DISCONTINUOUS GROUPS AND ALLIED TOPICS, III: ON A LEMMA ABOUT MATRICES

BY MAX ZORN

1. **Introduction.** In connection with the theory of hypercomplex units the following problem arises. Given a linear family $\mathfrak{A}$ of *real* matrices, which contains with every A its transpose A^T and with every regular A its inverse A^{-1}. Supposing that $\mathfrak{A}$ contains at least one unimodular matrix A_1, $|A_1| = 1$, to determine the unimodular matrices in $\mathfrak{A}$ for which the sum of the squares of the elements is a minimum. The results and proofs are extremely simple. The minimizing matrices are identical with the orthogonal matrices in the family. This is shown by application of Lagrange's multiplier rule. The more or less algebraic machinery may very well be extended to more general linear spaces. A. D. Michal expresses the opinion that the full theory might be generalized; yet it seems that at present the necessary existence theorems are not available. Apart from possible arithmetical applications the existence of at least one orthogonal matrix in such families is of a certain independent interest.

2. **Notations.** The determinant $|a_{ik}|$ of the matrix $A = (a_{ik})$ is as usual denoted by $|A|$, the inverse by A^{-1}, the transpose (a_{ki}) by A^T. The indices i, k range from 1 to n, the number of rows or columns of the matrices. The Greek index ranges from 1 to m, unless otherwise stated. The trace $t(A)$ is the sum $\sum a_{ii}$ of the diagonal elements. For example, the trace $t(I)$ of the unit matrix is equal to n.

A linear combination $\sum_{i,k} c_{ik}x_{ik}$ of symbols x_{ik} with coefficients c_{ik} is expressible in the form $t(XC^T)$, if we let $X = (x_{ik})$, $C = (c_{ik})$.

The sum of the squares of all elements is $t(XX^T)$. Since we shall deal only with real matrices, $t(XX^T)$ is positive and vanishes only for $X = 0$. A matrix A is orthogonal if $X^{-1} = X^T$ and if $|X| = 1$. For an orthogonal matrix XX^T is I, and $t(XX^T) = n$.

3. **Differentials.** We shall use x_{ik} as independent variables and dx_{ik} as their differentials.

For a matrix $F = (f_{ik})$ of functions we define $dF = (df_{ik})$. In particular we shall have to make use of

$$dX = d(x_{ik}) = (dx_{ik}),$$
$$dt(XC^T) = t(dXC^T),$$
$$dt(XX^T) = 2t(X^TdX).$$

Received August 17, 1938; presented to the American Mathematical Society, April 9, 1938; cf. Bulletin of the American Mathematical Society, vol. 44(1938), page 354.