APPROXIMATION THEORIES FOR MEASURE PRESERVING TRANSFORMATIONS

BY

PAUL R. HALMOS

1. **Introduction.** The purpose of this paper is to introduce approximative notions into the theory of measure preserving transformations in the hope that they will turn out to be useful tools in investigating some of the outstanding problems of ergodic theory.

By a measure preserving transformation one means, ordinarily, a one-to-one mapping of a measure space on itself which is such that both it and its inverse preserve the measurability and the numerical measure of measurable sets. Since, however, it is customary in measure theory (and, from the point of view of purity of method, highly desirable) to identify two transformations, or functions, or sets, which agree except for a set of measure zero, the treatment below discusses not measure spaces but rather their Boolean algebras, consisting of the measurable sets modulo sets of measure zero. It is known that such "measure algebras" are simply characterizable in algebraic terms and that from this point of view measure preserving transformations are nothing other than automorphisms of the underlying algebra. A consistent adherence to this point of view led me to some simple but illuminating facts (Theorem 1) which exhibit the reason why some common measure theoretic devices work.

Three notions of approximation—that is, three topologies—are introduced into the set G of all measure preserving transformations. For the reader familiar with von Neumann's fundamental work on Hilbert space two of these are best described as the analogues (and, indeed, relativizations) of the "strong neighborhood" and the "uniform" topologies for bounded operators. The third (called the "metric" topology below) is defined in terms of a distance function, the distance between two transformations being the measure of the set of points where they differ.

The results proved are of three types. The first type is purely technical: I investigate the relations of the various topologies to each other and to the group structure of G. It is not surprising that G turns out to be a topological group with more or less decent properties in all three cases (Theorems 2 and 7). The interesting and nontrivial fact along these lines is that the uniform and the metric topologies are the same (Theorem 10).

The second type of result is the one I consider most important: it asserts that arbitrary measure preserving transformations may be approximated by

Presented to the Society, April 24, 1943; received by the editors March 10, 1943.

transformations with comparatively simple properties (for example, by transformations of finite period). These results (Theorems 3, 4, and 8) are the analogues of various theorems asserting the density of step functions in sundry function spaces and will, I hope, find similar use.

The third and final type of result, generally an easy consequence of the approximation theorems just mentioned, is motivated by the oldstanding conjecture that "in general a measure preserving transformation is ergodic." I adopt the usual, by now classical, interpretation of "in general" (that is, "except for a set of first category") and show that the adage is true in the neighborhood topology(1) (Theorem 6) and false in the metric topology (Theorem 9). I hope in the near future to be able to apply these same methods to prove analogous theorems about the much more difficult class of mixing transformations.

2. **The Boolean algebra.** Let B be a Boolean σ-algebra(2) in which the null and unit elements, the Boolean operations of union and intersection, the inclusion relation, and the complement of an element a are denoted by the familiar symbols 0, 1, $\cup$, $\cap$, $\subset$, and a', respectively. It is well known that if addition and multiplication are defined in B by the formulas

$$a + b = (a \cap b') \cup (a' \cap b), \qquad ab = a \cap b,$$

then B becomes an algebraic ring(3). In this ring multiplication is commutative ($ab = ba$), every element is idempotent ($a^2 = a$), and addition is modulo 2 ($a + a = 0$). It follows that the relations

$$a + b = a - b, \qquad a' = 1 - a,$$
$$a \cup b = a + b + ab, \qquad a - b = a' - b'$$

are identities in B; in the sequel their right and left sides will be used interchangeably, the choice in any particular case being guided by convenience and intuitive content alone.

Suppose moreover that there is defined on B a numerically valued, positive, countably additive, finite measure: the measure of an element a is to be denoted by $|a|$(4). The existence of such a measure has a profound effect on

(1) The first theorem of this type is due to J. C. Oxtoby and S. M. Ulam, *Measure-preserving homeomorphisms and metrical transitivity*, Ann. of Math. (2) vol. 42 (1941) p. 880. Their topology is, however, very different from mine and depends on the topological and metric (as opposed to purely measure theoretic) structure of the underlying space.

(2) See Garrett Birkhoff, *Lattice theory*, Amer. Math. Soc. Colloquium Publications, vol. 25, New York, 1940, pp. 29 and 88.

(3) See M. H. Stone, *The theory of representations for Boolean algebras*, Trans. Amer. Math. Soc. vol. 40 (1936) p. 43.

(4) In other words: $|a| = 0$ is equivalent to $a = 0$, $a_i a_j = 0$ for $i \neq j$, $i, j = 1, 2, \cdots$ implies $|\bigcup_i a_i| = \sum_i |a_i|$, and $|1| < \infty$. Boolean algebras satisfying all these conditions are called *measure algebras*. Cf. Dorothy Maharam, *On homogeneous measure algebras*, Proc. Nat. Acad. Sci. U.S.A. vol. 28 (1942) p. 108.

the algebraic structure of B. Possibly the most surprising result is that the presence of the measure implies that B is complete in the sense of lattice theory. If, in other words, E is an arbitrary, not necessarily countable, subset of B then there exists an element $a \in B$ such that $x \subset b$ for all $x \in E$ is equivalent to $a \subset b$(5). This fact will be used quite frequently in the work that follows: the element a so defined will be denoted by $a = \sup E$.

The most usual form in which such suprema are encountered is the following. If E is any subset of B denote by E^* the set of all those elements of B every nonzero subelement of which contains a nonzero subelement belonging to E. (In other words, E^* is the set of all elements arbitrarily small parts of which belong to E.) Similarly, denote by E_* the set of all those elements of B *every* nonzero subelement of which belongs to E. Then the lattice completeness of B implies the following.

THEOREM 1. *E^* is always a principal ideal (that is, there is an element $e^* \in B$ such that $a \in E^*$ if and only if $a \subset e^*$); E_* is a principal ideal if and only if for all subsets A of E_*, containing a nonzero element, $\sup A \in E$. The generator $e^* = \sup E^*$ of E^* and the element $\tilde{e}_* = \sup (E')_*$ (which is the generator of $(E')_*$ in case $(E')_*$ is a principal ideal) are each other's complements*(6).

Proof. Observe first that $0 \in E^*$ and $0 \in E_*$. This paradoxical fact is due to the wording of the definitions of E^* and E_*: since 0 contains no nonzero subelements it vacuously satisfies the conditions of both definitions.

To prove that E^* is a principal ideal observe that if $a \in E^*$ and $b \subset a$ then $b \in E^*$. In consequence of this fact it is sufficient to prove that $e^* = \sup E^*$ belongs to E^*. Suppose therefore that $a \subset e^*$, $a \neq 0$. Then (from the definition of e^*) there must be an element $b \in E^*$ for which $ab \neq 0$. Since $ab \subset b$ and $b \in E^*$ there is an element $c \subset ab$, $c \neq 0$, for which $c \in E$. Since $c \subset ab \subset a$, this shows that every nonzero $a \subset e^*$ contains at least one nonzero c in E.

The sufficiency proof for E_* is similar to the above: once more $a \in E_*$ and $b \subset a$ implies $b \in E_*$, so that it is sufficient to prove that $e_* = \sup E_*$ belongs to E_*. Take $a \subset e_*$, $a \neq 0$; then $a = \sup \{b: b \subset a, b \in E_*\}$(7), whence, by the assumed condition, $a \in E$ and therefore $e_* \in E_*$. The necessity of the condition is even easier to see: if E_* is a principal ideal and $A \subset E_*$ then $\sup A \in E_*$, and consequently, since every nonzero element of E_* belongs to E, $\sup A \in E$.

Finally, if $a \in E^*$ and $b \in (E')_*$ then $ab = 0$. For otherwise $ab \subset b$ would imply that all nonzero subelements of ab are in E' and $ab \subset a$ would imply that some nonzero subelement of ab is in E. Hence e^* and $\tilde{e}_*$ are disjoint; to prove that their union is 1, suppose that $a \neq 0$ is disjoint from $\tilde{e}_*$. Then no nonzero subelement of a can be in $(E')_*$. Hence every nonzero subelement

(5) See Birkhoff, op. cit., p. 100.

(6) The Boolean algebra symbols, E', $E \cup F$, and so on, are used for the algebra of *subsets* of B, as well as for the abstract algebra of *elements* of B.

(7) The symbol $\{x: \cdots\}$ stands, as usual, for "the set of all x's for which $\cdots$."

of a contains a nonzero subelement in E; that is, $a \in E^*$ or $a \subset e^*$. This completes the proof of Theorem 1.

COROLLARY. *If, for every subset A of E,* sup A *belongs to E then E_* is a principal ideal.*

This follows immediately from Theorem 1 and the remark (already used above) that every nonzero element of E_* belongs to E.

The presence of a measure serves also to introduce a natural metric topology into B: the distance between a and b may be defined by $|a-b|$. It is easy to verify that the distance axioms are satisfied; a nontrivial conclusion is that B thereby becomes a *complete* metric space(8). The following inequalities will be useful later:

$$\left| |a| - |b| \right| \leqq |a-b|, \tag{1}$$

$$|(a_1+b_1)-(a_2+b_2)| \leqq |a_1-a_2| + |b_1-b_2|, \tag{2}$$

$$|a_1b_1 - a_2b_2| \leqq |a_1-a_2| + |b_1-b_2|, \tag{3}$$

$$|(a_1 \cup b_1)-(a_2 \cup b_2)| \leqq |a_1-a_2| + |b_1-b_2|. \tag{4}$$

The first of these follows from the relations $a=ab+ab'$ and $b=ab+a'b$, (2) is a consequence of the additivity of measure, (3) follows from (2) and the fact that $|ab| \leqq |a|$, and (4) follows from (3) by taking complements. It follows, of course, from these inequalities of Lipschitz type that each of the functions $|a|$, $a+b$, ab, and $a \cup b$ is a (uniformly) continuous function of all its arguments.

Finally the following three normalizing assumptions about B will be made throughout, unless the contrary is explicitly stated.

(*) $|1|=1$.

(**) B is non-atomic (that is, for every $a \in B$, $a \neq 0$, there is an element $b \in B$ such that $b \neq 0$, $b \neq a$, $b \subset a$).

(***) As a metric space B is separable.

Most of the interesting results of this paper remain true, though not interesting, in case (**) is not assumed. The cardinal number restriction (***) is merely a matter of convenience. One of its main purposes is to concretize the object of study (that is, the algebra B). It is known that the axioms of a measure algebra together with (*), (**), and (***) are categorical: any system satisfying them is isomorphic to the measure algebra of all measurable sets modulo sets of measure zero of, say, the unit interval(9). This representation is frequently useful and will be exploited below.

(8) The proof of this fact is very similar to the proof of the Riesz-Fischer theorem and can, in fact, be made to follow from it.

(9) See for example Paul R. Halmos and John von Neumann, *Operator methods in classical mechanics*. II, Ann. of Math. (2) vol. 43 (1942) p. 335.

3. **The neighborhood topology.** An *automorphism* of B is a one-to-one mapping T of B onto itself such that

$$|Ta| = |a|, \qquad Ta' = (Ta)',$$
$$T(\bigcup_i a_i) = \bigcup_i Ta_i, \qquad T(\bigcap_i a_i) = \bigcap_i Ta_i \text{(10)}.$$

Under the operation of composition the set G of all automorphisms of B is a group.

Let S be any element of G and take $a \in B$, $\epsilon > 0$. Writing

$$N(S) = N(S; a, \epsilon) = \{T: |Sa - Ta| < \epsilon\},$$

a unique topology is defined in G by the requirement that the collection of all sets of the form $N(S)$ be a subbase for the open sets(11). I shall call this the *neighborhood topology* of G(12); sets of the form $N(S)$ are *subbasic neighborhoods* of S, and finite intersections of subbasic neighborhoods are *basic neighborhoods* of S.

THEOREM 2. *In the neighborhood topology G is a complete topological group satisfying the first countability axiom*(13).

Proof. (1) If S_1, $S_2 \in G$ and $S_1 \neq S_2$ then there is an $a \in B$ such that $\epsilon = |S_1a - S_2a| > 0$. Consequently $N(S_1) = \{T: |S_1a - Ta| < \epsilon\}$ is a neighborhood of S_1 which does not contain S_2, so that G is a T_0-space.

(2) Consider any S_0, $T_0 \in G$ and any subbasic neighborhood N_0 of their "quotient" $S_0T_0^{-1}$,

$$N_0 = N(S_0T_0^{-1}; a, \epsilon) = \{R: |S_0T_0^{-1}a - Ra| < \epsilon\}.$$

Write $b = T_0^{-1}a$ (so that $a = T_0b$) and consider any $S \in N(S_0; b, \epsilon/2)$ and $T \in N(T_0; b, \epsilon/2)$. Then

$$\begin{aligned} |S_0T_0^{-1}a - ST^{-1}a| &\leqq |S_0T_0^{-1}a - ST_0^{-1}a| + |ST_0^{-1}a - ST^{-1}a| \\ &= |S_0b - Sb| + |b - T^{-1}T_0b| \\ &= |S_0b - Sb| + |Tb - T_0b| < \epsilon. \end{aligned}$$

Consequently ST^{-1} is a continuous function of both its arguments.

---

(10) It is known that the conditions $|Ta| = |a|$ and $T0 = 0$ are sufficient to imply that T is an automorphism. This fact is a member of a class of theorems of which a much better known specimen is the assertion that an isometry of a finite dimensional unitary space, which leaves the origin fixed, is a unitary transformation, that is, an automorphism of the space. Such theorems greatly reduce the labor of verifying that a given transformation is an automorphism.

(11) See Solomon Lefschetz, *Algebraic topology*, Amer. Math. Soc. Colloquium Publications, vol. 27, New York, 1942, p. 6.

(12) This topology is the analogue of the "strong" topology of operators on Hilbert space and may in fact be obtained from the strong topology by relativization.

(13) For the notion of completeness for topological groups see André Weil, *Sur les espaces a structure uniforme et sur la topologie générale*, Paris, 1938, p. 29.

(3) Since G has already been shown to be a topological group it is sufficient to investigate first countability in the neighborhood of the identity. Let $a_1, a_2, \cdots$ be a countable dense set in B; write $N_{ij}=N(I;\ a_i,\ 1/j)$, $i, j=1, 2, \cdots$; then the N_{ij} are a subbase at I (the identity). For given any $N=N(I;\ a,\ \epsilon)$, choose i so that $|a-a_i|<\epsilon/3$ and choose j so that $1/j<\epsilon/3$. Then, if $T\in N_{ij}$,

$$|a-Ta| \leqq |a-a_i|+|a_i-Ta_i|+|Ta_i-Ta|<\epsilon,$$

so that $N_{ij}\subset N$; q.e.d.

(4) In a topological group satisfying the first countability axiom completeness is equivalent to sequential completeness. In order to prove that every convergent sequence has a limit, I first remark that $\{T_n\}$ is convergent if and only if the sequence $\{T_na\}$ of elements of B is convergent for every $a\in B$. Consequently, if $\{T_n\}$ is convergent a unique T is defined by $Ta=\lim_n T_na$. It is easy to see that $Ta'=(Ta)'$ and $T(a_1\cup a_2\cup\cdots\cup a_n)=Ta_1\cup Ta_2\cup\cdots\cup Ta_n$; the continuity of $|a|$ implies then that T is measure preserving and therefore that $T\in G$. This concludes the proof of Theorem 2.

In order to state one of the principal results of this paper it is convenient (though not necessary) to assume for a moment that the elements of B are measurable sets in the unit interval (or, more precisely, residue classes of such sets modulo sets of measure zero). Call an interval $(k/2^n,\ [k+1]/2^n)$, $k=0, 1, \cdots, n-1$; $n=0, 1, 2, \cdots$, a *dyadic interval of rank n*, and a union of such intervals a *dyadic set of rank n*. A *permutation P* (or, more precisely, a dyadic permutation of rank n) is a one-to-one transformation of the interval which maps each dyadic interval of rank n into itself or into another one by an ordinary translation. The reason for introducing the interval was to avoid an abstract and complicated description of this comparatively simple class (the dyadic sets) of elements of B and the related class (the dyadic permutations) of elements of G. In what follows only one property of dyadic sets will be used (in addition, that is, to their simple algebraic structure) and that is that they form a (countable) dense set in B.

THEOREM 3. *In the neighborhood topology permutations are dense in* G[14].

Proof. (1) The idea of the proof is simple. It is to be proved that given any n elements $a_1, \cdots, a_n$ of B and any automorphism $T\in G$ there is a

[14] This theorem has the rank of a "folk theorem," that is, one that most measure theorists conjecture and prove in one form or another. It has a satisfying intuitive content: it says that, in the limit, every measure preserving transformation is obtained by cutting up the space (with an ordinary pair of Euclidean scissors) into a finite number of pieces and then merely permuting the pieces. The first precise formulation of this result (not the one above) I heard from John von Neumann in November, 1940. His formulation appears below in Theorem 8. The first published version (different from both von Neumann's and mine) is due to Oxtoby and Ulam, op. cit. p. 919.

permutation P which affects the a_i approximately the same way as T. To do this consider the atoms of the (finite) Boolean algebra generated by $a_1, \cdots, a_n, Ta_1, \cdots, Ta_n$. Approximate each of these atoms (and consequently all unions of such atoms) by dyadic sets and then permute the dyadic sets as necessary. The somewhat cumbersome details follow.

(2) Let $b_1, \cdots, b_m$ be a decomposition of 1 (that is, $b_ib_j=0$ for $i\neq j$ and $b_1\cup \cdots \cup b_m=1$), and let δ be any positive number. Let $\bar{b}_1, \cdots, \bar{b}_m$ be dyadic elements such that $|b_i-\bar{b}_i|<\delta$, $i=1, \cdots, m$. Write

$$\tilde{b}_i = \bar{b}_1' \cdots \bar{b}_{i-1}'\bar{b}_i\bar{b}_{i+1}' \cdots \bar{b}_m',$$

and observe that the disjointness of the b_i implies that

$$b_i = b_1' \cdots b_{i-1}'b_ib_{i+1}' \cdots b_m'.$$

Consequently

$$\begin{aligned} |b_i - \tilde{b}_i| \leqq |b_1' - \bar{b}_1'| + \cdots + |b_{i-1}' - \bar{b}_{i-1}'| + |b_i - \bar{b}_i| \\ + |b_{i+1}' - \bar{b}_{i+1}'| + \cdots + |b_m' - \bar{b}_m'| < m\delta \end{aligned}$$

(since $|a-b|=|a'-b'|$). The $\tilde{b}_i$, being finite intersections of dyadic elements, are themselves dyadic; moreover they are pairwise disjoint. Hence $\tilde{b}=\tilde{b}_1\cup \cdots \cup\tilde{b}_m$ is dyadic. Also, since $b_1\cup \cdots \cup b_m=1$, it follows that

$$|1 - \tilde{b}| \leqq \sum_i |b_i - \tilde{b}_i| < m^2\delta.$$

Split the complement $1-\tilde{b}$ of the dyadic element $\tilde{b}$ into m equal dyadic pieces and add one of these pieces to each b_i: denote the elements so obtained by b_i^*. Then $|b_i^*-\tilde{b}_i|<m^2\delta/m=m\delta$, so that

$$|b_i - b_i^*| \leqq |b_i - \tilde{b}_i| + |\tilde{b}_i - b_i^*| < 2m\delta.$$

Finally let $(i_1, i_2, \cdots, i_k)$ be any subset of $(1, 2, \cdots, m)$. Then

$$|(b_{i_1}\cup \cdots \cup b_{i_k}) - (b_{i_1}^*\cup \cdots \cup b_{i_k}^*)| \leqq \sum_j |b_{i_j} - b_{i_j}^*| < 2km\delta \leqq 2m^2\delta.$$

To sum up: corresponding to each decomposition $\{b_i\}$ of 1 there is a dyadic decomposition $\{b_i^*\}$ of 1 such that every union of the b's is arbitrarily close to the corresponding union of the b^*'s.

(3) Suppose now that $b_1, \cdots, b_m$ and $c_1, \cdots, c_m$ are two decompositions of 1 with the property $|b_i|=|c_i|$, $i=1, \cdots, m$. Then there exist arbitrarily close dyadic decompositions $\{b_i^*\}$ and $\{c_i^*\}$ with $|b_i^*|=|c_i^*|$. To see this apply (2) to the b's and c's separately and then alter the c^*'s (say) by subtracting from the fat ones and adding to the thin ones until the desired result is achieved. The hypothesis $|b_i|=|c_i|$ insures that the approximating property of the c_i^* will not be lost during the alterations.

(4) Consider now any automorphism $S \in G$, and any (basic) neighborhood N of S,

$$N(S) = \{T: |Sa_i - Ta_i| < \epsilon, i = 1, \cdots, n\}.$$

Let $b_1, \cdots, b_m$ be the atoms of the Boolean algebra generated by $a_1, \cdots, a_n$, and write $c_i = Sb_i$, $i = 1, \cdots, m$. According to (3) there are dyadic decompositions $\{b_i^*\}$ and $\{c_i^*\}$ such that every union of b's (or of c's) differs by less than $\epsilon/2$ from the corresponding union of the b^*'s (or of the c^*'s) and such that $|b_i^*| = |c_i^*|$. Construct a dyadic decomposition of 1 so fine that each b_i^* and each c_i^* is a union of intervals of this decomposition. Since the b_i^* (as well as the c_i^*) are pairwise disjoint and since (having the same measure) corresponding ones contain the same number of these small dyadic intervals, it is clear that there exists a permutation P for which $Pb_i^* = c_i^*$.

(5) Each a_j, $j = 1, \cdots, n$, is a union of certain b_i's: denote by a_j^* the corresponding union of the b_i^*'s. Then $|Sa_j - Pa_j| \leqq |Sa_j - Pa_j^*| + |Pa_j^* - Pa_j|$. Since Sa_j is a union of c's and Pa_j^* is the corresponding union of c^*'s, each of the two terms of the right side of the last written inequality is dominated by $\epsilon/2$, so that $P \in N$; q.e.d.

COROLLARY. *In the neighborhood topology G satisfies the second countability axiom.*

For in any topological group (uniform structure) the first countability axiom and the existence of a countable dense set are together equivalent to the second countability axiom. The set of permutations is countable.

The next theorem is to be used in proving Theorem 6 below; it is however of a certain interest of its own as a sharpening of Theorem 3.

THEOREM 4. *In the neighborhood topology cyclic permutations are dense in G.*

Proof. Since it is already known that permutations are dense it is sufficient to prove that if P is any permutation and

$$N = \{T: |Pa_i - Ta_i| < \epsilon, i = 1, \cdots, n\}$$

is any neighborhood of it then there is a cyclic permutation $Q \in N$. It is, moreover, no loss of generality to assume that the a_i are dyadic since the neighborhoods for which this is true form a base for open sets. Since in this case the Pa_i are also dyadic, there exists a decomposition of 1 into dyadic intervals b_j of sufficiently high constant rank such that all a_i and all Pa_i are unions of the b_j. Write P in the ordinary cyclic notation as a permutation on the b's indicating all cycles (even those of length one):

$$P: (b_1^{(1)} \cdots b_{n_1}^{(1)})(b_1^{(2)} \cdots b_{n_2}^{(2)}) \cdots (b_1^{(p)} \cdots b_{n_p}^{(p)}).$$

(It is true, but irrelevant, that $n_1 + n_2 + \cdots + n_p$ is a power of two.) Choose

m so large that $1/2^m < \epsilon/2p$, and consider the decomposition of 1 into the dyadic intervals of rank m. All $b_i^{(j)}$ contain the same number, say q, of these small dyadic intervals:

$$b_i^{(j)} = {}^1c_i^{(j)} \cup {}^2c_i^{(j)} \cup \cdots \cup {}^qc_i^{(j)}.$$

A unique (cyclic) permutation Q is defined by the conditions

$$\begin{aligned}
&{}^kc_i^{(j)} \to {}^kc_{i+1}^{(j)}, && \text{except for } i = n_j;\\
&{}^kc_{n_j}^{(j)} \to {}^{k+1}c_1^{(j)}, && \text{except for } k = q;\\
&{}^qc_{n_j}^{(j)} \to {}^1c_1^{(j+1)}, && \text{except for } j = p;\\
&{}^qc_{n_p}^{(p)} \to {}^1c_1^{(1)}.
\end{aligned}$$

It follows that for $i \neq n_j$, $Qb_i^{(j)} = Pb_i^{(j)}$, and $|Qb_{n_j}^{(j)} - Pb_{n_j}^{(j)}| = 2/2^m < \epsilon/p$. Since any union of b's contains at most p of the $b_{n_j}^{(j)}$, and since a_i is such a union for all $i = 1, \cdots, n$, it follows finally that $|Qa_i - Pa_i| < p \cdot \epsilon/p = \epsilon$, so that $Q \in N$, q.e.d.

The proof above establishes also the following corollary.

COROLLARY. *If P is any permutation and $N = N(P)$ any dyadic neighborhood, $N(P) = \{T: |Pa_i - Ta_i| < \epsilon, i = 1, \cdots, n\}$, then there is a positive integer m and a cyclic permutation Q of rank m such that $Q \in N$ and such that each a_i is a union of the dyadic intervals of rank m.*

To state the next theorem it is necessary to recall the definition of an ergodic automorphism. $T \in G$ is *ergodic* if $Ta = a$ implies $a = 0$ or 1. The notion of an ergodic automorphism is the weakest formulation, which is simultaneously useful and precise, of the intuitive concept of a transformation which is a thorough shuffling.

THEOREM 5. *In the neighborhood topology ergodic automorphisms are dense in G.*

Proof. In view of Theorem 3 and the corollary to Theorem 4 it will be sufficient to prove the following statement: given any cyclic permutation Q of rank m there is an ergodic transformation T which agrees with Q on all dyadic elements of rank m. To prove this let $c_0, c_1, \cdots, c_{r-1}$ be the dyadic intervals of rank m ($r = 2^m$) with the notation so chosen that $Qc_i = c_{i+1}$ for $i = 0, 1, \cdots, r-1 \pmod r$. Let S be an ergodic automorphism of the Boolean algebra of all subelements of c_0([15]), and define, for any $a \in b$, $a = ac_0 \cup ac_1 \cup \cdots \cup ac_{r-1}$,

([15]) That such an S exists is well known: it follows, for example, from the possibility of representing B on the unit interval, as described above.

$$Ta = QS(ac_0) \cup Q(ac_1) \cup Q(ac_2) \cup \cdots \cup Q(ac_{r-1}).$$

The verification that T is an automorphism is quite mechanical; it is necessary only to prove that T is ergodic. Suppose therefore that $Ta=a$. This implies that $(Ta)c_i=ac_i$ for $i=0, 1, \cdots, r-1$; since $(Ta)c_i=T^i(ac_0)$ it follows that $ac_i=T^i(ac_0)$ for all i (mod r). Using, however, the definition of T,

$$T(ac_0) = QS(ac_0),\ T^2(ac_0) = Q^2S(ac_0), \cdots, T^r(ac_0) = Q^rS(ac_0) = S(ac_0).$$

Consequently $ac_0=T^r(ac_0)=S(ac_0)$, whence, by the ergodicity of S, $ac_0=0$ or $ac_0=c_0$. Since $ac_i=T^i(ac_0)$ it follows that either $ac_i=0$ for all i or else $ac_i=c_i$ for all i. In other words $a=0$ or 1; q.e.d.

THEOREM 6. *In the neighborhood topology the set E of ergodic automorphisms is a residual G_δ.*

Proof. Since E is already known to be dense and since a dense G_δ is residual, it will be sufficient to prove that E is a G_δ. To prove this I shall make use of the following fact([16]): an automorphism $T\in G$ is ergodic if and only if, for every a and $b\in B$, $\lim_n n^{-1}\sum_{k=0}^{n-1}|a\cdot T^kb| = |a|\cdot|b|$. Let $a_1, a_2, \cdots$ be a dense sequence in B; write

$$E(i, j, m, n) = \left\{T: \left| n^{-1}\sum_{k=0}^{n-1} |a_i\cdot T^ka_j| - |a_i|\cdot|a_j| \right| < m^{-1}\right\}$$

and

$$F = \bigcap_i\bigcap_j\bigcap_m\bigcup_n E(i, j, m, n).$$

It follows trivially that $E\subset F$. Suppose therefore that $T\in G$ is not ergodic; it is to be shown that T does not belong to F. Since T is not ergodic there is an element $b\in B$ for which $Tb=b$ and $|b|\cdot|b'|>0$. Write $\delta=|b||b'|/8$, choose i so that $|a_i-b'|<\delta$, choose j so that $|a_j-b|<\delta$, and choose m so that $1/m<4\delta$. Then

$$\Big|\,|a_i\cdot T^ka_j| - |b'\cdot T^kb|\,\Big| \leqq |a_i\cdot T^ka_j - b'\cdot T^kb| \leqq |a_i - b'| + |a_j - b| < 2\delta,$$

for all k and therefore

$$\left| n^{-1}\sum_{k=0}^{n-1} |a_i\cdot T^ka_j| - n^{-1}\sum_{k=0}^{n-1} |b'\cdot T^kb| \right| < 2\delta$$

for all n. Also

([16]) The first use of this result for a very similar purpose is due to Oxtoby and Ulam, op. cit. p. 904. For a proof of the equivalence of this limiting condition to ergodicity see Eberhard Hopf, *Ergodentheorie*, Berlin, 1937, p. 30. The condition has a very natural intuitive interpretation. Two elements a and b of the measure algebra B are called *independent* if $|ab|=|a||b|$. (This definition is motivated by probability theory.) Hence the condition says that an element moving under the influence of an ergodic automorphism tends, on the average, to become independent of every fixed element.

$$\left| \, |b'|\,|b| - |a_i|\,|a_j| \, \right| < |b'| \cdot \left| \, |b| - |a_j| \, \right| + |a_j| \cdot \left| \, |b'| - |a_i| \, \right| < 2\delta,$$

whence it follows (by the familiar trick of adding and subtracting both $n^{-1}\sum_{k=0}^{n-1}|a_i \cdot T^k a_j|$ and $|a_i|\,|a_j|$) that

$$\left| \, n^{-1}\sum_{k=0}^{n-1} |b' \cdot T^k b| - |b'|\,|b| \, \right| < \left| \, n^{-1}\sum_{k=0}^{n-1} |a_i \cdot T^k a_j| - |a_i|\,|a_j| \, \right| + 4\delta.$$

Since $Tb = B$ implies $T^k b = b$ for all k, the left side of this inequality is equal to $|b'|\,|b|$ and consequently to 8δ. Hence, finally,

$$\left| \, n^{-1}\sum_{k=0}^{n-1} |a_i \cdot T^k a_j| - |a_i|\,|a_j| \, \right| > 4\delta > m^{-1}$$

for all n: this proves that T is not in F, q.e.d.

Two comments are in order concerning the interpretation of Theorem 6.

I. Category results are usually stated for *complete* metric spaces only, whereas the topology of G was defined in terms of neighborhoods. It is easy to see, however, that the ordinary proof of Baire's theorem goes through with only verbal changes in any uniform space satisfying the first countability axiom. More than this is true: such a uniform space is always metrizable([17]). For topological groups a still stronger result is true: the first countability axiom implies the existence of a metric invariant under all left translations([18]). It is easy to exhibit such a metric for the group G discussed above. Let $a_1, a_2, \cdots$ be a dense sequence in B; the distance function $d(S, T) = \sum_{n=1}^{\infty}(1/2^n)|Sa_i - Ta_i|$ has all the desired properties. Since, however, this rather artificial construction did not seem to throw any light on the structure of G it was omitted from the main body of the discussion.

The ideal situation would be the existence of a metric for G invariant simultaneously under right and left translations. The following (unfortunately quite involved and indirect) argument will show that this is not the case. I shall need to make use of van Dantzig's necessary and sufficient condition for the (invariant) metrizability of a separable topological group([19]) and of the following fact concerning compact groups. If Γ is any compact separable topological group, denote by T_α, for $\alpha \in \Gamma$, the (Haar) measure preserving transformation defined by $T_\alpha \gamma = \alpha\gamma$. T_α may be looked upon as an automorphism of the Boolean algebra B of all (Haar) measurable sets modulo sets of measure zero; the correspondence $\alpha \to T_\alpha$ is a homeomorphic isomor-

([17]) See André Weil, op. cit. p. 16.

([18]) See Shizuo Kakutani, *Über die Metrisation der topologischen Gruppen*, Proceedings of the National Academy of Japan vol. 12 (1936) p. 82.

([19]) The condition is that $T_n \to I$ should imply $S_n T_n S_n^{-1} \to I$ for all sequences $\{S_n\}$. See D. van Dantzig, *Zur topologischen algebra*. I. *Komplettirungstheorie*, Math. Ann. vol. 107 (1932) p. 16.

phism between all Γ and a part of G([20]) if G is thought of as topologized by the neighborhood topology. If S is any continuous automorphism of the compact group Γ then S is a (Haar) measure preserving transformation([21]) (so that S is also an automorphism of B, $S \in G$); a trivial computation shows that $ST_\alpha S^{-1} = T_{S(\alpha)}$. Putting all this together, the impossibility of invariant metrizability will be proved if I can exhibit a topological group Γ, a sequence of continuous automorphisms S_n of Γ, and a sequence α_n of elements of Γ, such that α_n converges to 1, whereas $S_n(\alpha_n)$ does not converge to 1. (For then T_{α_n} converges to I in G whereas $S_n T_{\alpha_n} S_n^{-1} = T_{S_n(\alpha_n)}$ does not.) Such an example is easy to construct. Let Γ be the two-dimensional toral group (that is, the set of all pairs (ξ, η), $0 \leqq \xi, \eta \leqq 1$; the group operation is coordinatewise addition modulo 1). Let S_n be the automorphism defined by the unimodular matrix

$$\begin{pmatrix} n^2 & n+1 \\ n-1 & 0 \end{pmatrix},$$

that is, $S_n(\xi, \eta) = (n^2\xi + (n+1)\eta, (n-1)\xi)$, and $\alpha_n = (0, 1/2(n+1))$. Clearly $\alpha_n \to (0, 0)$ (the identity of Γ) and $S_n(\alpha_n) = (1/2, 0)$.

II. Baire's theorem, as is well known, is often used to give existence proofs. Theorem 6, however, is certainly not presented in the spirit of an existence theorem; in fact the existence of an ergodic automorphism was used in the course of the proof. The purpose of the theorem is rather to give further support (along the lines of the work of Oxtoby and Ulam([22])) to the familiar conjecture that "in general a measure preserving transformation is ergodic." There is, however, no implication between Theorem 6 and the corresponding result of Oxtoby and Ulam: they define a stronger topology and I consider a wider class of transformations.

4. **The metric topology.** There are several other interesting and natural topologies on the group G of automorphisms; in this section I shall discuss one of these. Since this topology is defined by a metric I shall refer to it as the *metric topology* of G, in distinction to the previously discussed neighborhood topology. Using the notations and results of Theorem 1 it is not difficult to define the metric. Given S and T in G, write $E = E(S, T) = \{a: Sa \neq Ta\}$, then the distance between S and T is given by $d(S, T) = |\sup E^*(S, T)|$. In words: the distance between S and T is the measure of the largest element on arbitrarily small subelements of which S and T are different.

([20]) This result is usually stated in terms of the "strong" topology of operators on Hilbert space; the translation from the usual proof to one applicable in the present case is, however, quite trivial. See footnote 13 above, and André Weil, *L'integration dans les groupes topologiques et ses applications*, Paris, 1938, p. 141.

([21]) See Paul R. Halmos, *On automorphisms of compact groups*, Bull. Amer. Math. Soc. vol. 49 (1943) p. 619.

([22]) See Oxtoby and Ulam, op. cit. p. 876.

The lack of immediate intuitive content of this definition is a typical example of the one disadvantage of the Boolean algebra formulation as compared with the more customary "point" formulation. If S and T had been measure preserving transformations on a measure space, I could have defined the distance between S and T to be the measure of the set of points where they differ. Not having any points to talk about I had to adopt the above circumlocution. This process is not as bad, however, as it seems, since in proving "almost everywhere" statements points always have to be ignored and the only effective tools are the algebraic properties of sets of positive measure.

THEOREM 7. *In the metric topology G is a complete topological group and, in fact, the metric $d(S, T)$ is invariant under both left and right translation.*

Proof. (1) The first thing to prove is that $d(S, T)$ satisfies the distance axioms. It is clear that $d(S, T) \geqq 0$ and $E(S, T) = E(T, S)$ implies that $d(S, T) = d(T, S)$. If $d(S, T) = 0$ then $E^*(S, T)$ contains only the element 0. Hence every nonzero $a \in B$ contains a nonzero b no nonzero subelement of which is in E. In other words, every $a \neq 0$ has a nonzero subelement belonging to $(E')_*$. Since E' satisfies the hypotheses of the corollary to Theorem 1, $(E')_*$ is a principal ideal; write $e = \sup (E')_*$. Then for every $a \neq 0$, $ae \neq 0$: this means that $e = 1$, that is, that every a is in E'. Hence E is empty and $S = T$. To prove, finally, the triangle inequality, take $R, S, T \in G$ and write $A = E(R, S)$, $B = E(S, T)$, $C = E(R, T)$. Clearly $(C')_* \supset (A')_* \cap (B')_*$. Writing $a = \sup (A')_*$, $b = \sup (B')_*$, $c = \sup (C')_*$, it follows that $c \supset ab$, so that $c' \subset a' \cup b'$. From the second half of Theorem 1, $a' = \sup A^*$, $b' = \sup B^*$, $c' = \sup C^*$, whence $d(R, T) \leqq d(R, S) + d(S, T)$.

(2) Since $E(S, T) = E(RS, RT)$, it is clear that $d(S, T) = d(RS, RT)$. Moreover $S^{-1}a = T^{-1}a$ $(= b$, say) if and only if $Sb = Tb$ $(= a)$; consequently $E'(S^{-1}, T^{-1})$ is the transform under S (or under T) of $E'(S, T)$. It follows that $d(S^{-1}, T^{-1}) = d(S, T)$. Since in any group a metric which is invariant under two of the three obvious group operations (right translation, left translation, taking inverses) is also invariant under the third, d has the stated invariance properties.

(3) If $\{T_i\}$ is a Cauchy sequence in G, write $E_{ij} = \{a: T_i a \neq T_j a\}$; then $d(T_i, T_j) = |\sup E_{ij}^*| \to 0$. Choose m_k so that for $i, j > m_k$, $|\sup E_{ij}^*| < 1/2^k$; write $n_1 = m_1$, $n_k = \max (n_{k-1} + 1, m_k)$. Define also

$$e_k = \sup (E_{n_k n_{k+1}})^*, \; d_k = e_k \cup e_{k+1} \cup e_{k+2} \cup \cdots, \; c_k = d_k'.$$

Since $c_k \subset e_m'$ for $m \geqq k$, $a \subset c_k$ implies $T_{n_k} a = T_{n_{k+1}} a = \cdots$. Since $|d_k| \leqq |e_k| + |e_{k+1}| + \cdots < 1/2^{k-1}$, $c_k \to 1$, clearly $c_1 \subset c_2 \subset c_3 \subset \cdots$. Consequently a unique automorphism T is defined by the conditions $Ta = T_{n_k} a$ for $a \subset c_k$; clearly $d(T, T_{n_k}) \leqq |d_k| < 1/2^{k-1}$. The Cauchy property of the $\{T_i\}$ implies, as usual, that $T_n \to T$. This completes the proof of Theorem 7.

G is not separable in the metric topology. For, representing B on the unit

interval as usual, let T_α $(0 \leqq \alpha < 1)$ be translation by α (mod 1). Clearly $d(T_\alpha, T_\beta) = 1$ if $\alpha \neq \beta$(23).

The remainder of this section is dedicated to deriving the analogues of the approximation theorem (Theorem 3) and the category theorem (Theorem 6) of the previous section. For this purpose it is necessary to go into some detail concerning "local" properties of automorphisms: this is done in the lemmas below.

For any $T \in G$ write $E_n = E_n(T) = \{a : T^n a = a\}$. According to the corollary to Theorem 1, $(E_n)_*$ is a principal ideal; write $e_n = \sup (E_n)_*$. It will be convenient to use the following terminology: if $a \subset e_n$, *T has period n in a* (in particular if $a \subset e_1$, *T is the identity in a*); if T has some finite period in a, *T is periodic in a*; if T is periodic in 1, *T is periodic*; if T is not periodic in any $b \subset a$, $b \neq 0$, *T is nowhere periodic in a*; if T is nowhere periodic in 1, *T is nowhere periodic.*

LEMMA 1. *If nowhere in a $(a \neq 0)$ does T have a period smaller than n, $n = 1, 2, \cdots$, then there is at least one $b_n \subset a$, $b_n \neq 0$, such that $b_n T^i b_n = 0$ for $0 < i \leqq n-1$.*

Proof. For $n = 1$ the lemma is vacuously true; I proceed by induction. If, for $n \geqq 2$, b_{n-1} has already been found then, since $b_{n-1} \subset a$, T does not have the period $n-1$ in b_{n-1}. Consequently there is a nonzero subelement c of b_{n-1} for which $c \neq T^{n-1}c$. It follows (using the fact that T preserves measure) that $c \cdot T^{n-1}c' \neq 0$; write $b_n = c \cdot T^{n-1}c'$. Since $b_n \subset T^{n-1}c'$ and $T^{n-1}b_n \subset T^{n-1}c$, $b_n \cdot T^{n-1}b_n = 0$; the fact that $b_n \subset b_{n-1}$ implies that $b_n \cdot T^i b_n = 0$ for $0 < i \leqq n-2$.

LEMMA 2. *If T has nowhere a period smaller than n, $n = 2, 3, \cdots$, then there is at least one b, $b \neq 0$, such that $b, Tb, \cdots, T^{n-1}b$ are pairwise disjoint and $1/n \geqq |b| \geqq 1/(2n-1)$.*

Proof. (1) This lemma is a sharpening of Lemma 1 (at least for the case $a = 1$) in that seemingly more is required of b than of b_n, and in that b is quantitatively estimated. Two-thirds of this improvement is trivial. First: if $b \cdot T^i b = 0$ for $i = 1, \cdots, n-1$ then $b, Tb, \cdots, T^{n-1}b$ are pairwise disjoint since for any $i < j$, $T^i b \cdot T^j b = T^i(b \cdot T^{j-i}b) = 0$. Also: if $b, Tb, \cdots, T^{n-1}b$ are pairwise disjoint then $b \cup Tb \cup \cdots \cup T^{n-1}b \subset 1$ shows that $n|b| \leqq 1$. It is only the lower inequality that is not trivial.

(2) Consider the set $E = \{a : a \cdot T^i a = 0, i = 1, \cdots, n-1\}$. Lemma 1 shows that E is not empty. With the ordering $(\subset)$ of B, E is a partially ordered set; let $A \subset E$ be a linearly ordered subset (that is, for $a_1, a_2 \in A$ either $a_1 \subset a_2$ or $a_2 \subset a_1$). Write $a^* = \sup A$; then $T^i a^* = \sup T^i A$; I assert that $a^* \in E$. For, if $a_1 \in A$ and $T^i a_2 \in T^i A$, then either $a_1 \subset a_2$ and $a_2 T^i a_2 = 0$ or $a_2 \subset a_1$ and $a_1 T^i a_1 = 0$ implies $a_1 \cdot T^i a_2 = 0$, so that $\sup A$ and $\sup T^i A$ are indeed disjoint

(23) This example, as well as many other properties of the metric topology, is reminiscent of the "uniform" topology of operators on Hilbert space; the connection between them will become clearer in the next section.

for $i=1, \cdots, n-1$. These considerations prove that the hypothesis, and therefore the conclusion, of Zorn's lemma([24]) is valid; E contains at least one maximal element b.

(3) Using this maximal $b \in E$ write $a = b \cup Tb \cup \cdots \cup T^{n-1}b$. Then for every nonzero $c \subset a'$, $b \cdot T^i c \neq 0$ for at least one $i=1, \cdots, n-1$. For if this were not true, in other words if $0 \neq c \subset a'$, $b \cdot T^i c = 0$ for $i=1, \cdots, n-1$, then by Lemma 1 there is a nonzero subelement $d \subset c$ for which $d \cdot T^i d = 0$, $i=1, \cdots, n-1$. Hence $(b \cup d) \cdot T^i(b \cup d) = b \cdot T^i b \cup d \cdot T^i b \cup b \cdot T^i d \cup d \cdot T^i d = 0$ for $i=1, \cdots, n-1$; but this is impossible since it contradicts the maximality of b. In other words: for every $c \subset a'$, $c(T^{-1}b \cup T^{-2}b \cup \cdots \cup T^{-n+1}b) \neq 0$; consequently $a' \subset T^{-1}b \cup \cdots \cup T^{-n+1}b$ and $|a'| \leqq (n-1)|b|$. Since $1 = b \cup Tb \cup \cdots \cup T^{n-1}b \cup a'$, $1 = n|b| + |a'| \leqq (2n-1)|b|$, q.e.d.

LEMMA 3. *If T is nowhere periodic and ϵ is positive, then there exists an automorphism $S \in G$, an element $a \in B$, and a positive integer n such that*

(i) $Sa = a$,

(ii) *S has the period n in a,*

(iii) $|a| > 1/2$,

(iv) $d(S, T) \leqq \epsilon$.

Proof. Choose n so that $2/n \leqq \epsilon$, and choose b, in accordance with Lemma 2, so that $b \cdot T^i b = 0$ for $i=1, \cdots, n-1$, and $|b| \geqq 1/(2n-1)$. Write $a = b \cup Tb \cup \cdots \cup T^{n-1}b$, as above; then $|a| \geqq n/(2n-1) > 1/2$. For $d \subset b \cup Tb \cup \cdots \cup T^{n-2}b$ define $Sd = Td$; for $d \subset T^{n-1}b$ define $Sd = T^{-n+1}d$. These requirements determine a unique automorphism S of the algebra of subelements of a; it is clear that (i), (ii), and (iii) are satisfied.

Consider $e = a' \cdot T^{-1}a'$ and $Te = Ta' \cdot a'$. Then $a' = e \cup a' \cdot T^{-1}a$ and $a' = Te \cup a' \cdot Ta$, so that $|a' \cdot T^{-1}a| = |a' \cdot Ta|$. Consequently there is an automorphism $R \in G$ for which $R(a' \cdot T^{-1}a) = a' \cdot Ta$. For $d \subset e$ define $Sd = Td$ and for $d \subset a' \cdot T^{-1}a$ define $Sd = Rd$; these requirements (together with the ones already described in the preceding paragraph) determine a unique automorphism $S \in G$. Since $S = T$ on all subelements of $b \cup Tb \cup \cdots \cup T^{n-2}b \cup e$, $d(S, T) \leqq |b| + |a' \cdot T^{-1}a|$. Since $a' \cdot T^{-1}a \subset T^{-1}b$, $d(S, T) \leqq 2|b| \leqq 2/n \leqq \epsilon$, q.e.d.

THEOREM 8. *In the metric topology periodic automorphisms are dense in G*([25]).

Proof. The main idea of this proof is to show that Lemma 3 remains valid even if the hypothesis of nowhere periodicity is removed. Assuming for a moment that this has already been accomplished, I shall show how the theorem follows from it. Given T and ϵ, apply this sharpened form of Lemma 3 to T and $\epsilon/2$, obtaining an automorphism $S_1 \in G$, an element $a_1 \in B$, and an in-

([24]) See Lefschetz, op. cit. p. 5.

([25]) My thanks are due to R. H. Fox for several valuable discussions of this theorem and its proof. Among other things he discovered the original version (which I subsequently modified slightly) of the statement and proof of Lemma 3.

teger $n_1>0$, such that $S_1a_1=a_1$, S_1 has the period n_1 in a_1, $|a_1|>1/2$, and $d(S_1, T)\leqq\epsilon/2$. Next apply the same result to $\epsilon/4$ and S_1 considered as an automorphism of the algebra of subelements of a_1' only, obtaining S_2, $a_2\subset a_1'$, and n_2 such that $S_2a_2=a_2$, S_2 has the period n_2 in a_2, $|a_2|\geqq(1/2)|a_1'|$ and $d(S_1, S_2)\leqq\epsilon/4$. (Since the algebra of subelements of a_1' is not normalized, the form of the third condition had to be modified slightly.) Repeating this process sufficiently often, say k times, a stage is reached where $|a_1'\cdots a_k'|<\epsilon$. A unique automorphism S is defined by the conditions $Sd=S_id$ for $d\subset a_i$, $i=1,\cdots,k$, $Sd=d$ for $d\subset a_1'a_2'\cdots a_k'$. For this S, $d(S, T)<\epsilon/2+\epsilon/4+\cdots+\epsilon/2^k+\epsilon<2\epsilon$, and S has everywhere the period $n_1n_2\cdots n_k$.

In order to obtain the required sharpening of Lemma 3, consider once more the sets $E_n=\{a:T^na=a\}$, and the elements $e_n=\sup(E_n)_*$. Since $TE_n=E_n$, it follows that $Te_n=e_n$ and consequently, writing $e=e_1\cup e_2\cup e_3\cup\cdots$, that $Te=e$. Given $\epsilon<|e|/2$ choose n so that $|e_1\cup\cdots\cup e_n|>|e|-\epsilon/2$; define T_0 by $T_0d=Td$ for $d\subset e_1\cup\cdots\cup e_n$ and for $d\subset e'$, and $T_0d=d$ for $d\subset e\cdot e_1'\cdots e_n'$. Clearly $d(T_0, T)\leqq\epsilon/2$. If $|e_1\cup\cdots\cup e_n|>1/2$ the desired result is already achieved (with S replaced by T_0 and n replaced by $n!$); if $|e_1\cup\cdots\cup e_n|<1/2$, a single application of Lemma 3 to $\epsilon/2$ and T_0 considered on the subelements of e' only, in the way already described above, concludes the proof.

The "size" of the set of ergodic automorphisms is easily determined by means of Theorem 8.

THEOREM 9. *In the metric topology the set of ergodic automorphisms is nowhere dense in G.*

Proof. Let K be any sphere in G; it follows from Theorem 8 that K contains a periodic automorphism S of, say, period n. Choose $\epsilon>0$ so that $\epsilon<1/n$ and so that the sphere K_1 of radius ϵ about S is contained in K. The nowhere denseness follows from the fact that no T in K_1 is ergodic. To prove this choose $T\in K_1$ and consider $E=E(S, T)=\{a: Sa\neq Ta\}$ and $e=\sup E^*$. Then $d(S, T)=|e|<\epsilon<1/n$; if $|e|=0$, $S=T$ and T is obviously not ergodic. If $|e|>0$, write $a=e\cup Se\cup S^2e\cup\cdots\cup S^{n-1}e$. Then $Sa=a$ and $0<|a|<1$. Since $e\subset a$, S and T agree on all subelements of a', so that $Ta'=Sa'=a'$ and $0<|a'|<1$. In other words, every $T\in K_1$ has a nontrivial invariant element and is therefore nonergodic, q.e.d.

5. **The uniform topology.** The metric discussed in the previous section is not perhaps one that most people would consider "natural." The topology it defines is, for one thing, very strong (that is, has many open sets); so strong in fact that it is practically discrete. It is very hard for two automorphisms to be close in this topology: the assertion that any particular set (such as the set of periodic automorphisms) is dense reveals a deep structural property of automorphisms. By the same token, however, it is very easy for any particular set (such as the set of ergodic automorphisms) to be nowhere dense and

any assertion concerning topological smallness (for example, being nowhere dense or having first category) is essentially a property of the topology only. The metric topology is, however, of considerable interest because, as I shall presently show, it coincides with one of the most natural topologies of G and is much easier to handle analytically.

This natural topology is also defined by a metric; by analogy with the space of bounded operators on Hilbert space I shall call it the *uniform topology*. The distance between two elements S and T of G is defined by

$$\delta(S, T) = \sup \{ |Sa - Ta| : a \in B \}.$$

The elementary properties of δ are so easy to see that it is not worthwhile to state them in a formal theorem. That δ satisfies the distance axioms is a well known fact in general metric spaces. Since $|RSa - RTa| = |Sa - Ta|$, it is clear that δ is invariant under left translation in G. Also, as a runs through B, so does $T^{-1}a$, and

$$|S^{-1}a - T^{-1}a| = |S(S^{-1}a - T^{-1}a)| = |a - S(T^{-1}a)| = |T(T^{-1}a) - S(T^{-1}a)|;$$

it follows that $\delta(S^{-1}, T^{-1}) = \delta(S, T)$, and consequently that δ is invariant under all the group operations. The interesting fact concerning δ, namely the coincidence of the uniform and the metric topologies of G, is a consequence of the following theorem.

THEOREM 10. $(2/3)d(S, T) \leqq \delta(S, T) \leqq d(S, T)$.

Proof. Since both d and δ are invariant under the group operations it is sufficient to prove the theorem for the case $S = I$. To prove the upper inequality, suppose that $d(I, T) = \alpha$; then there is an element $a \in B$ of measure α such that $Ta = a$ and for all $b \subset a'$, $Tb = b$. Hence, for any b,

$$\begin{aligned} |Tb - b| &\leqq |T(ab) - (ab)| + |T(a'b) - (a'b)| = |T(ab) - (ab)| \\ &= |a(Tb - b)| \leqq |a| = \alpha = d(I, T). \end{aligned}$$

It follows that $\delta(I, T) = \sup \{ |Tb - b| : b \in B \} \leqq d(I, T)$. This inequality is rather natural. It shows that every δ-sphere contains the corresponding d-sphere of the same radius; in other words that the metric topology is stronger (has more open sets) than the uniform topology. This is in accordance with my earlier remarks to the effect that no reasonable topology ought to be stronger than the metric topology.

To prove the lower inequality it is necessary to examine in a little more detail the way in which an arbitrary automorphism T is made up of periodic pieces. Consider, as in the proof of Theorem 8, the elements

$$e_k = \sup (\{a : T^k a = a\}_*), \qquad k = 1, 2, \cdots.$$

Write $\bar{e}_1 = e_1$, $\bar{e}_k = e_k e'_{k-1} e'_{k-2} \cdots e'_2 e'_1$, and $\bar{e}_0 = e'_1 e'_2 e'_3 \cdots = \bar{e}'_1 \bar{e}'_2 \bar{e}'_3 \cdots$.

(Observe that, according to Theorem 1, $|\bar{e}_1'| = d(I, T)$.) Since $Te_k = e_k$ for $k=1, 2, \cdots$, it follows that $T\bar{e}_k = \bar{e}_k$ for $k=0, 1, 2, \cdots$; clearly $1 = \bar{e}_0 \cup \bar{e}_1 \cup \bar{e}_2 \cup \cdots$. T is nowhere periodic in $\bar{e}_0$, T has period k in $\bar{e}_k$ for $k=1, 2, \cdots$, and, in $\bar{e}_k$, T has nowhere a period smaller than k. Hence Lemma 1 may be applied to each nonzero $\bar{e}_k$, with a, n replaced by $\bar{e}_k$, k for $k=1, 2, \cdots$, and a, n replaced by $\bar{e}_0$, 2 for $k=0$. The conclusion is that for each k there is an element $b_k \subset \bar{e}_k$ ($b_k \neq 0$ if $\bar{e}_k \neq 0$) such that $b_k \cdot T^i b_k = 0$ for $0 < i \leqq k-1$, $k=1, 2, \cdots$, and $b_0 \cdot Tb_0 = 0$ for $k=0$. Lemma 2, applied to the (unnormalized) algebra of all subelements of $\bar{e}_0$, shows that b_0 may be chosen so that $|b_0| \geqq (1/3)|\bar{e}_0|$. (In case $\bar{e}_0 = 0$ this is, of course, trivial.) For larger values of k the corresponding estimate for b_k is not good enough for my purpose; I shall show, in fact, that for $k=1, 2, \cdots$ b_k can be chosen so that $|b_k| = |\bar{e}_k|/k$. To see this it is necessary only to apply Zorn's lemma, exactly as in step (2) of the proof of Lemma 2. This procedure shows that the b_k may be assumed maximal. If now $|b_k| < |\bar{e}_k|/k$, then Lemma 1 may be applied to $\bar{e}_k(b_k' \cdot Tb_k' \cdots T^{k-1}b_k')$, leading to a contradiction of the maximality of b_k.

Now write $a_0 = b_0$, $a_1 = b_1$ $(= \bar{e}_1)$; $a_{2k} = \bigcup_{i=0}^{k-1} T^{2i} b_{2k}$, $a_{2k+1} = \bigcup_{i=0}^{k-1} T^{2i} b_{2k+1}$, $k=1, 2, \cdots$. Then, for all $k \neq 1$, $a_k \cdot Ta_k = 0$. Moreover for $k=1, 2, \cdots$,

$$|a_{2k}| = k|b_{2k}| = (k/2k)|\bar{e}_{2k}| = |\bar{e}_{2k}|/2,$$

$$|a_{2k+1}| = k|b_{2k+1}| = (k/(2k+1))|\bar{e}_{2k+1}| \geqq |\bar{e}_{2k+1}|/3;$$

consequently for all $k=0, 1, 2, \cdots$, $|a_k| \geqq |\bar{e}_k|/3$.

The desired inequality between d and δ is a consequence of these last written estimates. Writing $a = a_0 \cup a_2 \cup a_3 \cup a_4 \cup \cdots$, it follows that $a \cdot Ta = 0$ and $|a| = |a_0| + |a_2| + |a_3| + \cdots \geqq (|\bar{e}_0| + |\bar{e}_2| + |\bar{e}_3| + \cdots)/3 \geqq |\bar{e}_1'|/3$. Consequently $\delta(I, T) \geqq |Ta - a| \geqq 2|\bar{e}_1'|/3 = 2d(I, T)/3$, q.e.d.

The translation of the unit interval by 1/3, mod 1, shows that both bounds in Theorem 10 are best possible.

In order for S and T to be close in the uniform (and hence in the metric) topology $|Sa - Ta|$ has to be uniformly small, whereas closeness in the neighborhood topology requires only that $|Sa - Ta|$ be small for a finite number of a. Clearly then the uniform topology is no weaker than the neighborhood topology: that they do not coincide, and consequently that the uniform topology is really stronger, follows from the already demonstrated topological differences between them (such as separability, density of ergodic automorphisms, invariant metrizability, and so on).

SYRACUSE UNIVERSITY,
SYRACUSE, N. Y.