

On some cases of Matrices with Linear Invariant Factors. By
H. F. BAKER. Received and Read December 11th, 1902.

When an important theorem is capable of very brief proof it is desirable that the fact should be widely known. This more than their novelty may perhaps justify the following lines. We denote a matrix of n rows and columns by a single capital letter, say M ; that obtained from it by changing rows into columns by \bar{M} ; that obtained from it by changing every element into its conjugate complex by M_0 . And a row of n single quantities is denoted by a small letter, such as x or y ; their conjugate complexes by x_0 or y_0 . The symbol Mx is the row of n quantities obtained by combining the element of each row of M , in turn, with the elements of x ; while Mxy is the single quantity obtained by combining the elements of Mx with those of y . We utilize the obvious identity $Mxy = \bar{M}yx$, and assume the fundamental theorem as to invariant factors explained in detail in § 13 of the preceding paper "On the Integration of Linear Differential Equations."

1. For a matrix M such that $M^p = 1$, p being a positive integer, if θ be a root of the determinantal equation $|M - \rho| = 0$, we can find x so that $Mx = \theta x$, and hence $M^p x = \theta^p x$; so that $\theta^p = 1$, or θ is a p -th root of unity. We cannot, however, find y , other than zero and linearly independent of x , such that $My = \theta y + x$, or else we should have in succession

$$M^2 y = \theta(\theta y + x) + \theta x = \theta^2 y + 2\theta x, \quad M^3 y = \theta^3(\theta y + x) + 2\theta^2 x = \theta^3 y + 3\theta^2 x,$$

$$\text{and finally} \quad M^p y = \theta^p y + p\theta^{p-1}x,$$

$$\text{leading to} \quad p\theta^{p-1}x = 0,$$

which is absurd, provided $p \neq 0$ and $\theta \neq 0$, of which the latter follows from $M^p = 1$, giving $|M| \neq 0$.

Thus the invariant factors of the matrix are linear, and a matrix μ can be found such that $\mu^{-1}M\mu$ consists only of diagonal elements, the roots of $|M - \rho| = 0$. It is known (E. H. Moore, *Math. Ann.*, Vol. L., 1898, p. 214) that, if M_1, M_2, \dots be the generating matrices of a finite group, a matrix m can be found such that the matrices $N_1 = m^{-1}M_1m$, $N_2 = m^{-1}M_2m$, ... satisfy the relation $\bar{N}_0 N = 1$. Thus the result is a particular case of that proved under (3) below.

2. If a matrix M satisfy the relation $\bar{M} = M_0$ so that, if it consist of real elements, it is a symmetrical matrix, and if θ be a root of $|M - \rho| = 0$, we have for a suitable row x the sequences

$$Mx = \theta x, \quad Mxx_0 = \theta xx_0;$$

$$M_0x_0 = \theta_0x_0, \quad M_0x_0x = \theta_0x_0x, \quad \bar{M}_0xx_0 = \theta_0xx_0, \quad Mxx_0 = \theta_0xx_0,$$

and can infer $(\theta - \theta_0)xx_0 = 0, \quad \theta = \theta_0,$

so that the roots θ are necessarily real; but we cannot have an equation $(M - \theta)y = x$, or else the real quantity

$$xx_0 = (M - \theta)yx_0 = (M_0 - \theta)y_0x = (\bar{M}_0 - \theta)xy_0 = (M - \theta)xy_0 = 0$$

would be zero.

Thus the invariant factors of the matrix M are necessarily linear, and the roots are real.

If A, B be two matrices of the n -th order such that $\bar{A} = A_0, \bar{B} = B_0$, and Bxx_0 only vanishes if each element of x is zero, so that $|B| \neq 0$, and θ be a root of $|A - \rho B| = 0$, we can, from $(B^{-1}A - \theta)x = 0$, infer

$$Axx_0 = \theta Bxx_0, \quad \bar{A}x_0x = \theta \bar{B}x_0x, \quad A_0x_0x = \theta B_0x_0x, \quad Axx_0 = \theta_0 Bxx_0,$$

and hence $(\theta - \theta_0)Bxx_0 = 0$, showing that θ is real; but we cannot have $(B^{-1}A - \theta)y = x$, or else the real quantity

$$\begin{aligned} \bar{B}_0xx_0 = B_0x_0x = \bar{B}x_0x = Bxx_0 &= (A - \theta B)yx_0 = (A_0 - \theta B_0)y_0x \\ &= (\bar{A}_0 - \theta \bar{B}_0)xy_0 = (A - \theta B)xy_0 = 0; \end{aligned}$$

so that the invariant factors of $B^{-1}A$ are linear.*

* If θ, θ' be two different roots of $|B^{-1}A - \rho| = 0$, equations

$$Ax = \theta Bx, \quad Ax' = \theta' Bx'$$

give $\theta Bxx'_0 = Ax'_0 = \bar{A}x'_0x = A_0x'_0x = \theta' B_0x'_0x = \theta' \bar{B}_0xx'_0 = \theta' Bxx'_0$,

and hence $(\theta - \theta')Bxx'_0 = 0$, so that $Bxx'_0 = 0$. While, if, for a repeated root, we have $Ax = \theta Bx, Ay = \theta By, Az = \theta Bz, \dots$ and put $\eta = y + \lambda x$, where λ is a single quantity, we have $Bx\eta_0 = Bxy_0 + \lambda_0 Bxx_0$; thus we may suppose $Bxy_0 = 0$, giving $Bxz_0 = 0$; then, putting $\zeta = z + \mu y + \nu x$, when μ, ν are single quantities, we have $Bx\zeta_0 = Bxz_0 + \nu_0 Bxx_0, By\zeta_0 = Bxy_0 + \mu_0 Byy_0$; so we may suppose $Bxz_0 = 0, Byz_0 = 0$, and so on; and this leaves the sets x, y, z, \dots independent. We may then suppose, if μ be the matrix of which any column consists of one of the n independent sets satisfying, for the various values of ρ , the equations $(B^{-1}A - \rho)x = 0$, that, when $i \neq j$,

$$0 = \sum_{r,j} B_{ri} \mu_{rj} (\mu_{rj})_0 = \sum_r (B\mu)_{ri} (\mu_{rj})_0 = (\bar{\mu}_0 B\mu)_{ji},$$

and denote the real quantity $(\bar{\mu}_0 B\mu)_{ii}$ by ϕ_i , and the diagonal matrix $\bar{\mu}_0 B\mu$ by Φ . Then it is a well known consequence of the definition of μ that the matrix $\mu^{-1}B^{-1}A\mu$

3. If S be a matrix such that the bilinear form Sxx_0 does not vanish unless every element of x is zero, and M be a matrix such that $\bar{M}_0 SM = S$, a particular case being when $S = 1$, and thereof a particular case being the ordinary orthogonal transformation when M consists of real elements; then, if $|M - \theta| = 0$ and $Mx = \theta x$, $t = Mx$, $t_0 = \bar{M}_0 x_0$, we have

$$Sxx_0 = \bar{M}_0 SMxx_0 = SMxt_0 = Stt_0 = \theta\theta_0 Sxx_0,$$

so that $\theta\theta_0 = 1$ and the roots θ are of modulus unity. But we cannot have $My = \theta y + x$, or else, noticing that the determinant of S , and therefore of M , and therefore θ , is not zero, we could infer the sequences

$$y = \theta M^{-1}y + M^{-1}x, \quad Sy = \theta SM^{-1}y + SM^{-1}x = \theta \bar{M}_0 Sy + \bar{M}_0 Sx,$$

$$\bar{M}_0 Sy = \theta_0 Sy - \theta_0 \bar{M}_0 Sx,$$

$$\bar{S}M_0 y_0 y = \bar{M}_0 Syy_0 = (\theta_0 Sy - \theta_0 \bar{M}_0 Sx)y_0 = \theta_0 Syy_0 - \theta_0 \bar{M}_0 Sxy_0;$$

$$\bar{S}M_0 y_0 y = \bar{S}(\theta_0 y_0 + x_0)y = (\theta_0 \bar{S}y_0 y + \bar{S}x_0 y = \theta_0 Syy_0 + Sx_0 y;$$

$$\text{and hence deduce} \quad H = \theta_0 \bar{M}_0 Sxy_0 + \bar{S}x_0 y = 0,$$

of which, however, the left side H is the same as

$$\begin{aligned} \theta_0 \bar{S}M_0 y_0 x + \bar{S}x_0 y &= \theta_0 \bar{S}(\theta_0 y_0 + x_0)x + \bar{S}x_0 y \\ &= \theta_0^2 \bar{S}y_0 x + \bar{S}x_0 y + \theta_0 \bar{S}x_0 x \\ &= \theta_0^2 Sxy_0 + \bar{S}x_0 y + \theta_0 Sxx_0 \\ &= \theta_0^2 \bar{M}_0 SMxy_0 + \bar{S}x_0 y + \theta_0 Sxx_0 \\ &= \theta_0^2 \theta \bar{M}_0 Sxy_0 + \bar{S}x_0 y + \theta_0 Sxx_0 \\ &= H + \theta_0 Sxx_0, \end{aligned}$$

giving the impossible result $Sxx_0 = 0$.

is a diagonal matrix, say Θ , having only the roots of $|A - \rho B| = 0$ in its diagonal, each to its own multiplicity; thus

$$\mu^{-1} B^{-1} (A - \rho B) \mu = \Theta - \rho, \quad \bar{\mu}_0 (A - \rho B) \mu = \Phi \Theta - \rho \Phi,$$

wherein, as the equations $(A - \theta B)x = 0$ only determine the ratios of the elements of x , the real quantities ϕ_i in Φ are arbitrary. Putting $\Phi \Theta = \Psi$, $\tau = \mu \Phi$, this is equivalent to

$$(A - \rho B) \tau \tau_0 = (\Psi - \rho \Phi) t t_0,$$

whereby the bilinear forms $A\tau\tau_0$, $B\tau\tau_0$, are simultaneously transformed, each to contain only n terms, of the form $\psi_i t_i (t_i)_0$. In particular when A , B are real symmetric matrices and the quadratic form Bx^2 does not vanish unless $x = 0$, the equations $(A - \theta B)x = 0$ give only real elements for μ , and the real quadratic forms $A\tau^2$, $B\tau^2$ are hereby transformed simultaneously into sums of real positive or negative squares.

Thus the invariant factors of $|M - \rho| = 0$ are linear. (Cf. Loewy, *Nova Acta Kais. Leop. Carol. Deut. Ak. der Naturf.*, LXXI., Halle, 1898, where another proof, involving somewhat more detailed considerations, is given.)

4. One simple example of the results in (2) and (3) deserves notice from its connexion with the theory of algebraically integrable linear differential equations. If α be a matrix as in (2) for which $\bar{\alpha} = \alpha_0$ and $\Omega(\alpha t^{-1})$ denote the matrix

$$\Omega\left(\frac{\alpha}{t}\right) = 1 + \alpha Q(t^{-1}) + \alpha^2 Q(t^{-1}(Q t^{-1})) + \alpha^3 Q(t^{-1}(Q t^{-1}(Q t^{-1}))) + \dots$$

where Q denotes integration from an arbitrary position $t = t_0$, and the integration be extended completely round $t = 0$, back to t_0 , it is easy to see that $\Omega\left(-\frac{\bar{\alpha}}{t}\right)$ is the conjugate complex of $\Omega\left(\frac{\alpha}{t}\right)$.

For an incomplete circuit the inverse of the matrix $\Omega\left(\frac{\alpha}{t}\right)$ can be shown to be the transposed of $\Omega\left(-\frac{\bar{\alpha}}{t}\right)$. Thus, if M denote the complete value of $\Omega\left(\frac{\alpha}{t}\right)$, we have*

$$\bar{M}_0 M = 1;$$

which is simply $\sum_0^\infty \frac{(2\pi i \alpha)^m}{m!}$, and it can be shown that the correspondence of the roots θ , $\phi = e^{2\pi i \theta}$ of the equations $|\alpha - \theta| = 0$, $|M - \phi| = 0$ is complete.

5. If E denote the matrix occurring in the theory of continuous groups for which the general element is

$$E_{\rho\sigma} = \sum_{r=1}^r c_{\sigma\rho} e_r \quad (\rho, \sigma = 1 \dots r),$$

where the constants $c_{\sigma\rho}$ are such that we have identically (*Proc. Lond. Math. Soc.*, Vol. xxxiv., p. 93)

$$Ee' + E'e = 0, \quad EE'e'' + E'E''e + E''Ee' = 0,$$

there is a result of importance (Killing, *Math. Ann.*, Vol. xxxiii., 1889, p. 5) which we can prove briefly by means of a result proved in a previous note (*Proc. Lond. Math. Soc.*, Vol. xxxiv., p. 351). We have identically $Ee = 0$, and, of the determinantal equation $|E + \rho| = 0$,

* [June 17th, 1903.—And conversely any matrix M satisfying this equation can be so written as $\Omega(\alpha t^{-1})$; in particular for a real orthogonal transformation $\alpha = i\beta$, where β is a skew symmetrical real matrix.]

zero is a root of at least multiplicity unity. If, for general values of e_1, \dots, e_r , we form the descending series of positive integers which are the exponents of the invariant factors corresponding to this root zero, the theorem is that the last of them is necessarily unity; so that the last invariant factor is linear. To prove this it is sufficient to show that a set e' of quantities independent of e cannot be found such that, beside $Ee = 0$, we have $Ee' = e$. First, when the root zero of the equation $|E + \rho| = 0$ is of multiplicity r for all values of e_1, \dots, e_r , the equation $Ee' = e$ would give $(E' + 1)e = 0$, and so the impossible equation $|E' + 1| = 0$. Next, when the root zero is at least of multiplicity k ($0 < k < r$) for all values of e_1, \dots, e_r , but is of no higher multiplicity for at least one particular set of values e_1, \dots, e_r , if then e' can be found such that $e = Ee' = -E'e$, and if the series $1 + E + \frac{E^2}{2!} + \dots$ be denoted by Δ_e , we have with arbitrary λ , since $E'^2 e = -E'e = e$, $E'^3 e = E'e = -e$, ..., putting $f = \lambda e'$,

$$\begin{aligned}\Delta_f e &= e + \lambda E'e + \frac{\lambda^2 E'^2 e}{2!} + \frac{\lambda^3 E'^3 e}{3!} + \dots \\ &= e - \lambda + \frac{\lambda^2 e}{2!} - \frac{\lambda^3 e}{3!} + \dots \\ &= \mu e,\end{aligned}$$

where $\mu = \exp(-\lambda)$. It has been proved, however, that, if $e' = \Delta_f e$, then

$$E' = \Delta_f E \Delta_f^{-1};$$

thus here we have $\mu E = E_{\mu} = \Delta_f E \Delta_f^{-1}$,

and so $|E + \rho| = |\mu E + \rho|$.

The coefficients in the equation $|E + \rho| = 0$ are therefore unaltered when e_1, \dots, e_r are replaced respectively by $\mu e_1, \dots, \mu e_r$. As $\mu = \exp(-\lambda)$ and λ is arbitrary, this can only be so if the equation reduces to $\theta'' = 0$, which by hypothesis is not the case.

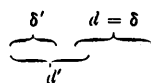
6. The result in (5) is for general values of e_1, \dots, e_r . Consider an integrable group, and let e_1, \dots, e_r be a particular set for which $P = e_1 X_1 + \dots + e_r X_r$ is the infinitesimal transformation such that when P' is any other the combinant (P', P) is zero or a constant multiple of P , so that we have $Ee' = \lambda e$; if then $Ee'' = \mu e$, and therefore $E(\mu e' - \lambda e'') = 0$, and neither of λ, μ be zero, we can replace one of e', e'' by $\mu e' - \lambda e''$. In other words, there is at most only one independent transformation P' for which the combinant (P', P) is not

zero. But the proof in (5) avails when $Ee' = \lambda e$ is not zero to show that for this special e_1, \dots, e_r the equation $|E + \rho| = 0$ has no root but zero; thus, besides the equation $Ee' = \lambda e$, we have $(r-2)$ equations $Ef = 0$, together with $Ee = 0$. In other words, the first invariant factor of $|E + \rho| = 0$ is of exponent 2, and the remaining ones are linear, and the matrix E satisfies the equation $E^2 = 0$. For a group known to be necessarily integrable, such that $|E + \rho| = 0$ reduces to $\rho^r = 0$ for all values of e_1, \dots, e_r , it can be shown that a set e_1, \dots, e_r exists, not identically zero, for which $E = 0$.

Overlapping Intervals. By W. H. Young. Received November 30th, 1902. Read December 11th, 1902. Revised March 10th, 1903.

1. Given any set of overlapping intervals, we will show how to determine a countable set from among them which by themselves determine the most important properties of the given set.

Take, first, any one of the intervals, and let us denote it by d or δ . Then either there is no interval of the given set which abuts or overlaps with d on the left, or else there is such an interval. In the latter case we denote by δ' the part of this interval which extends beyond d to the left, and by d' the interval itself, which coincides with δ' if d' abuts with d , and which otherwise contains δ' as a part, and has its left-hand end point coincident with that of δ' .



Proceeding in this way towards the left, we must ultimately either come to an interval of the given set having no interval abutting or overlapping with it on the left, or else the parts of intervals $\delta', \delta'', \delta''', \dots$ must get smaller and smaller without limit, and define a limiting point P external to all of them, and therefore external to the intervals d', d'', d''', \dots of the given set. Such a point P may, however, be internal to some other interval of the given set; in this case we choose out any one of the intervals containing P , say D . There will only be a finite number of the intervals d, d', d'', \dots