REVERSIBLE TRANSFORMATIONS

DANIEL C. LEWIS, JR.

1. Introduction and summary of results. The concept of a reversible transformation has played a fairly important part in the writings of G. D. Birkhoff on Dynamics. It is shown here in Theorems 1 and 2 that this concept is categorical if, and only if, it is formulated relative to a specified group. For, according to Theorem 1, lack of this specification leads to the conclusion that *every* transformation is reversible, while Theorem 2 provides an example to show that some, but not all transformations, are reversible when the group is suitably restricted. The same is true regarding the representation of a transformation as the product of two involutory transformations.

Theorem 3 states that any linear transformation of a finite dimensional vector space onto itself which is reversible in the group of linear transformations can be expressed as the product of two linear involutory transformations. It is not known if an analogous theorem holds for non-linear transformations and suitably restricted groups. A converse proposition is to the effect that any transformation expressible as the product of two involutory transformations is reversible in any group containing the two involutory factors. This proposition is well known and indeed is quite trivial.

The rest of the paper mainly concerns the utility of the property of reversibility in detecting points which are invariant under iterates of a reversible transformation. G. D. Birkhoff had used these methods when the transformation was factorable into two involutory transformations. It is shown here that this assumption of factorability is superfluous as long as the transformation is known to be reversible in a suitable group.

2. Definition and significance of reversibility. A one-to-one transformation T of a set S onto itself is said to be reversible in a group G of one-to-one transformations of S onto itself, if $T \in G$ and if there is a tranformations $U \in G$ such that $T^{-1} = UTU^{-1}$. In the important special case, where U is involutory, i.e., $U^{-1} = \mathrm{U}$, this relationship may be written $(UT)(UT) = I$, where I is the identity. Hence, if we set $V = UT$, we see that V is also involutory. Moreover we have $T = U^{-1}V = UV$, so that T is the product of two involutory transformations. Conversely, if T is the product of two involutory

Received October 20, 1960. This research was partially supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command, under Contract Number AF 49(638)-382. Reproduction in whole or in part is permitted for any purpose of the United States Government.

transformations, $T = UV$, with $U^2 = V^2 = I$, we can write $T^{-1} = VU = (U^2)VU = U(UV)U = U(UV)U^{-1} = UTU^{-1}$, so that T is reversible in the group generated by U and V.

These facts were noticed by G. D. Birkhoff [3] in whose writings there are repeated references to reversible transformations, especially the case in which $U = U^{-1}$. These begin with his celebrated paper of 1914 on the restricted problem of three bodies and end in a posthumous paper of 1945 written jointly with Jaime Lifshitz. The group G with which he was concerned was usually a group of analytic transformations. It should be emphasized that the group G under which a transformation is stated to be reversible should *always* be specified. For according to the following theorem *every* transformation is reversible in the group of all one-to-one transformations.

THEOREM 1. *Every one-to-one transformation T of a set S onto itself can be represented in the form $T = UV$, where U and V are involutory transformations of S onto itself. Hence, every one-to-one transformation of S onto itself is reversible in the group of one-to-one transformations of S onto itself.*

Comment. This theorem is not of a topological nature. S does not even have to be a topological space. If S is a topological space and if T is continuous, the U and V, whose existence is asserted by the theorem, do not have to be continuous.

Proof of Theorem. An "orbit" of a point P of S is defined as the set, $\cdots, T^{-2}P, T^{-1}P, P, T^1P, T^2P, \cdots$. P is called the "initial point" of the orbit. Evidently orbits of two points P and Q coincide with each other if and only if $Q = T^kP$ for some integer k (positive, negative, or zero); and, if this condition is not fulfilled, the two orbits are disjoint. S may thus be conceived as the union of a (possibly uncountably infinite) set W of disjoint orbits.

By the axiom of choice, we define a point-set M by attributing to it just one point from each orbit of W, and, in the sequel, we regard $P \in M$ as the initial point of its orbit. The transformation V is defined in such a way that every point of M is invariant under V and (more generally) in such wise that (with $P \in M$) it sends T^kP into $T^{-k}P$ for every integer k. This definition is consistent; for, if $T^kP = T^hP$ for distinct integers k and h, it follows that $P = T^{h-k}P$ and hence that $T^{-k}P = T^{-h}P$. This definition also makes V involutory; for it sends $T^{-k}P$ into T^kP. Moreover we observe that

$$TVTV = I. \tag{1}$$

For, if Q is an arbitary point of S, we many write $Q = T^kP$ for some

point $P \in M$ and some integer k, and then

$$\begin{aligned} V &\text{ sends } Q = T^k P \text{ into } T^{-k}P \\ T &\text{ sends } T^{-k}P \text{ into } T^{-k+1}P \\ V &\text{ sends } T^{-k+1}P \text{ into } T^{k-1}P \\ T &\text{ sends } T^{k-1}P \text{ into } T^k P = Q\ . \end{aligned}$$

Thus $TVTV$ sends Q into itself and (1) is established. We next define

$$U = TV, \tag{2}$$

which, according to (1), is involutory. Since V is also involutory, as previously noted, we have $V^{-1} = V$, so that from (2) we obtain $T = UV$, as we desired to prove.

We next state a theorem which definitely shows that, when the group G is more restricted, not every transformation is reversible. A fortiori it can not be expressed as a product of two involutory tranformations in the group.

THEOREM 2. *A linear transformation of an n-dimensional vector space onto itself, expressed by the equation $y = Ax$, where A is a non-singular $n \times n$-matrix, can not be reversible in the group G of all non-singular linear transformations unless the set of eigenvalues of A is unchanged when each eigenvalue is replaced by its reciprocal.*

Proof. If our linear transformation is reversible in the specified group, our definition requires the existence of a non-singular matrix B such that $A^{-1} = BAB^{-1}$. But the eigenvalues of A are the same as those of BAB^{-1}, and hence also the same as those of A^{-1}. But the eigenvalues of A^{-1} are the reciprocals of those of A. Hence the theorem.

We leave to the reader the simple task of showing that the set of reversible linear transformations (in the linear group) is far from vacuous.

A corollary of Theorem 2 is to the effect that a non-singular transformation T of class C^k ($k \geqq 1$) of a neighborhood of an invariant point in Euclidean n-dimensional space onto such a neighborhood is not necessarily reversible in the group of all such transformations of class C^k. For we can develop the functions defining the transformation T by Taylor's theorem. It is then easily seen that the linear terms define a tranformation of the type considered in Theorem 2, which is reversible if T is reversible. On the other hand a transformation T can be set up having arbitrary linear terms and hence certainly can not be reversible if the linear terms are deliberately chosen so as to violate the condition of Theorem 2.

Strictly speaking the discussion of the last paragraph concerns germs of transformations instead of transformations themselves. This is because not all transformations considered need be defined in the same neighborhood and hence need not form a group. However equivalence classes (the germs) of these transformations may be introduced in a well known and obvious way so as to form a group G. Namely, two transformations are said to be equivalent if they coincide in a neighborhood of the invariant point. A non-singular transformation T of a neighborhood of an invariant point 0 onto a neighborhood of 0 is said to be reversible in a group G of germs of transformations of such neighborhoods into themselves, if T belongs to a germ of G and if there is a transformation U also belonging to a germ of G such that $T^- = UTU^{-1}$.

3. On the factorization of reversible linear transformations.

THEOREM 3. *If a linear transformation is reversible in the group of linear transformations, it can be written as the product of two involutory linear transformations.*

Proof. Given that $A^{-1} = BAB^{-1}$, where A and B are non-singular $n \times n$-matrices it is sufficient to show that there is an involutory matrix C such that

$$A^{-1} = CAC \,. \tag{3}$$

For, then we could set $D = AC$, which, by (3), would be involutory and thus $A = DC^{-1} = DC$, giving the required factorization.

A slight extension of the argument used in the proof of Theorm 2 shows that the elementary divisors associated with a characteristic root λ of A must be, if A is reversible, similar to the elementary divisors associated with the characteristic root λ^{-1}. Hence, expressing A in Jordan canonical form, we write

$$A = \begin{pmatrix} \alpha & 0 & 0 & \cdots \\ 0 & \bar{\alpha} & 0 & \cdots \\ 0 & 0 & \beta & \cdots \\ \cdots & & & \\ \cdots & & & \end{pmatrix}. \tag{4}$$

Here $\alpha, \bar{\alpha}, \beta, \cdots$ stand for certain square blocks of elements, while 0 is a general symbol standing for a square or rectangular block of zeros. Moreover the explicit forms of α and $\bar{\alpha}$, which are matrices of the same order, are given as follows:

$$(5)\qquad \alpha = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \cdots & 0 \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 0 & 0 & 0 & & \lambda \end{pmatrix}, \qquad \bar{\alpha} = \begin{pmatrix} \lambda^{-1} & 1 & 0 & \cdots & 0 \\ 0 & \lambda^{-1} & 1 & \cdots & 0 \\ 0 & 0 & \lambda^{-1} & \cdots & 0 \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 0 & 0 & 0 & & \lambda^{-1} \end{pmatrix}.$$

If the eigenvalue λ is $+1$ or -1, it is self reciprocal, and the block $\bar{\alpha}$ may be missing from the scheme indicated in (4). Otherwise $\bar{\alpha}$ must be present as shown.

It is clear from (4) that our matrix A is the direct product of smaller matrices. Hence it is sufficient to establish (3) in the following two cases:

Case 1.

$$(6)\qquad A = \begin{pmatrix} \alpha & 0 \\ 0 & \bar{\alpha} \end{pmatrix}$$

where α and $\bar{\alpha}$ are given by (5) and λ need not be ± 1.

Case 2.

$$(7)\qquad A = \alpha$$

where α is again given by (5) but $\lambda = \pm 1$.

We consider first Case 1. Let $\gamma(\theta)$ denote the square matrix of the same order as α, the element in whose kth row and hth column is $\binom{h-1}{k-1}(=1)^{h-1}\theta^{h+k-2}$. Here we use the usual notation for the binomial coefficients, namely $\binom{p}{q} = \dfrac{p!}{q!(p-q)!}$ for integers p and q with $p \geqq 0$ and $0 \leqq q \leqq p$. If $q < 0$ or $q > p$, we set $\binom{p}{q} = 0$, a consequence of which is that all the elements below the main diagonal of $\gamma(\theta)$ are zero. We also similarly define a matrix $\zeta(\theta)$, of the same order as α, the element in whose kth row and hth column is $\binom{h}{k}(-1)^{h-1}\theta^{h+k-1}$.

We recall three elementary properties of the binomial coefficients:

$$\text{I}\qquad \binom{p+1}{q} = \binom{p}{q} + \binom{p}{q-1}$$

$$\text{II}\qquad \sum_{l=0}^{q}(-1)^l\binom{p}{q-l} = \binom{p-1}{q}$$

III $$\sum_{p=q}^{k}(-1)^p\binom{p}{q}\binom{k}{p}=(-1)^k\delta_{kq}\,,$$

where $\delta_{kq}=0$ if $k\neq q$ and $\delta_{qq}=1$. (Properties II and III many be established by applying the binomial theorem to the appropriate expressions in the following identities and then equating coefficients of like powers of z: $(1+z)^{-1}(1+z)^p\equiv(1+z)^{p-1}$ is used for II and $[1-(1+z)]^k\equiv(-z)^k$ is used for III).

Using Property I, we show by a routine calculation, the details of which are left to the reader, that $\alpha\gamma(\lambda)=\zeta(\lambda)$. Using Property II and the fact that

$$\bar{\alpha}^{-1}=\begin{pmatrix}\lambda-\lambda^2 & \lambda^3 & \cdots\\ 0 & \lambda-\lambda^2 & \cdots\\ 0 & 0 & \lambda & \cdots\\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \end{pmatrix}$$

we find similarly that $\gamma(\lambda)\bar{\alpha}^{-1}=\zeta(\lambda)$. Thus, we have

$$\gamma(\lambda)\bar{\alpha}^{-1}=\alpha\gamma(\lambda)$$

and we may similarly prove that

$$\gamma(\lambda^{-1})\alpha^{-1}=\bar{\alpha}\gamma(\lambda^{-1})\,.$$

Again a routine calculation based on Property III shows that $\gamma(\lambda)$ and $\gamma(\lambda^{-1})$ are inverses of each other. We therefore let $\Gamma=\gamma(\lambda)$ and $\Gamma^{-1}=\gamma(\lambda^{-1})$, so that $\bar{\alpha}^{-1}=\Gamma^{-1}a\Gamma$ and $\alpha^{-1}=\Gamma\bar{\alpha}\Gamma^{-1}$. Hence from (6) we see that

$$\begin{aligned}A^{-1}&=\begin{pmatrix}\alpha^{-1} & 0\\ 0 & \bar{\alpha}^{-1}\end{pmatrix}=\begin{pmatrix}\Gamma\bar{\alpha}\Gamma^{-1} & 0\\ 0 & \Gamma^{-1}\alpha\Gamma\end{pmatrix}\\ &=\begin{pmatrix}\Gamma & 0\\ 0 & \Gamma^{-1}\end{pmatrix}\begin{pmatrix}\bar{\alpha} & 0\\ 0 & \alpha\end{pmatrix}\begin{pmatrix}\Gamma^{-1} & 0\\ 0 & \Gamma\end{pmatrix}\\ &=\begin{pmatrix}\Gamma & 0\\ 0 & \Gamma^{-1}\end{pmatrix}\begin{pmatrix}0 & I\\ I & 0\end{pmatrix}\begin{pmatrix}\alpha & 0\\ 0 & \bar{\alpha}\end{pmatrix}\begin{pmatrix}0 & I\\ I & 0\end{pmatrix}\begin{pmatrix}\Gamma^{-1} & 0\\ 0 & \Gamma\end{pmatrix}\\ &=\begin{pmatrix}0 & \Gamma\\ \Gamma^{-1} & 0\end{pmatrix}A\begin{pmatrix}0 & \Gamma\\ \Gamma^{-1} & 0\end{pmatrix}.\end{aligned}$$

Since $\begin{pmatrix}0 & \Gamma\\ \Gamma^{-1} & 0\end{pmatrix}$ is obviously involutory, we have thus completed the proof of (3) in Case 1 by taking $C=\begin{pmatrix}0 & \Gamma\\ \Gamma^{-1} & 0\end{pmatrix}$.

The proof in Case 2 also uses the matrix $\gamma(\lambda)$ defined above; only

here, since $\lambda = \pm 1$, $\Gamma = \gamma(\pm 1)$ must be its own inverse. We then easily find from (7) that $A^{-1} = \alpha^{-1} = \Gamma\alpha\Gamma$, which establishes (3) also Case 2 by taking $C = \Gamma$.

Both cases having been settled, the proof of Theorem 3 is now complete.

It is not known if Theorem 3 remains true when, istead of considering linear transformations, we consider germs of non-singular transformations of a neighborhood of the origin into other neighborhoods of the origin, which tranformations are assumed to be either analytic or differentiable up to a certain order.

4. Invariant points of the iterates of a reversible transformation. The factorization of a reversible transformation T into two involutory factors has been considered useful in finding points that are invariant under the iterates of T in accordance with the following Theorem 4 proved by G. D. Birkhoff and Jaime Lifshitz. In Theorem 7, however, we give a generalization of Theorem 4, which shows that, for this purpose, we may dispense with such a factorization.

THEOREM 4. *Suppose that T, U, V are one-to-one transformations of a set S onto itself, that $T = UV$, and that U and V are both involutory. Let Λ be the set of points, each of which is invariant under U and let N be the set of points, each of which is invariant under V. Then any point common to $T^k\Lambda$ and $T^l\Lambda$ is invariant under $T^{2(l-k)}$; any point common to $T^k\Lambda$ and T^lN is invariant under $T^{2(l-k)-1}$; and any point common to T^kN and T^lN is invariant under $T^{2(l-k)}$. Here k and l represent any two integers, positive, negative, or zero.*

Notice that, like Theorem 1, this theorem is not of a topological nature; S does not have to be a topological space. We here omit the proof of Theorem 4 partly because, as stated above, a prior proof has been given by Birkhoff and Lifshitz (even though these authors did not explicitly formulate the theorem in precisely these terms) and partly because several of our following theorems contain Theorem 4 as a special case. It should also be stated that a proof of a special case of Theorem 4 also appears on the last page of a paper of G. D. Birkhoff written in 1931 (cf. Reference 3). It is remarkable that this simple proof is not really much simpler than the proof of the following considerably more general theorem.

THEOREM 5. *Suppose that T, U, W are one-to-one transformations of a set S onto itself and that T and U are related as follows*

$$T^{-1} = UTU^{-1}, \tag{8}$$

(so that T is reversible in any group of transformations containing both T and U). Let Λ be the set of points, each of which is invariant under U and let M be the set of points, each of which is invariant under W. Then any point common to $T^k\Lambda$ and T^lM is invariant under $T^lW^{-1}UT^{l-2k}$.

Comment. In comparing this theorem with Theorem 4, we take $V = UT$. Theorem 5 in the special case, in which U and hence V, because of (8), are involutory and in which $W = U$, $M = \Lambda$, reduces to the first part of Theorem 4. In the special case in which $W = V$, $M = N$, U, and hence V, being still involutory, Theorem 5 reduces to the second part of Theorem 4. The third and last part of Theorem 4 is, of course, an obvious consequence of the first part of Theorem 4, by interchanging the roles of T and T^{-1}, of U and V, and of Λ and N.

Proof of Theorem 5. From (8), we have $T^{-(l-k)} = UT^{(l-k)}U^{-1}$, which may also be written

$$T^{k-l}U = UT^{l-k}. \tag{9}$$

Suppose $P \in (T^k\Lambda) \cap (T^lM)$. Then there exist points Q and R, such that $Q \in \Lambda$ and $R \in M$ and such that

$$P = T^kQ \tag{10}$$

and

$$P = T^lR\,. \tag{11}$$

Moreover, by definition of Λ and M, we also have

$$UQ = Q \tag{12}$$

and

$$WR = R\,. \tag{13}$$

We wish to show that $T^lW^{-1}UT^{l-2k}P = P$.

Eliminating P from (10) and (11) we have

$$R = T^{k-l}Q\,. \tag{14}$$

Because of (12) and (13) we are permitted to replace in (14) Q by UQ and R by WR. Thus we have $WR = T^{k-l}UQ$. From (9) we get $WR = UT^{l-k}Q$. Remembering from (10) and (11) that $Q = T^{-k}P$ and $R = T^{-l}P$, we have $WT^{-l}P = UT^{l-2k}P$. Taking the image of this point under T^lW^{-1}, we obtain the desired result.

In order to obtain a more symmetrical theorem it is convenient to

note the following trivial

LEMMA. *Assuming that T and U satisfy* (8) *a necessary and sufficient condition that*

$$T^{-1} = WTW^{-1} \tag{15}$$

is that $W^{-1}U$ commute with T.

Proof. From (8) we see that (15) is valid if and only if $UTU^{-1} = WTW^{-1}$, which is equivalent to $(W^{-1}U)T = T(W^{-1}U)$, as stated in the Lemma.

THEOREM 6. *Let T, U, W satisfy* (8) *and* (15) *and let Λ and M be defined as in Theorem* 5. *Then any point common to $T^k\Lambda$ and T^lM is invariant under $W^{-1}UT^{2(l-k)}$ or, what is the same thing, $T^{2(l-k)}W^{-1}U$.*

Theorem 6, which is a trivial consequence of Theorem 5 and the lemma, is slightly less general than Theorem 5, but it has the advantage of remaining true when the elements of the pairs (U, W) and (l, k) are interchanged.

By taking $W = UT^m$, where m is any integer, we find that $U^{-1}W = T^m$, which certainly commutes with T, so that, by the lemma, (15) is valid. We many thus specialize Theorem 6 as follows:

THEOREM 7. *Suppose that T and U are any one-to-one transformations of a set S onto itself and that T and U are such that $T^{-1} = UTU^{-1}$ so that T is reversible in any group of transformations containing both T and U. Let Λ be the set of points, each of which is invariant under U and let M be the set of points, each of which is invariant under UT^m. Then any point common to $T^k\Lambda$ and T^lM is invariant under $T^{2(l-k)-m}$.*

This theorem, while more special than Theorem 5, is still more general than Theorem 4, as we see by considering the cases $m = 0$ and 1.

We have, of course, never assumed U to be involutory in Theorems 5, 6, or 7. Nevertheless it will appear in the sequel that both U and $V = UT$ must be involutory in a significantly important subset $\sum$ of S, if the intersection of $T^k\Lambda$ and T^lM is not empty.

THEOREM 8. *Suppose that the one-to-one transformations T and U of the set S onto itself are such that* (8) *hold. Let $\sum$ be the set of points, each of which is invariant under U^2. Then $\sum$ is invariant*

under T and U, and the set Λ defined in Theorem 5 is a subset of $\sum$.

Proof. As in the proof of Theorem 5, we first establish (9) in certain special cases so that we have both

$$T^{-1}U = UT \text{ and } TU = UT^{-1}. \tag{16}$$

If $P \in \sum$ we have (by definition of $\sum$) $U^2P = P$. Therefore $TU^2P = TP$. But, with the help of (16) we may write $TU^2P = (TU)UP = (UT^{-1})UP = U(T^{-1}U)P = U(UT)P = U^2TP$. Hence, we have $U^2TP = TP$. This shows that TP is invariant under U^2, and hence (by definition of $\sum$) $TP \in \sum$. Hence $\sum$ is invariant under T.

That $\sum$ is also invariant under U is even more obvious. For, from $U^2P = P$, we have immediately $U^3P = UP$ or $U^2(UP) = UP$.

Finally $\Lambda \subset \sum$ because, if $P \in \Lambda$, then $UP = P$ and hence $U^2P = UP = P$ so that also $P \in \sum$.

In non-vacuous applications of Theorem 5, the set $\sum$ is not empty, because then $T^k\Lambda$ has a point P in common with another set, namely, T^lM, and this implies the existence of a point $Q \in \Lambda \subset \sum$ such that (10) holds. By Theorem 8, all the sets $T\Lambda$, $T^2\Lambda$, $T^3\Lambda$, $\cdots$ are also subsets of $\sum$. Since $(T^k\Lambda) \cap (T^lM) \neq 0$, $\sum \cap (T^lM)$ is not empty and hence also $\sum \cap (T^{l-1}M)$, $\sum \cap (T^{l-2}M)$, $\cdots$, $\sum \cap M$ are not empty. It follows that in proving Theorem 5 we can restrict our transformations to the set $\sum$ on which the property $U = U^{-1}$ is valid, and where accordingly $T = UV$ with $V = V^{-1}$. When so restricted the original proofs of Birkhoff may be adapted to our more general situation. But such an alternative proof of Theorem 5 would not be simpler than the one actually offered.

As a final remark, it may be said that Theorem 5, general though it already is, can be made still more general. In fact, in following the proof of this theorem up to (14), we notice that we are permitted (thanks to (12) and (13)) to replace in (14) Q by $U^\alpha Q$ and R by $W^\beta R$, where α and β are arbitrary integers. This would lead to a more general theorem reducing to the stated theorem when $\alpha = \beta = 1$. I have not carried out a formulation of such more general theorems because of a doubt that they could be at all significant.

References

1. G. D. Birkhoff, *The restricted problem of three bodies*, Rendiconti del Circolo Matematico di Palermo, **39** (1914), 1-70, especially pp. 46-48.

2. _____, *Dynamical systems*, Amer. Math. Soc. Colloq. Publ, **9** (1927), viii + 295 pp., especially pp. 180-188.

3. _____, *A new criterion of stability*, Atti del Congresso Intern. dei Matem. Bollogna,

3-10 settembre 1928-vi, 1931, vol. **5**, pp. 5-13.
4. _____, *Sur le problème restreint des trois corps* (premier memoire), Annali della R. Scuola Normale Superiore di Pisa. s.2, **4** (1935), 267-306, especially pp. 290-300.
5. _____, *Sur le problème restreint dds trois corps* (second mémoire), ibid., **5** (1936), 1-42, especially pp. 1-25.
6. _____ and Jaime Lifshitz, *Ciertas transformaciones en la dinamica sin elementos periodicos*, Publicacione del Instituto de mathematica, Rosario, **6** (1945), 1-14, especially pp. 8,9.

ALL EXCEPT THE SECOND REFERENCE CAN BE FOUND IN VOLUMES I AND II OF BIRKHOFF'S COLLECTED MATHEMATICAL PAPERS.