JOURNAL

DE

MATHÉMATIQUES

PURES ET APPLIQUÉES

FONDÉ EN 1836 ET PUBLIÉ JUSQU'EN 1874

PAR JOSEPH LIOUVILLE

GEORGE D. BIRKHOFF

PAUL A. SMITH

Structure Analysis of Surface Transformations

Journal de mathématiques pures et appliquées 9e série, tome 7 (1928), p. 345-379.

<http://www.numdam.org/item?id=JMPA_1928_9_7__345_0>

Structure Analysis of Surface Transformations;

By George D. BIRKHOFF
(Cambridge, U. S. A.),
and Paul A. SMITH
(New York, U. S. A.).

Introduction.

It is intended to set forth in this paper certain general facts concerning the structure of one-to-one continuous transformations of surfaces into themselves, especially as regards the movement of points under indefinite iteration. For the most part, the transformations with which we shall deal are non-analytic. The restriction of analyticity, however, reduces the possible structural complexity to such an extent that something like a systematic structure analysis in the general analytic case can conceivably be developed. At the end of the paper we shall indicate in the briefest way some of possibilities with regard to a systematic study of this sort, altho the paper as a whole may be regarded as preliminary to such a study.

Surface transformations which are associated with certain types of dynamical problems have the property that they admit an invariant area integral. These « conservative » transformations, which have been studied by Poincaré (II)[1] and more extensively by Birkhoff (IV), possess a fundamental property of regional recurrence. In general,

[1] The roman numerals refer to the list of references found at the end of this paper.

however, the phenomenon of recurrence does not extend to the entire surface, but does nevertheless take place within certain invariant subsets. We shall undertake a study of these subsets, — the precise nature of the recurrence which takes place within them, the extent to which they may be considered as transforming conservatively, the general movement of the remaining points of the surface and questions of uniform approach. Finally, we shall apply our general principles in a brief examination of the structure of simple types of analytic transformations.

1. *Preliminary definitions. The general analytic case.* — Throughout this paper, the term « surface » (denoted consistently by S) will mean a closed orientable surface of arbitrary genus and the term « transformation » (denoted by T) will mean a one-to-one continuous sense-preserving transformation of such a surface into itself.

We shall designate by T_2, T_3, ... the successive powers of T, and by T_{-2}, T_{-3}, ... those of the inverse T_{-1}. An infinite sequence of points of the form

$$\ldots,\quad P_{-2},\quad P_{-1},\quad P,\quad P_1,\quad P_2,\quad \ldots$$

where P is an arbitrary point of S, and $P_n = T_n(P)$, will be called a *complete sequence.*

If two points P_α and $P_\beta (\alpha < \beta)$ of a complete sequence coincide, then so do $P_{\beta-\alpha}$ and P. Let k be the smallest positive integer for which $T_k(P) = P_k = P$. Then P will iterate periodically thru a set of k distinct points, and is therefore called a *periodic* point of the order k.

Any limit point of the infinite sequence P, P_1, P_2, — is called an *ω-limit point* of P and any limit point of the sequence P, P_{-1}, P_{-2}, ... is an *α-limit point.* In case P is periodic, each of its images is to be considered an α-and an ω-limit point, and there are no others.

A complete sequence together with its α-and ω-limit points will be called a *complete group.*

A set of points E is invariant under T if E and T(E) are identical point sets. A complete group, for example, is a closed invariant set. If a point P is contained in an invariant set, so is the complete sequence of P. An invariant set, however, need contain no invariant or periodic points. In case E is invariant under T_k, but not under T, $T_1, \ldots, T_{k-1}$, it is periodic, and of the order k.

Suppose that a simply connected region G on S has the property that it contains the transformed region T(G) together with the boundary of T(G). Then G, together with its boundary is contained in $T_{-1}(G)$. We shall call G a *contracting region* under T, or an *expanding region* under T_{-1}.

We shall now summarize certain facts and definitions concerning *analytic* surface transformations. In the neighborhood of an invariant point O, T may be represented in terms of a properly chosen coordinate system as follows:

$$\begin{aligned} u_1 &= au + bv + \ldots \\ v_1 &= cu + dv + \ldots \end{aligned} \qquad (ad - bc > 0),$$

where the right hand members are convergent power series with real coefficients, and u_1, v_1, are the coordinates of the transformed point.

Let μ and ν be the roots of the characteristic equation

$$\lambda^2 - (a + d)\lambda + ad - bc = 0.$$

We shall say that O is of *general type* provided that μ and ν are distinct and of modulus different from 1. In any other case O is of *special type*. If μ and ν are real, and $0 < \mu < 1 < \nu$, O is called a *directly unstable point;* if $\mu < -1 < \nu < 0$, O is *inversely unstable*. All other invariant points of general type are called *stable*. An invariant point which is inversely unstable under T, is directly unstable under T_2, while stable and directly unstable points retain their type under any power of T.

The definitions above are independent of the coordinate system. They extend, moreover, to periodic points of any order; for example, if O is of order k, we take the roots μ and ν relative to T_k.

Suppose O is an invariant point of stable type. Then points in the neighborhood of O converge toward O on indefinite iteration of T (or of T_{-1}) in such a way that the region bounded by any sufficiently small circle about O is contracting under T (or T_{-1}) (See reference to Lattés).

Suppose next that O is of directly unstable type. There abut at O four invariant curves or branches [1], — two *α-branches* whose points converge toward O on indefinite iteration of T_{-1}, and two *ω-branches*, whose points converge toward O on iteration of T. A point contained in a small neighborhood σ of O, but not on one of the invariant branches, is carried out of σ on repeated iteration of either T or T_{-1}.

The two α-branches are analytic continuations of each other at O, and taken together, form an invariant analytic curve *without singularities*, which at O crosses the corresponding curve formed by the two ω-branches. If two sufficiently small arcs a_1 and a_2 of the two α-branches respectively, abut at O, they will not intersect. It follows that the two α-branches can *nowhere* intersect each other; for if P were a point of intersection, an image of P under a sufficiently great power of T_{-1} would be in both a_1 and a_2 which is impossible. Moreover, an α-branch of O can not intersect an α-branch of some other point, say Q, for a point of intersection would be carried simultaneously toward O and Q on indefinite iteration of T_{-1}, which is impossible. The same holds for the ω-branches.

An inversely unstable invariant point has α-branches of order 2, — they are invariant under T_2 but not T. An unstable periodic point O of order k has α-and ω-branches of order k and $2k$, according as O is directly or inversely unstable. If O is of the former type, a point on any of its branches tends asymptotically toward the complete group of O on repeated iteration of T (or T_{-1}), and toward O itself, on iteration of T_k (or T_{-k}).

If we consider the totality of α-and ω-branches of all orders on S, it is clear from the discussion above, that no α-(or ω-) branch can intersect another α-(or ω-) branch. However, an α-branch may intersect an ω-branch, and the points of intersection in such a case are called *doubly asymptotic* (d. a.) points (Poincaré, II). If the two branches which intersect at a d. a. point P actually cross, i. e., are not coincident or merely tangent at P, then P will be said to be of *general type*; in the contrary case, P is of *special type*.

---

(1) First proved by Poincaré, (I). With regard to these invariant branches, see also Poincaré (II) and Lattès.

An analytic surface transformation which admits no d. a. or periodic points of special type will be said to belong to the *general analytic case*. The sense in which such transformations may be considered « general » is indicated by the fact that a d. a. point of special type can be converted into a number of d. a. points of general type by an arbitrarily slight modification of T, but not conversely; the same holds for the periodic points. We shall not, however, here consider the situation in further detail.

A d. a. point is called *homoclinic* (Poincaré, II) if the α-and ω-branches on which it lies issue from the same unstable periodic point or from two points belonging to one and the same periodic group. For convenience, let us say that homoclinic points of the former type are *simple*.

We shall have occasion to refer later to the following theorem :

In the general analytic case, an arbitrarily small neighborhood of a homoclinic point contains infinitely many periodic points.

A proof of this theorem is given by Birkhoff (V) for the case of simple homoclinic points [1]. The remaining cases are disposed of by the following lemma :

In the general analytic case, an arbitrarily small neighborhood of a homoclinic point contains a homoclinic point of simple type.

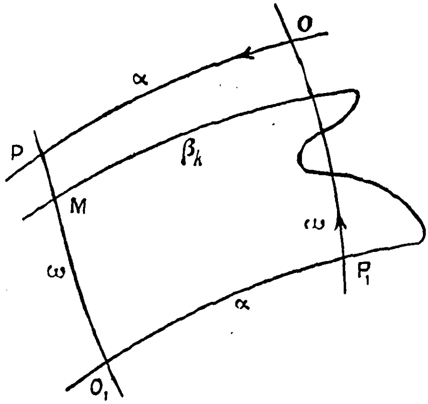
We shall briefly indicate the proof. Suppose for concreteness that P is a point of intersection of an α-branch of O with an ω-branch of $O_1 = T(O)$ where O is a directly unstable point of order 2. Thus P is a non-simple homoclinic point. The transform of the α-branch OP is the α-branch $O_1\ P_1$ and the transform of the ω-branch O_1 P is the ω-branch OP_1. (See *fig.* 1).

Now O is invariant under T_2, and P_1, being on an ω-branch of O is carried toward O on repeated iteration of T_1. Moreover, a point sufficiently close to O and on the proper side of the curve OP_1. is carried toward and beyond P on iteration of T_2, the successive images

(1) The proof given assumes the existence of an invariant integral, but can be extended to the general case: details for this will appear elsewhere.

remaining close to OP. Hence a small arc β crossing OP_1 at P_1, will eventually be carried into an arc β_k which follows along close to OP sufficiently far that it will cross PO_1 near P, say at M. Since

Fig. 1.



the d. a. point P_1 is of general type, the α-branch $O_1 P_1$ actually crosses OP_1 at P_1 and hence β may be taken as an arc of $O_1 P_1$. Hence β_k is also an arc of the α-branch $O_1 P_1$ and M is therefore a homoclinic point of simple type. This establishes the lemma for the case considered, and there is no difficulty in making the proof general.

2. *The central motions* (1).

Consider an arbitrary connected region σ on S. It may happen that σ is intersected by none of its images

$$\ldots, \quad \sigma_{-2}, \quad \sigma_{-1}, \quad \sigma_1, \quad \sigma_2, \quad \ldots$$

in which case, σ is called a *wandering region* and its points *wandering points*. A point of S which is contained in no wandering region is a *non-wandering* point.

No two images of a wandering region can intersect. For if σ_i and $\sigma_j (i < j)$ intersect, then so do σ and σ_{j-i} which is impossible. Consequently any image of a wandering region or point is again wandering, and any image of a non-wandering point is non-wandering.

(1) *Cf.* Birkhoff, IV.

Among the non-wandering points are the α-and ω-limit points. Suppose for example that L is an ω-limit point and let σ be an arbitrarily small neighborhood of L. Within σ there are infinitely many points of some sequence $P, P_1, P_2, \ldots$ If P_α and $P_\beta (\alpha < \beta)$ are two of these points in σ, the regions $\sigma_{\beta-\alpha}$ and σ obviously overlap, and hence σ can not be wandering.

The totality of non-wandering points of S *constitutes a non-null closed invariant set* M' *towards which all other points tend asymptotically on indefinite iteration of* T *or* T_{-1}.

In the first place, non-wandering points must exist on S, for there are always α-and ω-limit points on a closed surface. The set $S - M'$ consists only of inner points, and hence M' is closed. Since all images of a non-wandering point are non-wandering, M' is invariant. Finally, if P is a point of $S - M'$, the sequence $P, P_1, \ldots$ tends asymptotically toward M'. For otherwise a number $\delta > 0$ exists and an infinite subsequence of points $P_\alpha, P_\beta, \ldots$ each of which is at a distance greater than δ from M'. No limit point of the subsequence can belong to M'; but on the other hand, every such limit point is an ω-limit point and therefore non-wandering. This contradiction proves our assertion. If follows similarly that the sequence $P, P_{-1}, \ldots$ tends asymptotically toward M'.

The following theorem concerns the movement of $W' = S - M'$ as a whole, on indefinite iteration.

THEOREM 1. — *Not more than k points of a complete sequence of wandering points can be outside a given neighborhood* V *of* M', *where k depends only on the choice of* V.

Proof. Suppose the theorem false. Then there exist complete sequences which have more than N points in $W' - T$, where N is arbitrary. Thus, for every positive integer n, there is a set E^n consisting of at least n points taken from a complete sequence, and all contained in $W' - V$. We shall pick from each E^n a pair of points P^n and Q^n chosen such that the distance $P^n Q^n$ shall converge to zero with $1/n$. This is possible since the number of points in E^n grows indefinitely with n.

Now the set $W^1 - V$ is closed and any limit point L of the sequence $P^1, P^2, \ldots$, must therefore be wandering. But on the other hand a neighborhood of L, however small, will contain a pair P^n Q^n; and since Q^n belongs to the same sequence as P^n, a certain power of T or T_{-1} will carry P^n into Q^n, and hence σ into a region that overlaps σ. Therefore L is non-wandering, which is a contradiction.

It may of course happen that the set M^1 is identical with S. This is the case, for example, when T possesses an invariant integral of a certain type, as we shall see later.

Let us suppose now that M^1 is not identical with S, and let us take the set M^1 as fundamental instead of S. A connected region which contains points of M^1 will be called wandering *with respect to* M^1 if the set σM^1 of points common to σ and M^1 is intersected by none of its images under powers of T or T_{-1}. The points of M^1 which are contained in such a region are called wandering *with respect to* M^1, and their totality will be denoted by W^2. The set $M^2 = M^1 - W^2$ consists of the points which are non-wandering *with respect to* M^1. In case $M^1 = M^2$, we shall say that M^1 is non-wandering with respect to itself.

The complete analogy which exists between S, M^1, W^1, and M^1, M^2, W^2, will be seen immédiately; M^2 is a non-null invariant closed subset of M^1, and toward M^2 the points of W^2 tend asymptotically on indefinite iteration of T or T_{-1}.

In case M^2 is not identical with M^1, the process may be carried one step farther, yielding the set M^3 of points which are non-wandering with respect to M^2. We continue thus until we arrive at the set M^k which is non-wandering with respect to itself. In case, however, that no such set appears after a finite number of steps, we shall have an infinite sequence M^1, M^2, ... with $M^1 > M^2 > \ldots$ The set $M^\omega = M_1 M_2 \ldots$, is closed and not null, and our process applied to M^ω yields $M^{\omega+1}$, then $M^{\omega+2}$ and so on.

In this manner we obtain an ordered aggregate of point sets

$$M^1, \ \ldots, \ M^\omega; \ M^{\omega+1}, \ \ldots, \ M^{2\omega}, \ \ldots, \ M^{\omega^2}, \ \ldots.$$

Each set is a proper subset of all those preceeding it. Such an aggregate can be at most denumerable, and hence, when arranged as above in a well-ordered sequence, is associated with a definite ordinal r of Cantor's second ordinal class. Thus the sequence above terminates with M^r which therefore must be non-wandering with respect to itself. The points of M^r will be called *central points*, and a complete sequence of central points will be called a *central motion*.

Among the closed invariant point sets which are non-wandering with respect to themselves, the set M^r is *maximal* in the sense that every such set is contained in M^r. For if E is such a set, we have successively $E \leqq M^1$, $E \leqq M^2$, ..., $E \leqq M^\omega$, ..., and hence $E \leqq M^r$. Moreover any closed invariant set on S may be taken as the initial set in the above process, and hence must contain a subset which is non-wandering with respect to itself. It follows that *every closed invariant set contains at least one central motion*. This applies in particular to a complete group.

A study of the structure of M^r and the sequence M^1, M^2, ..., which determines M^r will occupy much of our attention in what follows. We shall first prove a fundamental recurrence property of M^r.

A point which is both an α-and ω-limit point of its own complete sequence will be called *pseudo-recurrent* (¹). The characteristic property of such a point is that it returns infinitely often into an arbitrarily small neighborhood of itself under indefinite iteration of T as well as T_{-1}. All images of a pseudo-recurrent point are pseudo-recurrent. Moreover, a pseudo-recurrent point is a central point, for its complete group is obviously non-wandering with respect to itself and is therefore contained in M^r.

The fundamental recurrence property of M^r may now be stated as follows :

THEOREM 2. — *The set E which consists of the pseudo-recurrent points together with the limit points of pseudo-recurrent points is identical with* M^r.

Proof. First, since pseudo-recurrent points are central motions and since M^r is closed, we have $E \leqq M^r$.

(¹) The term « recurrent » has been used elsewhere (Birkhoff, III) with a slightly different meaning. Recurrent points are pseudo-recurrent, but not conversely.

It remains to prove that $M^r \leqq E$. We shall show that an arbitrarily small neighborhood σ of a central point A contains at least one pseudo-recurrent point, so that A is either pseudo-recurrent, or a limit point of pseudo-recurrent points, and is therefore in E.

The set of central points contained in any small region, — for example σ, must intersect images of itself under powers of T and T_{-1}, for there are no wandering regions with respect to M^r. Hence there exist in σ a pair of central points P and Q which are images, one of the other, under some power of T. We shall assume that P preceeds Q ([1]).

Next choose about P a neighborhood p so small that both p and q, the corresponding neighborhood of Q, shall be contained in σ. There exists in p a pair P^1 and Q^1 of central points, images one of the other, under some power of T. We shall suppose this time that P^1 is preceeded by Q^1.

We shall describe one more step in detail. A neighborhood p^1 of P^1 is chosen so small that both p^1 and q^1 shall be contained in p. In p^1 are P^2 and Q^2, images one of the other under some power of T, and named so that P^2 preceeds Q^2. The important point in the choice of successive pairs P^i and Q^i is to name them in such a way that P^{2n} preceeds Q^{2n}, while P^{2n+1} is preceeded by Q^{2n+1}.

In continuing thus, we choose the successive neighborhoods $p, p^1, \ldots,$ in such a way that the diameter of p^i shall converge to zero as $i \to \infty$. By the manner in which these neighborhoods are defined, we have

$$\sigma > p > p^1 > p^2 > \ldots, \tag{1}$$

$$\sigma > q,\ p > q^1, \ldots, p^\alpha > q^{\alpha+1}, \ldots \tag{2}$$

Now there must exist at least one point L with the property that it lies in or on the boundary of each neighborhood of the sequence (1). We shall show that L is pseudo-recurrent, and thus establish our theorem.

We must show, then, that given an arbitrarily small neighborhood λ

([1]) i. e when the complete sequence to which P and Q belong is written according to increasing powers of T. In case P coincides with Q, we shall say that P precedes and is preceded by Q.

of L, there are in λ images of L under powers of T as well as T_{-1}. Let α be a positive integer chosen so large that p^{α} together with its boundary shall be contained in λ. Then by (1) and (2), the regions $p^{\alpha+1}$, $q^{\alpha+1}$, $p^{\alpha+2}$, $q^{\alpha+2}$, are all contained in p^{α} and hence in λ.

Suppose that α is even. Then $q^{\alpha+1}$ preceeds $p^{\alpha+1}$ and $p^{\alpha+2}$ preceeds $q^{\alpha+2}$. Therefore, since L is in or on the boundary of both $p^{\alpha+1}$ and $p^{\alpha+2}$, the image of L under some power of T_{-1} is in or on the boundary of $q^{\alpha+1}$, hence in λ; and the image of L under some power of T is in or on the boundary of $q^{\alpha+2}$, hence in λ. The situation is reversed if α is odd. This completes the proof.

We shall now prove a theorem concerning the distribution of pseudo-recurrent points in the following important special case: T is analytic and possesses an invariant integral defined over a closed invariant set E, where E is measurable in the sense of Lebesgue and of non-zero measure. We assume specifically that the integral is of the form

$$\int_{e} \varphi(P)\, d\sigma, \quad 0 < \alpha < \varphi(P) < \beta \qquad \text{on E, (β finite).}$$

the function φ being defined and measurable on E.

The set of pseudo-recurrent points contained in E *is measurable, and its measure is equal to* m (E).

Proof. Suppose e is a measurable subset of E, with $m(e) > 0$. Then e must intersect images of itself under powers of T and T_{-1}. For if the sets e, e_1, ..., were mutally exclusive, the sum

$$\sum_{i} \int_{e_i} \varphi(P)\, d\sigma$$

could not be finite, since $\int_{e} = \int_{e_1} = \int_{e_2} = \ldots$ But in contradiction to this we have

$$\sum_{i} \int_{e_i} \varphi(P)\, d\sigma \leqq \int_{E} \varphi(P)\, d\sigma \leqq \beta m(E).$$

Now let ε_1, ε_2, ... be a sequence of positive numbers converging to zero and let H^i consist of those points of E which never come within a distance ε^i from their initial positions on iteration of T. We shall show that $m(\mathrm{H}^i) = 0$ for every i. For suppose that the outer measure of H^1, say, is > 0. Then H^1 possesses a measurable subset $\overline{\mathrm{H}}$ of non-zero measure ([1]).

Now, we can obviously choose a simply connected region σ of diameter smaller than ε_1 for which $m(\sigma\overline{\mathrm{H}}) > 0$. By the remark above, $\sigma\overline{\mathrm{H}}$ must intersect an image of itself under some power of T, and this same power of T therefore carries some point of $\sigma\overline{\mathrm{H}}$ back into σ. This is impossible by definition of H^1. Hence $m(\overline{\mathrm{H}}) = m(\mathrm{H}^1) = 0$.

It follows that $m(\mathrm{K}) = m(\mathrm{E})$, where $\mathrm{K} = \mathrm{E} - (\mathrm{H}^1 + \mathrm{H}^2 + \ldots)$. Each point of K is clearly an ω-limit point of its own complete sequence. By entirely similar reasoning we arrive with a set L with $m(\mathrm{L}) = m(\mathrm{E})$, each of whose points is an α-limit point of its own complete sequence.

Since the sets K and L are both contained in E, and are in measure equal to $m(\mathrm{E})$, they must overlap to the extent that $m(\mathrm{KL}) = m(\mathrm{E})$. The points of KL are of course pseudo-recurrent, which establishes the theorem.

For a conservative transformation, E is identical with S, and therefore the measure of the pseudo-recurrent points equals the total surface area of S. This is closely related to the statement of Poincaré (II) that in certain dynamical problems, there exists stability in the sense of Poisson, except for « motions of zero probability ».

Theorem. — *At least one point of every set of k successive points of a complete sequence falls in a given neighborhood* V *of the set* M^r. *The value of k depends only on* V.

Proof. At least one point of every complete group must fall in V, since a complete group contains at least one central motion. Now if the theorem were false, there would be sequences of the form P, $\mathrm{P}_1, \ldots, \mathrm{P}_\mathrm{N}$, N being arbitrarily large, which have no points in V. Let

$$\begin{array}{llll} \mathrm{P}, & \mathrm{P}_1, & \ldots. & \mathrm{P}_{m_1}, \\ \mathrm{Q}, & \mathrm{Q}_1, & \ldots, & \mathrm{Q}_{m_2}, \\ \ldots & \ldots & \ldots & \ldots \end{array}$$

([1]) See Carathéodory, *Vorlesungen über reelle Funktionen.*

be an infinite succession of such sequences, with $m_1 < m_2 < \ldots$. Then clearly the complete group of any limit point of the set P, Q, ..., must be entirely outside of V, which is impossible.

3. *The sequence* $N^1, N^2, \ldots$ — We shall now define the set M^r of central motions by means of a new fundamental sequence, and in so doing, we shall reveal certain additional properties of M^r.

We shall employ a modification of our earlier process, which consists in considering those non-wandering points which are α- or ω- limit points. If to the set of α-and ω-limit points of S, we add the ordinary limit points of such points, we obtain a closed invariant set N^1. This set is of course contained in M^1, altho the two sets may be identical. (We shall consider later the conditions under which this must happen.)

By the same argument which we used for M^1, it follows that the points of $S - N^1$ tend asymptotically toward N^1 on indefinite iteration of T or T_{-1}; we do not, however, have a theorem analogous to Theorem 1, § 2 for this case.

Let us now take the set N^1 as fundamental.* An α- limit point *with respect to* N^1 is a limit point of some sequence $P, P_{-1}, \ldots$, *contained in* N^1. If to the α- and ω- limit points with respect to N^1, we add their limit points (in the ordinary sense) we obtain a closed invariant set N^2 contained in N^1; it is easily verified that N^2 is contained also in M^2.

In the light of the preceeding section, the manner of proceedure is clear. We arrive eventually with a closed invariant set N^s, with which the process terminates, — i. e. such that $N^{s+1} = N^s$.

THEOREM. — *The sets* N^s *and* M^r *are identical.*

Proof. Since N^s consists of α- and ω-limit points with respect to itself, and the ordinary limit points of such points, an arbitrarily small region σ which contains points of N^s, contains at least one limit point of some complete sequence contained in N^s. Hence σN^s must intersect images of itself under powers of T or T_{-1}. Thus N is non-wandering with respect to itself and therefore $N^s \leqq M^r$ (§ 2).

Next, it is clear from their definition that all pseudo-recurrent points belong to N^s. Since N^s is closed, limit points of pseudo-recurrent points are also contained in N^s. Therefore, by Theorem 2, § 2, $M^r \leq N^s$. Hence $M^r = N^s$.

Between the ordinals r and s, we have the relation $s \leq r$; it is probable that s may be actually less than r in certain cases.

We shall now prove a simple lemma preliminary to obtaining a further property of M^r:

LEMMA. — If Q is a limit point of a complete sequence Σ relative to T, it is also a limit point of a complete sequence Σ' relative to T_k, k being any integer, and Σ' being a subsequence of Σ.

Let Σ be the sequence $\ldots, P_{-2}, P_{-1}, P, P_1, \ldots$, and from it let us extract a subsequence $P_\alpha, P_\beta, \ldots$, which converges to Q. The subsequences

$$\ldots,\ P_{-2k+i},\ P_{-k+i},\ P_i,\ P_{k+i},\ P_{2k+i},\ \ldots \qquad (i = 0, 1, \ldots, k-1)$$

constitute a set of k complete sequences of T_k, and each is a subsequence of Σ. Taken together, these sequences contain all the points of Σ, and hence at least one of them contains infinitely many points of the sequence $P_\alpha, P_\beta, \ldots$ and so has Q for a limit point.

THEOREM 3. — *The sets $N^1, N^2, \ldots$, relative to T are identical respectively with the sets $\overline{N}^1, \overline{N}^2, \ldots$, relative to T_k. Hence the set of central points relative to T is identical to the set relative to T_k.*

Proof. If Q is an α- or ω- limit point of T_k, it is of course an α- or ω- limit point of T. By virtue of the lemma, the converse is also true. Hence $N^1 = \overline{N}^1$. Next, if Q is an α- or ω- limit point of T_k with respect to N^1, it is an α- or ω- limit point of T with respect to N^1. Again, the converse of this statement follows from the lemma, and we have $N^2 = \overline{N}^2$. For suppose Q to be a limit point of a complete sequence Σ of T, contained in N^1. Then by the lemma Q is also a limit point of a complete sequence Σ' of T_k, where Σ' is contained in Σ and hence in N^1. Thus Q is an α- or ω- limit point of T_k with respect to N^1, as stated. Proceeding in an entirely similar manner for N^2, $\overline{N}^2, \ldots$, the proof of our theorem is established.

It is questionable whether or not the theorem holds also for the sets M^1, M^2, ...

Before considering under what circumstances the sets M^1 and N^1 are identical, we shall introduce a new definition. Let σ be a connected region contained in the open set $S - N^1$. As we have seen, the points of σ tend asymptotically toward N^1 on indefinite iteration. Now it may happen that their totality tends toward N^1 *uniformly* on iteration of T (or T_{-1}). By this we mean that given arbitrary positive ε, there exists a positive integer K such that each point of each σ_k (or σ_{-k}), $k > K$, is within a distance ε from the closed set N^1. In such a case we shall call σ an ω- (or α-) *regular region*, and its points ω- (or α-) *regular points*. Points of $S - N^1$ which are contained in no such region will be called ω- (or α-) *irregular*.

THEOREM. — *Points of* $M^1 - N^1$, *if any exist, are* α- *and* ω-*irregular.*

Proof. Let P be such a point. An arbitrarily small neighborhood σ of P is intersected by at least one of its images under powers of T. Consequently σ contains a point Q, an image of which, say Q_k, $k > 0$, is also in σ. Let us choose a point pair such as QQ_k for each one of a sequence of regions $\sigma^1, \sigma^2, \ldots$, closing down on P. Let these pairs be

$$(\Sigma) \qquad Q^1, \quad Q^1_{m_1}, \quad Q^2, \quad Q^2_{m_2}, \quad \ldots$$

There can not be a finite upper bound for the sequence of positive integers $m_1, m_2, \ldots$. For if N were such a bound, infinitely many of the integers of the sequence are equal to some integer m, $0 < m \leqq N$, and from Σ we could extract a sequence of the form

$$Q^{\alpha_1}, \quad Q^{\alpha_1}_m, \quad Q^{\alpha_2}, \quad Q^{\alpha_2}_m, \quad \ldots$$

But since the sequences $Q^{\alpha_1}, Q^{\alpha_2}, \ldots$, and $Q^{\alpha_1}_m, Q^{\alpha_2}_m, \ldots$, both converge to P, it is clear that P_m coincides with P. Thus P is periodic, and hence belongs to N^1, contrary to hypothesis.

It follows that there can be extracted from Σ a sequence of the form

$$Q^{a_1}, \quad Q^{a_1}_{b_1}, \quad Q^{a_2}, \quad Q^{a_2}_{b_2}, \quad \ldots$$

with $0 < b_1 < b_2 < \ldots$. Each region σ^i contains all the points of this sequence from a certain rank on and hence it is clear that no σ^i could possibly tend uniformly toward N^1 on iteration to T. Hence P is ω- irregular. By an entirely similar argument, P is shown to be α- irregular. This completes the proof.

We shall have occasion later to refer to the following theorem.

THEOREM 4. —*If a connected region p contains non-wandering points, it is intersected by infinitely many of its images under powers of* T *as well as* T_{-1}.

Proof. Let P be a non-wandering point in p. Referring to the proof of the preceding theorem, let Q^1, $Q^1_{m_1}$, Q^2, $Q^2_{m_2}$, ..., be a sequence of the type (Σ) and converging to P. As we have shown, either there exists no finite upper bound for the sequence m_1, m_2, ..., or else P is periodic. Either situation leads to the stated conclusion.

4. *Invariant integrals.* — We have seen that an invariant set E over which there can be defined an invariant integral of a certain type must necessarily consist of central motions. It is probably not true, however, that conversely, an invariant integral may always be defined over an invariant subset of M^c. Suppose, for example, that M^c is identical with S. Then if σ is any connected region on S, the regions σ, σ_1, σ_2, ..., can not be mutually exclusive. But there is no apparent reason why the regions of some infinite subsequence σ_α, σ_β, ..., should not be mutually exclusive, — a situation which could not arise if T were conservative. Indeed, there can not be any purely topological condition for a metrical phenomenon such as conservatism. It will be worth while, however, to examine any available condition which will shed light on the structure of T.

Consider the region σ. In general there will be some image of σ whose area is smaller than that of σ. Hence on dividing S into a number of regions and choosing the proper image of each, S becomes compressed, in a sense, into an area smaller than its total surface area. We shall show that a necessary and sufficient condition that there exist invariant integrals of a certain type on S, or part of S, in that S *be not compressible into an arbitrarily small area.* This is an intuitive statement of the results of this section.

We shall assume now that T is analytic, and shall begin by introducing of function $\varphi(e)$, e being any measurable set on S, defined as follows : let e be divided into a finite number of mutually exclusive measurable sets δ^i,

$$e = \sum_i \delta^i, \quad \delta^i \delta^j = 0 \quad \text{for} \quad i \neq j.$$

Then $\varphi(e)$ is the lower bound of the sum

$$\sum_i m(\delta^i_{n_i})$$

with respect to all possible methods of subdivision of e into finite numbers of measurable sets, and all possible choices of the integers n_i. Here we make use of the fact that the property of measurability is preserved under analytic transformations.

It is clear that function φ may be identically zero, in which case, S is « compressible into an arbitrarily small area ». This happens, for example, when T is an analytic transformation of a sphere such that each circle parallel to the equator closes down on the north (or south) pole on indefinite iteration of T (or T_{-1}). On the other hand, for a transformation which preserves areas, we have $\varphi(e) = m(e)$.

In any case, it follows immediately from the definition, that

$$\varphi(e) \leqq m(e)$$

and hence $\varphi(e)$ is bounded, and totally continuous on S.

The importance of φ for our purposes is due to the following theorem :

$\varphi(e)$ is a completely additive function of measurable sets and is invariant under T.

Proof. We first prove the invariance of φ. Let e be an arbitrary measurable set and suppose that $\varphi(e) < \varphi(e_1)$. By a proper subdivision of e into a finite number of measurable sets, together with a proper choice of the corresponding integers n_i we obtain a sum

$$\sum_i m(\delta^i_{n_i})$$

which approximates $\varphi(e)$ to any desired degree of closeness. In particular, we may assume by virtue of the assumption $\varphi(e) < \varphi(e_1)$ that

$$\sum_i m(\delta'_{n_i}) < \varphi(e_1).$$

Now $\sum_i m(\delta'_1)_{n_i-1}$ is an approximating sum for $\varphi(e_1)$. Since its value is $\sum_i m(\delta^i_{n_i})$, the inequality above contradicts the fact that $\varphi(e_1)$ is the lower bound of its approximating sums. Hence $\varphi(e)$ can not be smaller than $\varphi(e_1)$. By interchanging e and e_1, the same argument shows that $\varphi(e)$ can not be greater than $\varphi(e_1)$. Hence $\varphi(e) = \varphi(e_1)$.

Next we wish to prove that if e and f are measurable sets without common points,

$$\varphi(e+f) = \varphi(e) + \varphi(f).$$

Let us choose approximating sums

$$\Sigma m(\alpha^i_{a_i}), \quad \Sigma m(\beta^i_{b_i}), \quad \Sigma m(\gamma^i_{c_i})$$

for e, f, and $e+f$ respectively. Regardless of how the first two sums are chosen, we can always choose the third such that

$$\Sigma m(\gamma^i_{c_i}) \leqq \Sigma m(\alpha^i_{a_i}) + \Sigma m(\beta^i_{b_i});$$

hence it follows that

$$\varphi(e+f) \leqq \varphi(e) + \varphi(f).$$

Now suppose that

$$\varphi(e+f) < \varphi(e) + \varphi(f)$$

Assuming, as we may, that $\Sigma m(\gamma^i_{c_i})$ approximates $\varphi(e+f)$ sufficiently closely, it follows that

(1) $$\Sigma m(\gamma^i_{c_i}) < \varphi(e) + \varphi(f).$$

New approximating sums for $\varphi(e)$ and $\varphi(f)$ are furnished by

$$\Sigma m(e\gamma^i_{c_i}) \quad \text{and} \quad \Sigma m(f\gamma^i_{c_i});$$

and since

$$\gamma^i = e\gamma^i + f\gamma^i,$$

it follows that

$$(2) \qquad \Sigma m(\gamma^i_{c_i}) = \Sigma m(e\gamma^i_{c_i}) + \Sigma m(f\gamma^i_{c_i}).$$

Moreover, $\varphi(e)$ and $\varphi(f)$ being lower bounds of their respective approximating sums, we have

$$\Sigma m(e\gamma^i_{c_i}) \geqq \varphi(e), \qquad \Sigma m(f\gamma^i_{c_i}) \geqq \varphi(f).$$

Combining these relations with (1), we obtain

$$\Sigma m(e\gamma^i_{c_i}) + \Sigma m(f\gamma^i_{c_i}) > \Sigma m(\gamma^i_{c_i})$$

which contradicts the equality (2). Hence

$$\varphi(e+f) = \varphi(e) + \varphi(f).$$

We now seek to define the circumstances under which the following situation will arise :

E is an invariant measurable set of non-zero measure; F(P) is a non-negative measurable function defined over E and possesses a finite upper bound M on E. Moreover F(P) may vanish at most on a set of measure zero.

Finally the integral

$$\int_e \mathrm{F(P)}\, d\sigma \qquad (e \leqq \mathrm{E})$$

is invariant under T. (It vanishes only when $m(e) = 0$, by the last assumption on F).

A necessary and sufficient condition for the existence of a set E *and an associated integral* $\int \mathrm{F(P)}\, d\sigma$ *is that* $\varphi(\mathrm{S}) > 0$.

We first assume the existence of E and $\int \mathrm{F(P)}\, d\sigma$ and prove that $\varphi(\mathrm{S}) > 0$.

By the assumptions on E and F(P), it follows that $\int_{\mathrm{E}} \mathrm{F(P)}\, d\sigma > 0$.

If, now, $\varphi(\mathrm{S}) = 0$, it will follow that $\int_{\mathrm{E}} \mathrm{F(P)}\, d\sigma = 0$ which is a contradiction. To show this, let $\Sigma m(\delta^i_{n_i})$ be an approximating sum for $\varphi(\mathrm{E})$. The assumption $\varphi(\mathrm{S}) = 0$ implies that $\varphi(\mathrm{E}) = 0$ and that $\Sigma m(\delta^i_{n_i}) < \varepsilon$.

Then

$$\sum_i \int_{\delta'_{n_i}} F(P)\, d\sigma \leqq M \sum_i m(\delta'_{n_i}) < \varepsilon M.$$

Now since

$$\int_E = \sum_i \int_{\delta^i} = \sum_i \int_{\delta^i_{n_i}},$$

(the integral being invariant) it follows that $\int_E$ is smaller than εM and is therefore equal to zero, wich is the desired contradiction.

That the condition is sufficient follows from the theorem (see Carathéodory, loc. cit.) that a bounded totally continuous additive function of measurable sets is expressible as the indefinite integral of any of its « derivatives ». The function φ is of this type and we may therefore write

$$\varphi(e) = \int_e D(P)\, d\sigma.$$

A derivative is defined as follows : To each point P is associated a sequence of neighborhoods λ^i_P of suitable type closing down on P and so chosen that as $i \to \infty$,

$$\frac{\varphi(\lambda^i_P)}{m(\lambda^i_P)}$$

shall converge to a unique limit. Different derivatives may result from different choices of the regions λ^i_P. All derivatives are summable functions, however, and any two of them differ at most on a set of measure zero.

In view of the inequality $\varphi(\lambda^i_P) \leqq m(\lambda^i_P)$ it follows that every derivative of φ has the upper bound 1.

The set G of points for which $D(P) > 0$ is measurable and its measure is greater than zero, since

$$\int_G D(P)\, d\sigma = \int_S D(P)\, d\sigma = \varphi(S) > 0.$$

Moreover, $\int_{G_k} D(P) > 0$ for ever k, on account of the invariance of $\varphi(e)$. Hence on G_k, $D(P)$ can vanish only over a set of measure zero. Therefore, on the invariant set $E = G + G_1 + G_{-1} + \ldots$, D(P) vanishes at most on a set of measure zero. The set E and the function D(P) taken over E, are precisely the set and associated function demanded by the theorem, and the sufficiency of the condition is thus established.

Let δ be an arbitrary measurable set of non-zero measure on S.

A sufficient condition for the existence of an invariant integral defined over the whole of S *is that*

$$\text{(A)} \qquad \frac{m(\delta_k)}{m(\delta)} > \xi > 0 \qquad (k = \pm 1, \pm 2, \ldots)$$

where ξ *is independent of* δ *and* k.

This is a corollary of preceeding theorem. In fact, an easy consequence of the condition (A) is that $\varphi(e)$ vanishes only when $m(e) = 0$.

We shall now add a few remarks concerning linear dependence. Let us suppose that T admits one or more invariant integrals of the type $\int_e F(P)\,d\sigma$, where, to simplify the discussion, the measurable function F(P) is assumed to be defined over the whole of S and is non-negative thruout, vanishing at most on a set of measure zero. Moreover F(P) will possess a finite upper bound on S. All invariant integrals in this discussion will be of the same type.

The invariant integrals $\int F^1(P)\,d\sigma, \ldots, \int F^n(P)\,d\sigma$ are *linearly dependent* if there exist constants $A^1, \ldots, A^n$ not all zero, such that the function

$$A^1F^1 + \ldots + A^nF^n$$

vanishes « almost everywhere » on S. If no such constants exist, the integrals are *linearly independent*. We shall see how the structure of T is influenced by the existence of several linearly independent integrals.

A transformation will be called *metrically transitive* if there exists no measurable invariant set E such that $0 < m(E) < m(S)$. A transformation of this type is also transitive in the ordinary sense; that is, for any two mutually exclusive connected regions α and β, some power of T can be chosen which will carry points of α into points of β. If for some α and β this were not the case, then E being the invariant set

$$\alpha + \alpha_1 + \alpha_{-1} + \dots,$$

we would have $0 < m(\mathrm{E}) \leqq m(\mathrm{S} - \beta) < m(\mathrm{S})$.

A necessary and sufficient condition that no two invariant integrals on S *be linearly independent is that* T *be metrically transitive.*

The condition is necessary. For suppose that every invariant integral depends linearly on the integral $\int \mathrm{F}(\mathrm{P})\,d\sigma$. Now if there existed an invariant set E with $0 < m(\mathrm{E}) < m(\mathrm{S})$, the invariant integral $\int \mathrm{G}(\mathrm{P})\,d\sigma$, where

$$\begin{aligned} \mathrm{G}(\mathrm{P}) &= \mathrm{F}(\mathrm{P}) \text{ on } \mathrm{E} \\ &= 2\mathrm{F}(\mathrm{P}) \text{ on } \mathrm{S} - \mathrm{E} \end{aligned}$$

is linearly independent of $\int \mathrm{F}(\mathrm{P})\,d\sigma$, which is impossible.

To prove that the condition is sufficient we shall show that if the invariant integrals

$$\mathrm{I}(e) = \int_e \mathrm{F}(\mathrm{P})\,d\sigma \qquad \text{and} \qquad \mathrm{J}(e) = \int_e \mathrm{G}(\mathrm{P})\,d\sigma$$

are linearly independent, T can not be metrically transitive.

Consider the derivatives

$$\mathrm{F}'(\mathrm{P}) = \lim_{i \to \infty} \frac{\mathrm{I}(\lambda_\mathrm{P}^i)}{m(\lambda_\mathrm{P}^i)} \qquad \mathrm{G}'(\mathrm{P}) = \lim_{i \to \infty} \frac{\mathrm{J}(\lambda_\mathrm{P}^i)}{m(\lambda_\mathrm{P}^i)},$$

where, as previously, λ_P^i denotes a sequence of neighborhoods of suitable type closing down on P. For every P, the sequence λ_P^i is assumed to be so chosen that the limits written above, as well as

$$\lim_{i \to \infty} \frac{\mathrm{I}[\mathrm{T}(\lambda_\mathrm{P}^i)]}{m[\mathrm{T}(\lambda_\mathrm{P}^i)]} \qquad \lim_{i \to \infty} \frac{\mathrm{J}[\mathrm{T}(\lambda_\mathrm{P}^i)]}{m[\mathrm{T}(\lambda_\mathrm{P}^i)]}$$

shall exist.

The functions F′ and G′ are equal almost everywhere to F and G respectively. Hence the measurable function $\Psi'(\mathrm{P})$, where

$$\Psi'(\mathrm{P}) = \frac{\mathrm{F}'(\mathrm{P})}{\mathrm{G}'(\mathrm{P})} = \lim_{i \to \infty} \frac{\mathrm{I}(\lambda_\mathrm{P}^i)}{\mathrm{J}(\lambda_\mathrm{P}^i)}$$

is well defined and finite on S, except at most on a set of measure zero, for the points of which we shall arbitrarily assign the value zero.

We shall show that $\Psi'(P)$ is « almost invariant », — i. e. that

$$\Psi'(P) = \Psi'(P_1)$$

except possibly for a set of points P of measure zero. For, the derivatives

$$F''(P_1) = \lim_{i \to \infty} \frac{I[T(\lambda_P^i)]}{m[T(\lambda_P^i)]} \qquad G''(P) = \lim_{i \to \infty} \frac{J[T(\lambda_P^i)]}{m[T(\lambda_P^i)]}$$

are equal almost everywhere to $F'(P_1)$ and $G'(P_1)$ respectively. Hence the function

$$\Psi''(P_1) = \frac{F''(P_1)}{G''(P_1)} = \lim_{i \to \infty} \frac{I[T(\lambda_P^i)]}{J[T(\lambda_P^i)]}$$

is equal to $\Psi'(P_1)$ almost everywhere. But since

$$I[T(\lambda_P^i)] = I(\lambda_P^i), \qquad J[T(\lambda_P^i)] = J(\lambda_P^i),$$

it follows that $\Psi'(P) = \Psi''(P_1)$ for all points P, which proves our assertion.

Let us denote by (a, b), $(0 \leq a \leq b)$, the set of points for which $a \leq \Psi'(P) \leq b$. By the preceeding paragraph, each image of (a, b) under powers of T or T_{-1} is of same measure as (a, b). For, the points P for which $\Psi'(P_1) \neq \Psi'(P)$ are at most of measure zero. From this it follows that for every set (a, b) there exists an invariant set of same measure, namely the set

$$(a, b) + T(a, b) + T_{-1}(a, b) + \ldots.$$

Now since $\Psi'(P)$ is non-negative and finite thruout, we have

$$m(S) \leq m(0, 1) + m(1, 2) + \ldots.$$

Hence there exists a positive number B such that $0 < m(0, B)$. But for no finite number C can we have $m(C, C) = m(S)$. For then we would have $F'(P) = CG'(P)$ and hence $F(P) = CG(P)$ almost everywhere, which contradicts the assumption of linear independence of I and J. It follows that there exists a positive number D such that $0 < m(0, D) < m(S)$. Since there exists a measurable invariant set of measure equal to $m(o, D)$, T can not be metrically transitive. This completes the proof.

Do there actually exist transformations which are metrically transitive? A simple example is the transformation of a torus given in terms of angular coordination by

$$T: \varphi_1 = \varphi + h, \qquad \Theta_1 = \Theta + k.$$

The constants h and k are incommensurable with 2π and with each other. T moreover admits the invariant integral $\int\int d\varphi\, d\Theta$ which is reducible to an integral of the type considered. The proof that T has the property in question offers no difficulty.

5. *Regular regions.* — Returning to the definition of ω- regular points (§ 3), we see that if any exist, their totality constitutes a set of inner points, and falls therefore into a set of maximal connected regions or *components*. Each point on the boundary of a component either belongs to N^1 or is ω-irregular. It follows also from the definition that the transform of an ω-irregular point is again ω-irregular. This holds also for points of N^1, and hence the transform of a component is again a component. Moreover, *a component* C *is either wandering or else periodic or invariant*, for if C intersects C_k, then C and C_k must be identical.

THEOREM. — *A component of* ω- (*or* α-) *regular points is either simply or doubly connected.*

In carrying out the proof, we shall take S to be of genus zero, altho the theorem holds for any genus ([1]).

Let C be a component of ω- regular points (essentially the same argument will hold for an α- regular component) and γ its boundary. There is on S at least one invariant point O ([2]), and we shall consider it as the « point at ∞. » We can then distinguish between the interior and exterior of a simple closed curve drawn on S and not passing

([1]) A proof of this will appear elsewhere.

([2]) First proved by Brouwer (I).

thru O, — in particular, of any simple closed curve contained in C or any image of C.

We shall consider separately the two cases, (I) C is wandering, and (II) C is periodic or invariant.

I. We shall show in this case that there can not be drawn in C a simple closed curve enclosing points of γ, from which it will follow that C is simply connected.

Let α be a simple closed curve in C, and A the region interior to α. Let k be the smallest non-negative integer, if any exists, for which A_k intersects A. Since α lies in a wandering region, no two of its images can intersect, and hence A is expanding or contracting under T_k; we shall assume the former, the argument being quite similar for both cases. Thus A_k contains A, but has no points in common with $A_1, A_2, \ldots, A_{k-1}$. The limit region

$$D = A + A_k + A_{2k} + \ldots$$

is simply connected and invariant under T_k. It is clear moreover that the k regions

$$D_i = A_i + A_{k+i} + A_{2k+i} + \ldots \qquad (i = 0, 1, \ldots, k-1)$$

are each of same type as D, and are mutually exclusive.

Now consider in D the ring $r = (\alpha\alpha_k)$, i. e. the region bounded by α and α_k. The images $r_1, r_2, \ldots, r_{k-1}$ are contained respectively in $D_1, D_2, \ldots, D_{k-1}$, while r_k is adjacent to r, r_{k+1} to $r_2, \ldots$, etc. Clearly r is wandering, and hence contains no points of N^1. Moreover, the points of r are ω-regular. For since the area of r_n converges to zero with $1/n$, the points of r_n tend asymptotically and uniformly toward α_n, which in turn, since α is in C, tends uniformly toward N^1.

It follows that r contains no points of γ, nor of any image of γ. Hence r lies in C, because α does, and also in C_k because α_k does. But this is impossible since C_k can not intersect C. We conclude therefore, that no integer k with the stated property exists, and A must accordingly be wandering. But then the area of A_n converges to zero with $1/n$ and hence A tends uniformly to N^1, since its boundary does. Therefore A contains no ω-irregular points and hence no points of γ. Since α is an arbitrary closed curve in C, it follows that C is simply connected.

II. We shall assume for simplicity that C is invariant, — the argument for the periodic case being essentially the same.

Let us suppose that there can be drawn two non-intersecting simple closed curves α and β whose interior regions A and B both contain points of γ, but have no points in common. We shall show that this situation is impossible, thus proving that C is at most doubly connected.

The closed set α_n tends uniformly toward N^1 as n increases indefinitely. Hence for a sufficiently large positive integer k, α_n fails to intersect α when $n \geqq k$. Then by theorem 4, § 3, A contains no non-wandering points and hence no points of N^1. Moreover, no two regions of the sequence.

$$(\Sigma) \qquad A, \quad A_k, \quad A_{2k}, \quad \ldots$$

can intersect. Hence by the same argument used above, the regions A_{jk} tend uniformly toward N^1 as j increases without limit. This situation holds for each one of the sequences

$$A_i, \quad A_{k+i}, \quad A_{2k+i}, \quad \ldots \qquad (i = 0, 1, \ldots, k-1)$$

since each is of same type as Σ. The totality of these sequences includes all the regions A, A_1, A_2, ... which makes it clear that A consists of ω-regular points. Since, as we have shown, A contains no points of N^1, it follows that A contains no points of γ, contrary to the choice of A. Thus the assumption that $A_n (n \geqq k)$ fails to intersects α, it follows that A is contracting or expanding under some T_m, $m \geqq k$, and hence contains a point U invariant under T_m (Brouwer, I). By precisely the same reasoning, there is a point V in B, invariant under $T_{m'}$, $m' \geqq k$.

Let us join α and β by a simple arc τ in C. By tracing a contour about the set $\alpha + \beta + \tau$ and sufficiently close thereto, we obtain a simple closed curve δ in C, whose interior region D contains no points of γ other than those in A and B. The set $D - (A + B + \alpha + \beta)$ consists entirely of ω-regular points.

We may of course apply the same reasoning as above to D. Hence if s is a sufficiently large multiple of mm', the transformation T_s leaves invariant the points U and V and admits D as an expanding or contracting region.

Consider now the ring r bounded by δ and $T_s(\delta)$. The rings $T_s(r)$ and $T_{-s}(r)$ are adjacent to r, one within and one without. The limit region R consisting of r together with all its images (boundaries included) under positive and negative powers of T_s, is a doubly connected region, invariant under T_s. Let I and E be its inner and outer boundaries.

We now examine separately the two possibilities as to D.

(a) *D is expanding under* T_s. In this case I contains no ω-regular points. For any neighborhood ρ of a point P on I contains points in the limit ring R, and these points are carried by iteration of T_s toward E, whereas all images of P remain on I. Hence from a certain rank on, each member of the sequence

$$\rho, \quad \rho_s, \quad \rho_{2s}, \quad \ldots$$

must intersect δ, which is at a non-zero distance from N'. This makes it clear that P can not be ω-regular under T_s; nor then, under T.

Now I is contained in D. But since I contains no ω-regular points it can not intersect the set $D-(A+B+\alpha+\beta)$. Hence I is contained in $A+B$ and is at a non-zero distance from $\alpha+\beta$. But this is impossible since every approximating curve $T_{ps}(\delta)(p=1, 2, \ldots)$ encloses the points U and V, and therefore contains points not in A or B. This contradiction excludes the possibility that D be expanding under T_s.

(b) *D is contracting under* T_s. Whatever points of N' there may be in D lie in $A+B$. Hence the curve δ, which tends toward N' uniformly on iteration of T_s must eventually be contained in A or B. But this is impossible, since each image of δ encloses the points U and V.

The contradiction in this final possibility shows that there can not exist curves α and β with the stated properties, which completes the proof.

On a surface of genus zero, a regular component of wandering type is simply connected, as follows from the proof of the theorem. It can be shown that this result holds for all surfaces of genus different from 1. For a torus, however, as we shall show later by an example, the regular components of wandering type may be doubly connected.

In a regular component C there are no invariant or periodic points. Hence if C is of invariant simply connected type, it follows from a theorem of Brower (III) that within C, T is topologically equivalent to a translation.

If C is invariant and doubly connected, its boundary consist of two continua, at least if S is of genus zero. Moreover, from the proof of the preceding theorem, each point of C tends asymptotically towards one and the same of these continua, on indefinite iteration of T or T_{-1} according as C is ω-or α-regular.

Let W be the set of all ω-irregular points and A the set of all α-irregular points.

THEOREM 5. — *On a surface of genus zero, each point of* W (*or* A) *is connected to* N^1 *thru* W (*or* A).

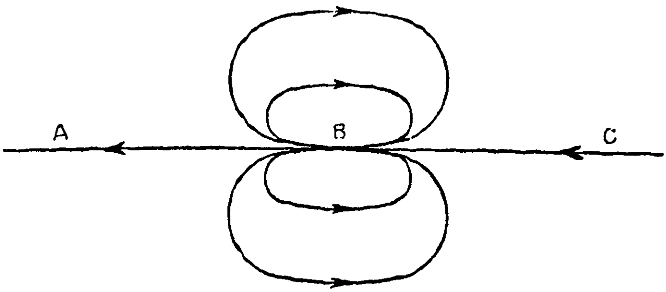
For otherwise it would be possible to draw a simple closed curve ρ enclosing points of W (A) but not of N^1. Thus ρ must lie entirely in some ω-(α-) regular component C. Since boundary points of C are enclosed by ρ, C must be doubly connected and therefore of periodic type. Hence the inner boundary of C, being a closed periodic set, must contain points of N^1, which is impossible. This establishes the theorem.

We shall now consider some simple examples displaying various types of regular components.

The first is a transformation of a sphere illustrated schematically in fig. 2. Here we have taken the plane with a single point at ∞ for our representation of S. The points B and ∞ are the only invariant points, and the set N^1 contains only these points. The motion of the remaining points is indicated by the arrows. Clearly the points on the arc BA∞ (excepting B and ∞) are ω- irregular while all others are ω- regular. Similarly the points on BC∞ are α- irregular and all others are α-regular. (We see here that irregular points need not be *both* α- and ω- irregular.) Hence the region whose boundary consists of BA∞ (BC∞) is an ω- (α-) component of simply connected invariant type. Finally, the example shows how each irregular point is connected to N' thru a continuum of irregular points of same type (theorem 5).

Let us next consider an example in which the regular components are of wandering type. It is easily shown (See Poincaré, I) that if a sense-preserving transformation of a circle C into itself admits no invariant or periodic points, it must be of one of two types. In the first type, N' coincides with C, while in the second, N' is a perfect nowhere dense set on C. Let t be a transformation of the second type. The set C — N' consists of a denumerable infinity of wander-

Fig. 2.



ing open arcs; suppose one of these is δ. Since the end-points of δ and of each image of δ are in N', and since the length of δ_k converges to zero with $1/k$, it follows that S is an α- and ω- regular « component » of wandering type.

Suppose now that S is a sphere, and that C is a great circle, undergoing the transformation t. We may extend t to the whole of S by letting each circle parallel to C undergo the corresponding congruent transformation. It is clear that corresponding to the wandering arcs of C, we now have wandering simply connect regions, and they are α- and ω- regular components of S.

Suppose finally that S is a torus with angular coordinates θ and φ. By letting the circles $\varphi =$ const. undergo congruent transformations of the same type as t, we obtain a transformation of S in which the regular components are wandering rings bounded by circles of the family $\theta = \text{const}$. As we have already pointed out, regular components of this type can only exist on surfaces of genus 1.

6. *The general analytic case.* — Transformations of the general analytic case (§ 1) are free from certain structural complexities and therefore seem best suited for study, in an attempt at systematic structure analysis. In this paper, we can only make a beginning, and must moreover, limit ourselves to the simplest case, — that in which the central motions are finite in number.

THEOREM 6. — *In the general analytic case, there must exist at least two central motions.*

Proof. There exists in any case at least one central motion M. We shall show that in the general analytic case there must exist further central motions.

If the complete sequence M is not pseudo-recurrent, then there are further central motions by theorem 2, § 2. Hence we may suppose M to be pseudo-recurrent. If M contains infinitely many points, its complete group would be a perfect set and would therefore contain central motions other than M, since M is at most denumerable. Hence we may suppose M to consist of a single periodic group. For simplicity let us suppose that M contains but a single point. There is no difficulty in extending the remainder of our argument to the more general situation.

If the invariant point M is of stable type, it is contained in a small expanding or contracting region σ. Hence the closed set $S - \sigma$ is transformed into part of itself by T or T_{-1}, and therefore contains a closed invariant subset which must contain further central motions (§ 2). Thus we may suppose M to be of unstable type. More explicitly, we may suppose M to be of directly unstable type, for if inversely unstable, we may replace T by T_2 in the remainder of the argument, making use of theorem 3, § 3.

We shall show that there exist central points other than the directly unstable point M. Consider the sequence $N^1, N^2, \ldots, N^s = M$ (§ 3) and suppose for the moment that $s > 1$. We may assume that each N^i $(i < s)$ has the property that from it there can be extracted an infinite sequence of points converging to M. For if this were not true for N^1, say, $N^1 - M$ would be a closed (or finite) invariant set and would contain central motions other than M. Let $p^1, p^2, \ldots,$ be a sequence of this type extracted from N^1.

Near M choose four points A, B, C, D on the four invariant branches respectively abutting at M. Let E be the closed set consisting of the four arcs AA_1, BB_1, ..., of the four invariant branches. We assert that E intersects each N^i, $i > s$. For suppose a point Q is very close to M. If Q is on one of the invariant branches, some image of Q will certainly fall on E. In the contrary case, Q will move along close to an α- (ω-) branch on repeated iteration of T_{-1} (T), (see § 1). Hence some image of Q will fall very close to E. From this it follows that out of the set P^1, P^2, ..., and its images, we can extract an infinite sequence of the form.

$$P^1_{\alpha_1}, \quad P^2_{\alpha_2}, \quad \ldots$$

converging to a point L on E. The sequence is contained in N^1, and hence so is L, since N^1 is closed. This proves our assertion for N^1, and the same reasoning applies for each $N^i (i < s)$.

Suppose that the sequence 1, 2, ..., ω, ..., of ordinals less than s possesses no last element. Their N^s consists of those points which are common to all the sets $N^1, N^2, \ldots, N^i, \ldots, (i < s)$. But $EN^1 \geqq EN^2 \geqq \ldots$, and hence there is a closed set of points common to the closed sets $EN^i (i < s)$. This set is of course EN^s and hence N^s contains points other than M; that is, there are central motions other than M.

There remains only the case in which the sequence of ordinals less than s possesses a last element $s - 1$. If Q is a point of N^s (or a point of S different from M, in case $s = 1$) the sequence Q, Q_1, ..., and Q, Q_{-1}, ..., both converge to M. If we recall (§ 1) that points in a small neighborhood σ of M and not lying on any invariant branch of M are carried out of σ on repeated iteration of T or T_{-1}, it becomes clear that Q is doubly asymptotic to M in the sense of § 1, and is in fact a homoclinic point. Hence by the theorem of § 1, there exist central motions other than M. This completes the proof.

Let us suppose that S is a sphere, and examine the structure of T. in the case when there are exactly two central motions.

To every invariant point of S there is associated a number i called its index, and in the general analytic case i may only take the value $+1$ or -1. More explicitly, the directly unstable points are of index -1 and all others of index $+1$. On a sphere, the sum of the indices of the invariant points [1] is always 2. (With regard to the statements above, see Birkhoff, II). Hence, in the general analytic case there are at least two invariant points. But in the case under consideration there are exactly two, say P and Q, since each is a central motion. There are no further periodic points of any order. The index of each point is 1; hence each is of stable or inversely unstable type. If one or both were of the latter type, the sum of the indices of the invariant points of T_2 would be < 2 which is impossible; hence both are of stable type.

About P may be drawn a small circle enclosing a region expanding under T or T_{-1}, — suppose under T. The boundary Σ of the simply connected limit region

$$\sigma + \sigma_1 + \sigma_2 + \ldots,$$

must contain Q, for otherwise, being a closed invariant set, it would contain central motions other than P and Q.

If Σ contains points other than Q, let ρ be a small expanding (or contracting) region containing Q. The closed set $\Sigma - \Sigma_\rho$ is carried into a part of itself by T (or T_{-1}) and therefore contains central motions other than P and Q, which is impossible. Hence $\Sigma = Q$.

Thus S is divided by a system of concentric analytic closed curves into a system of adjacent rings $\ldots, r_{-2}, r_{-1}, r, r_1, r_2, \ldots$. *The rings* $r, r_1, \ldots$, *and* $r, r_{-1}, \ldots$, *close down respectively on the two invariant points, and each* r_n *is carried by* T *into the adjacent ring* r_{n-1}.

In the general analytic case in which the central motions are finite in number, it is probable that a complete structure analysis can be effected, as we shall now indicate.

Let us again take S be a sphere and assume that the number of central motions is finite and greater than 2. From the proof of

---

[1] Assuming them to be finite in number, which they are in the general analytic case.

theorem 6, it is clear that the central points are finite in number, each central motion being either an invariant point or a periodic point group. Let k be the least common multiple of the orders of the various periodic groups. The central points are invariant under T_k.

Suppose that of these invariant points, p are stable, m directly unstable and n inversely unstable, and let p', m', n' represent the corresponding numbers relative to T_{2k}. Then referring to the discussion above relative to indices,

$$p - m + n = 2, \qquad p' - m' + n' = 2.$$

Moreover (§ 1) $p = p'$ and $m \leqq m'$. Hence $n \leqq n'$. But under T_{2k} there are no points of inversely unstable type; hence $n = n' = 0$, and from this we have $p \geqq 2$. Let the totality of stable points be $Q^1, \ldots, Q^p$, $(p \geqq 2)$.

Containing each Q^i, there is a small region σ^i expanding under T_k or T_{-k}. Each Q^i is therefore contained in a simply connected limit region A^i,

$$A^i = \sigma + T_{\varepsilon k}(\sigma) + T_{2\varepsilon k}(\sigma) + \ldots$$

where ε is $+1$ or -1 according as σ is expanding under T_k or T_{-k}. Each A^i is invariant under T_k.

Let the boundary of A^i be denoted by $\alpha^i (i = 1, 2, \ldots, p)$. It is clear that if σ^1, say, consists of only one point, that point together with Q are the only central motions of T_k and hence of T (theorem 3, § 3), whereas we are assuming that the number of central motions is greater than 2. Hence each α^i is a closed periodic continuum and hence contains central points.

Each α^i contains at least one periodic point of directly unstable type.

Proof. Consider α^1, and suppose σ^1 is *expanding* under T_k. Let us suppose moreover that α^1 contains no points of directly unstable type. Then the central points which do lie on α^1 are of stable type. Suppose one of them to be Q^2. Then σ^2 must be *contracting* under T_k. For if σ^2 were expanding, the points of σ^2 would tend toward Q^2 on repeated iteration of T_{-k}; but σ^2 contains points of A^1, and those points must tend toward Q^1 on repeated iteration of T_{-k}, wich would be impossible.

Since then, σ^2 is contracting under T_k and therefore expanding under T_{-k}, the closed set $\alpha^1 - \alpha^1\sigma^2$ is transformed by T_{-k} into a part of itself; hence it contains a closed subset invariant under T_k and therefore contains points of stable type other than Q^1 and Q^2. Suppose that one of these is Q^3. Then the same argument used for σ^2, shows that σ^3 is contracting under T_k. Hence $\alpha^1 - \alpha^1\sigma^2 - \alpha^1\sigma^3$ contains a further stable points, say Q^4. Continuing thus, the set of stable points is eventually exhausted. When this stage is reached, one more application of process must yield a point of directly unstable type on α^1. This completes the proof.

Suppose that the order of the region A^1 is m. Then the regions $A^1, \ldots, A^{m-1}$ are of same type as A^1 and together with A^1 form a periodic set of mutually exclusive regions. Similarly, each A^i belongs to a periodic set of this sort. Two of these sets, however, may overlap.

One method of proceedure would be to study the properties of a maximal set M of periodic sets, all the regions of which are mutually exclusive. Within M, the structure of T is known. On the boundary of M are a number of directly unstable periodic points, and it can be shown that certain of the α — and ω — branches of each of these points must also be contained in M. Finally, the remainder of S falls into a set of connected regions of periodic or wandering type, and these in turn, break up into regular components of various types. It is hoped that a complete analysis of this case, as well as more complex cases will soon be accomplished.

References.

Birkhoff. — I. Quelques théorèmes sur le mouvement des systèmes dynamiques (*Bulletin de la Société mathématique de France*, vol. 40, 1912).

II. Dynamical systems with two degrees of freedom (*Transactions of the American Mathematical Society*, vol. 18, 1917).

III. Surface transformations, and their dynamical applications (*Acta mathematica*, vol. 43, 1920).

IV. Ueber gewisse Zentralbewegungen dynamischer Systeme (*Göttinger Nachrichten*, 1926).

V. On the periodic motions of dynamical systems (*Acta mathematica*, vol. 50, 1927).

BROUWER. — I. Continuous one-one transformations of surfaces in themselves (*Proceedings of the Section of Sciences, Koninklijke Academie van Wetenschappen te Amsterdam*, vol. 11-15, 1908-1912).

II. Beweis des ebenen Translationssatzes (*Mathematische Annalen*, Bd 72, 1912).

LATTÈS. — Sur les équations fonctionnelles qui définissent une courbe ou une surface invariante par une transformation (*Annali di Matematica*, 3e série, vol. 13, 1907).

POINCARÉ. — I. Sur les courbes définies par les équations différentielles (*Journal de Mathématiques*, 3e série, vol. 7-8, 1881-1882, et 4e série, vol. 1-2, 1885-1886).

II. Méthodes nouvelles de la Mécanique céleste, vol. 3.