - 1$ and of degree less than or equal to $n_1, n_2, \dots, n_p - 1, \dots, n_R$ in the elements of the matrices $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_p, \dots, \mathbf{A}_R$ respectively. By successive application of this procedure, where p may take any appropriate value from 1 to R , we see that $\text{tr } \Pi$ and hence any scalar invariant of $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_R$ may be expressed as a polynomial in $\text{tr } \mathbf{A}_r$ ($r = 1, 2, \dots, R$) and $\text{tr } \mathbf{A}_r \mathbf{A}_s$ ($r, s = 1, 2, \dots, R$), where it may be borne in mind that $\text{tr } \mathbf{A}_r \mathbf{A}_s = \text{tr } \mathbf{A}_s \mathbf{A}_r$.

REFERENCES

- W. PRAGER, 1945, Jour. of Appl. Physics, **16**, 837.
M. REINER, 1945, Amer. J. Math. **67**, 350.
R. S. RIVLIN & J. L. ERICKSEN, 1955, J. Rat'l. Mech. & Anal. **4**, 323.
J. A. TODD, 1948, Phil. Trans. A **241**, 399.

Acknowledgment. Much of the approach developed in this paper arose as a result of discussion with Dr. J. L. ERICKSEN and Dr. R. TOUPIN, while the previous paper by RIVLIN & ERICKSEN (1955) was being written. The author is therefore indebted to Drs. ERICKSEN and TOUPIN for significant contributions to the ideas expressed in the present paper.