BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 8, Number 1, January 1983

SIMULTANEOUS SIMILARITY OF MATRICES

SHMUEL FRIEDLAND

Let M_n be the set of $n \times n$ matrices over the algebraically closed field k, G_n the general linear group in M_n, $M_{n,m} = M_n \times \cdots \times M_n (m+1 \text{ times})$. G_n acts naturally on $M_{n,m}$ by the conjugation $TM_{n,m}T^{-1}$. For $\alpha = (A_0, \ldots, A_m) \in M_{n,m}$ denote by $\text{orb}(\alpha)$ the orbit of α in $M_{n,m}$,

$$\text{orb}(\alpha) = \{\beta \in M_{n,m}, \beta = T\alpha T^{-1} = (TA_0T^{-1}, \ldots, TA_mT^{-1}), T \in GL_n\}.$$

It is a well-known problem to classify $\text{orb}(\alpha)$ for $m \geq 1$. See for example [**2**]. Rosenlicht in [**3**] outlined a general classification based on the ideas of algebraic geometry. The classification consists of a finite number of steps. In each step we get an algebraic irreducible variety V in $M_{n,m}$ which is invariant, that is $TVT^{-1} = V$ for all $T \in G_n$. Then, we consider $k(V)^G$—the field of rational functions on V which are invariant, i.e. these functions are constant on $\text{orb}(\alpha)$. It follows that $k(V)^G$ is finitely generated, let us say by $\chi_1, \ldots, \chi_j$. Then there exists locally closed algebraic invariant set V^0 in V such that for any $\alpha \in V^0 \chi_1, \ldots, \chi_j$ are well defined on $\text{orb}(\alpha)$ and the values of χ_k, $k = 1, \ldots, j$, on $\text{orb}(\alpha)$ determine this orbit uniquely in V^0.

The purpose of this announcement is to describe explicitly the open invariant varieties V^0 together with the invariant rational functions $\varphi_1, \ldots, \varphi_k$ defined on V^0 such that the values of $\varphi_1, \ldots, \varphi_k$ on $\text{orb}(\alpha)$ determine a finite number of orbits. We also describe some results on orbits in $S_{n,m} = S_n \times \cdots \times S_n$ ($m+1$ times) (S_n = the set of $n \times n$ complex symmetric matrices) under the action of O_n-complex orthogonal group in M_n.

For $\alpha = (A_0, \ldots, A_m)$, $\beta = (B_0, \ldots, B_m)$ let $\text{adj}(\alpha, \beta)$: $M_n \to M_{n,m}$ be a linear operator given by $\text{adj}(\alpha, \beta)(X) = (A_0X - XB_0, \ldots, A_mX - XB_m)$.

We identify $\text{adj}(\alpha, \alpha)$ with $\text{adj}(\alpha)$. Let $r(\alpha, \beta)$ and $r(\alpha)$ be the ranks of $\text{adj}(\alpha, \beta)$ and $\text{adj}(\alpha)$ respectively. Then $r(\alpha)$ is the first discrete invariant of $\text{orb}(\alpha)$ and it gives the dimension of the manifold $\text{orb}(\alpha)$. Suppose that $\beta \in \text{orb}(\alpha)$. Then one easily shows that $r(\alpha, \beta) = r(\alpha)$. Fix α and consider all $\xi \in M_{n,m}$ which satisfy the inequality

$$\mathcal{X}(\alpha) = \{\xi, r(\alpha, \xi) \leq r, \ \xi = (X_0, \ldots, X_m) \in M_{n,m}\}. \tag{1}$$

The set $\mathcal{X}(\alpha)$ is an algebraic set in $M_{n,m}$ which can be given by

$$N(r) = \binom{n^2}{r+1}\binom{n^2 \quad (m+1)}{r+1} \text{polynomial equations.}$$

Received by the editors December 22, 1981.

1980 *Mathematics Subject Classification*. Primary 14D25, 14L30, 15A21.

Key words and phrases. Simultaneous similarity, invariant functions, symmetric matrices.

0273-0979/82/0000-1031/$01.50

Indeed, in tensor notation, $\mathrm{adj}(\alpha, \xi)$ is represented as the following matrix

$$\mathrm{adj}(\alpha, \xi) = (I \otimes A_0 - X_0^t \otimes I, \ldots, I \otimes A_m - X_m^t \otimes I)$$

where X^t denotes the transposed matrix of X. Let $f_1(\alpha, \xi), \ldots, f_p(\alpha, \xi)$, $p = N(r)$ be all $(r+1)\times(r+1)$ minors of $\mathrm{adj}(\alpha, \xi)$. Then (1) is given by the equations $f_i(\alpha, \xi) = 0$, $i = 1, \ldots, N(r)$. Let $\mathcal{W}_r$ be a linear space of all polynomials $p(\xi)$—in the $(m+1)n^2$ entries of $X_0, \ldots, X_m$ of degree $d \leq r+1$. Denote by $u_1 = u_1(\xi), \ldots, u_{s(r)} = u_{s(r)}(\xi)$ the standard basis in $\mathcal{W}_r$. Then

$$f_i(\alpha, \xi) = \sum_{j=1}^{s(r)} \pi_{ij}^{(r)}(\alpha) u_j(\xi), \qquad i = 1, \ldots, N(r). \tag{2}$$

Put $\pi^{(r)}(\alpha) = (\pi_{ij}^{(r)}(\alpha))$, $i = 1, \ldots, N(r)$, $j = 1, \ldots, s(r)$. Then

$$\rho(\alpha) = \mathrm{rank}\, \pi^{(r(\alpha))}(\alpha)$$

is the second discrete invariant of $\mathrm{orb}(\alpha)$. Define

$$V_{r,\rho}^0 = \{\alpha, \alpha \in M_{n,m}, \mathrm{rank}\, \mathrm{adj}(\alpha) = r, \mathrm{rank}\, \pi^{(r)}(\alpha) = \rho\}. \tag{3}$$

Then $V_{r,\rho}^0$ is an open algebraic set in $M_{n,m}$. (It may be empty for some choices of r and ρ.)

Finally, we recall that two $p \times q$ rectangular matrices A and B are row equivalent ($A \sim B$) if there exists a nonsingular matrix Q such that $B = QA$. Any $p \times q$ matrix A can be brought to the unique row-echelon form E using the elementary row operations. $E = (e_{ij})$ is characterized by $1 \leq p_1 < \cdots < p_\rho \leq q$, $\rho = \mathrm{rank}\, E$, since $e_{ip_i} = 1$, $e_{jp_i} = e_{iq} = 0$ for $j < i$, $q < p_i$ and $i > \rho$. The integers $p_1, \ldots, p_\rho$ are called the discrete invariants of E and the entries e_{ij}, $p_i < j$, $j \neq p_{i+1}, \ldots, p_\rho$, for $i = 1, \ldots, \rho$ are called the continuous invariants of E. Once $p_1, \ldots, p_\rho$ are specified these invariants are given as well-determined rational functions of entries of E.

THEOREM 1. *Assume that $V_{r,\rho}^0$ is nonempty. Let $\alpha, \beta \in V_{r,\rho}^0$. If $\beta \in \mathrm{orb}(\alpha)$ then $\pi^{(r)}(\alpha) \sim \pi^{(r)}(\beta)$. Moreover, there are at most $\kappa = r^{(n^2-r)(mn^2+n^2-r)}$ distinct orbits $\mathrm{orb}(\alpha_1), \ldots, \mathrm{orb}(\alpha_k)$ such that $\pi^{(r)}(\alpha_1), \ldots, \pi^{(r)}(\alpha_k)$ have the same row-echelon form.*

SKETCH OF THE PROOF. We first note that if $\beta \in \mathrm{orb}(\alpha)$ then $\pi^{(r)}(\alpha) \sim \pi^{(r)}(\beta)$. Indeed, since $\beta = T\alpha T^{-1}$ the tensor representation of $\mathrm{adj}(\alpha, \xi)$ yields $\mathrm{adj}(\beta, \xi) = T_1\, \mathrm{adj}(\alpha, \xi)\, \mathrm{diag}\{T_1^{-1}, \ldots, T_1^{-1}\}$, $T_1 = I \otimes T$. The Cauchy-Binet formula implies that any minor of $\mathrm{adj}(\beta, \xi)$ is a linear combination of all $(r+1) \times (r+1)$ minors of $\mathrm{adj}(\alpha, \xi)$ and the coefficients in this dependence are functions of T, i.e. independent of ξ! Whence the subspace spanned by the rows of $\pi^{(r)}(\alpha)$ contains the rows of $\pi^{(r)}(\beta)$. Interchanging the roles of α and β we get $\pi^{(r)}(\alpha) \sim \pi^{(r)}(\beta)$. Fix α. We then show the existence of a neighborhood $D(\alpha)$ such that the conditions $\beta \in D(\alpha)$ and $\pi^{(r)}(\alpha) \sim \pi^{(r)}(\beta)$ imply that $\beta \in \mathrm{orb}(\alpha)$. For that, in the matrix $\mathrm{adj}(\alpha)$ pick up a nonzero $r \times r$ minor. We then consider the corresponding r linear equations out of $(m+1)n^2$ equations $A_i X - X A_i = 0$, $i = 0, \ldots, m$. This r-system has $n^2 - r$ free parameters x_{ij}, $(i,j) \in \mathcal{A}(X = (x_{ij}))$. Since $X = I$ is a solution, the above system has the unique solution $X = I$ whose free parameters are given by $x_{ij} = \delta_{ij}$, $(i,j) \in \mathcal{A}$. Consider the same r-equations in a more general system $A_iX - XB_i = 0$, $i = 0, \ldots, m$. Thus, there exists a neighborhood $D(\alpha)$ of α in $M_{n,m}$ such that for any $\beta \in D(\alpha)$ the above r-system is linearly independent and has the unique solution $X(\alpha, \beta)$, $x_{ij} = \delta_{ij}$, $(i,j) \in \mathcal{A}$ with $\det X(\alpha, \beta) \neq 0$. Suppose that $\pi^{(r)}(\alpha) \sim \pi^{(r)}(\beta)$. So each $(r+1) \times (r+1)$ minor of $\operatorname{adj}(\beta, \alpha)$ is a linear combination of all $(r+1) \times (r+1)$ minors of $\operatorname{adj}(\alpha, \alpha)$ which are equal to zero! So $\operatorname{rank} \operatorname{adj}(\alpha, \beta) = \operatorname{rank} \operatorname{adj}(\beta, \alpha) \leq r$. If in addition $\beta \in D(\alpha)$ then $\operatorname{rank} \operatorname{adj}(\alpha, \beta) = r$ and the matrix $X(\alpha, \beta)$ must satisfy all $(m+1)n^2$ equalities $A_iX - XB_i = 0$, $i = 0, \ldots, m$. So $\beta \in \operatorname{orb}(\alpha)$. Consider finally the variety $\mathcal{X} = \mathcal{X}(\alpha)$. Let $\mathcal{X} = \bigcup_{i=1}^{k} \mathcal{X}_i$ be the decomposition of $\mathcal{X}$ into irreducible components. To this end we show that each $\mathcal{X}_i$ contains at most one orbit. Assume that $\alpha \in \mathcal{X}_1$ and let $\mathcal{X}_1^0$ be the open manifold of all regular points of $\mathcal{X}_1$. The above arguments prove that $D(\alpha) \cap \mathcal{X}_1^0 \subset \operatorname{orb}(\alpha)$. On the other hand $\operatorname{orb}(\alpha) \subset \mathcal{X}_1$. As $\mathcal{X}_1^0$ and $\operatorname{orb}(\alpha)$ are connected we deduce that $\operatorname{orb}(\alpha) = \mathcal{X}_1^0$. A simple degree argument shows that $k \leq \kappa$. Therefore we have at most κ distinct orbits. □

Let $1 \leq p_1 < p_2 < \cdots < p_\rho \leq q = s(r)$. Let $V^0_{r,\rho,p_1,\ldots,p_\rho}$ be the set of all $\alpha \in V^0_{r,\rho}$ whose row-echelon form of $\pi^{(r)}(\alpha)$ has the discrete invariants $p_1, \ldots, p_\rho$. Then the entries e_{ij}, $p_i < j$, $j \neq p_{i+1}, \ldots, p_\rho$, $i = 1, \ldots, \rho$, in the row-echelon form the invariant rational functions which determine the $\operatorname{orb}(\alpha)$ up to κ orbits at most. In fact, we conjecture that if α and β lie in the same connected component of $V^0_{r,\rho}$ and $\pi^{(r)}(\alpha) \sim \pi^{(r)}(\beta)$ then $\operatorname{orb}(\alpha) = \operatorname{orb}(\beta)$.

For $\alpha \in S_{n,m}$ let $\operatorname{sorb}(\alpha) = \{\beta, \beta = T\alpha T^{-1},\ T \in O_n\}$, $\alpha(z) = \sum_{i=0}^{m} A_i z^i$, where $z \in C$ (the field of complex numbers). Let $p(\lambda, z) = \det(\lambda I - \alpha(z))$ be the characteristic polynomial of α. Clearly $p(\lambda, z)$ is invariant on $\operatorname{sorb}(\alpha)$ or $\operatorname{orb}(\alpha)$. It can be shown that for most $\alpha \in S_{n,m}$ the equation $p(\lambda, z) = 0$ (α is fixed) will have n distinct λ roots for all except a finite number of z, possibly $z = \infty$ ($p(\lambda, \infty) = \det(\lambda I - A_m)$) and at those exceptional points the equation $p(\lambda, z) = 0$ will not have triple roots. We call such α and corresponding $p(\lambda, z)$ simple.

THEOREM 2. *There are at most* $2^{(n-1)(mn-1)}$ *distinct* $\operatorname{sorb}(\alpha_1), \ldots, \operatorname{sorb}(\alpha_k)$ *such that all these orbits have the same simple characteristic polynomial.*

We conjecture that if $A_0, \ldots, A_m$ are real symmetric then $\operatorname{sorb}(\alpha)$ is determined by its characteristic polynomial up to a finite number of orbits.

The detailed results are given in [**1**].

REFERENCES

1. S. Friedland, *Simultaneous similarity of matrices*, Advances in Math. (to appear).
2. I. M. Gelfand, *The cohomology of infinite dimensional Lie algebras, some questions of integral geometry*, Actes, Congr. Internat. Math. **1** (1970), 95–111.
3. M. Rosenlicht, *Some basic theorems on algebraic groups*, Amer. J. Math. **78** (1956), 401–443.

INSTITUTE OF MATHEMATICS, HEBREW UNIVERSITY, JERUSALEM, ISRAEL

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN-MILWAUKEE, MILWAUKEE, WISCONSIN 53201