

POINCARÉ RECURRENCE AND NUMBER THEORY

BY HARRY FURSTENBERG

Introduction. Poincaré is largely responsible for the transformation of celestial mechanics from the study of individual solutions of differential equations to the global analysis of phase space. A system of differential equations such as those which embody the laws of Newtonian mechanics generates a one-parameter group of transformation of the manifold that represents the set of states of a dynamical system. The evolution of the dynamical system in time corresponds to a particular solution of the system of differential equations; it also corresponds to an orbit of the group of transformations acting on a single state. The efforts of the classical analysts in celestial mechanics had been directed to extracting by analytical means as much information as possible about the individual solutions to the system of differential equations. Poincaré's work gave impetus to a global approach which studies the totality of solutions and shifts attention to the transformation group of phase space.

Two of Poincaré's achievements which can be traced to this point of view are his theory of periodic solutions and his recurrence theorem. In the former of these, the topological nature of the phase space plays a key role; Poincaré showed how an essentially topological analysis of orbits, under certain conditions, can be used to establish the existence of a periodic solution curve. In his recurrence theorem, Poincaré demonstrated how measure-theoretic ideas, particularly the idea of a measure-preserving group of transformations, lead to the existence of numerous "approximately periodic"—or, "recurrent"—solution curves. The impact of these ideas is felt today in the establishment of two new disciplines: topological dynamics and ergodic theory. In topological dynamics one abstracts from the classical setup the topological space representing the totality of states of a dynamical system, together with the group of homeomorphisms corresponding to the evolution of the system from its position at time 0 to its position at time t . For ergodic theory, the phase space is replaced by an abstract measure space and the "dynamics" come from the action of a group of measure-preserving transformations of the measure space.

Initially these abstract settings for dynamical theory were contemplated in order to shed light on classical dynamics by focusing on those aspects of dynamical systems that were pertinent to the phenomena being studied. At the same time, however, the scope of dynamics was considerably broadened by the generality of the new theory, and the groundwork was laid for

Presented to the Symposium on the Mathematical Heritage of Henri Poincaré, April 7–10, 1980; received by the editors April 15, 1981.

1980 *Mathematics Subject Classification*. Primary 05A05, 28D05, 58F11; Secondary 10A99.

© 1981 American Mathematical Society
0002-9904/81/0000-0500/\$07.00

applying similar ideas to areas that, on the face of it, are quite unrelated to dynamics.

We shall be interested in two particular dynamical theorems and their ramifications when taken in the broadest possible context. One of these is the Poincaré recurrence theorem alluded to previously; the other is a topological analogue due to Birkhoff. Choosing a nonconventional model of a dynamical system rather than a classical model, we will obtain results of interest in number theory. For example we will see that van der Waerden's theorem on arithmetic progressions is a consequence of an appropriate generalization of Birkhoff's recurrence theorem. A more recent result is that of Szemerédi stating that a subset of the integers having positive upper density contains arbitrarily long arithmetic progressions. This will be seen to relate in a natural way to an extension of Poincaré's recurrence theorem.

For a comprehensive treatment of the results described here the reader is referred to [3]–[5].

1. Dynamical systems and measure-preserving systems. A dynamical system evolving in time is described by a one-parameter group $\{T_t, -\infty < t < \infty\}$, $T_{t+s} = T_t \circ T_s$, acting on a space X . Since we are interested primarily in asymptotic behavior as $t \rightarrow \infty$, the dynamical aspects are usually already reflected in the action of the subgroup $\{T_n, n \in \mathbf{Z}\}$, $\mathbf{Z} = \text{integers}$. For our purposes, then, a *dynamical system* will consist of a space X on which a one-one transformation T acts, thereby generating a group $\{T^n, n \in \mathbf{Z}\}$ of transformations.

The space X will be either a topological space (in our discussion, a metric space) in which case we require T to be a homeomorphism, or a measure space in which case T will be required to be measure preserving. The central phenomenon to be studied is that of *recurrence*. For this phenomenon to occur one must assume some kind of boundedness of X . When X is a topological space, the appropriate assumption is that X is compact. When X is a measure space, the boundedness of X is expressed by requiring X to have finite measure.

Let us then define formally a *dynamical system* as a pair (X, T) , X being a compact metric space, and T a homeomorphism of X . A *measure-preserving system* will be a quadruple $(X, \mathfrak{B}, \mu, T)$ where X is an abstract space, \mathfrak{B} is a σ -algebra of subsets of X , μ is a probability measure in \mathfrak{B} , and T is a measure-preserving transformation of (X, \mathfrak{B}, μ) . These last two conditions mean that $\mu(A) \geq 0$ for $A \in \mathfrak{B}$, with $\mu(X) = 1$, and for $A \in \mathfrak{B}$, $T^{-1}A \in \mathfrak{B}$ with $\mu(T^{-1}A) = \mu(A)$.

In the topological context recurrence is a property of individual points of the space X . We shall be studying several versions of this notion, the most basic of which we shall simply call *recurrence*. We say a point $x \in X$ is a *recurrent point* of (X, T) if for some sequence $n_k \rightarrow \infty$, $T^{n_k}x \rightarrow x$. Equivalently we could say that a point x is recurrent if for each neighborhood V of x , $T^n x \in V$ for some $n > 0$.

Closely related to this is the notion of recurrence implicit in Poincaré's recurrence theorem.

THEOREM 1.1. *Let $(X, \mathfrak{B}, \mu, T)$ be a measure-preserving system, and let $V \in \mathfrak{B}$ with $\mu(V) > 0$. There exists some point $x \in V$ with $T^n x \in V$ for some $n > 0$.*

The proof is extremely simple. Assume no point $x \in V$ returned to V . Then $T^{-n}V \cap V = \emptyset$ for all $n > 0$, and so $T^{-n}V \cap T^{-m}V = \emptyset$ whenever $n \neq m$. But the sets $T^{-m}V$ all have the same measure $\mu(V) > 0$, and they cannot be disjoint since $\mu(\bigcup_{n=1}^{\infty} T^{-n}V) \leq \mu(X) = 1$.

Actually Theorem 1.1 implies the stronger statement that almost every point (excepting a set of measure 0) $x \in V$ returns sometimes to V . For this set of points is, in any case, measurable (i.e., in \mathfrak{B}) and we can consider the subset $V' \subset V$ of points that never return to V . *A fortiori* they never return to V' and so, by Theorem 1.1, $\mu(V') = 0$.

Now take X to be a separable metric space as well as a measure space. Cover X by countably many balls of radius $\varepsilon/2$, and apply the foregoing to each ball. We conclude that almost every point X returns to within ε of itself. Since $\varepsilon > 0$ is arbitrary, we conclude that *almost every point of X is recurrent*.

It can be shown that any (compact) dynamical system (X, T) possesses some measure on Borel sets which is preserved by T . Poincaré's theorem then implies that *for any dynamical system (X, T) there exist points $x \in X$ which are recurrent*. It is Birkhoff who initiated the study of dynamical systems in the context of metric spaces [2], and he gave a purely topological proof of this statement. Let us present a streamlined proof of this result which we call the *Birkhoff recurrence theorem*.

THEOREM 1.2. *If X is a compact metric space, T a continuous map of X to itself, then there exists some point $x \in X$ with $T^{n_k} x \rightarrow x$ for some sequence $n_k \rightarrow \infty$.*

PROOF. Consider the family of closed sets $Y \subset X$ satisfying $TY \subset Y$. It is easy to see that Zorn's lemma applies to this family (on account of compactness) so that there exists a minimal closed invariant set. Let Y_0 be such a set and let $y \in Y_0$. Then $Y_1 = \overline{\{T^n y, n = 1, 2, 3, \dots\}}$ is again a closed invariant set contained in Y_0 . By minimality $Y_1 = Y_0$; hence $y \in \overline{\{T^n y, n = 1, 2, 3, \dots\}}$. Thus y is a recurrent point.

Actually Birkhoff proved more than this. In fact Birkhoff uses the term recurrence for a stronger version of this notion that we shall call "uniform recurrence". Birkhoff's argument shows that every compact dynamical system possesses uniformly recurrent points.

DEFINITION. A point $x \in X$ is *uniformly recurrent* for the dynamical system (X, T) if, for any neighborhood V of x , there is a length $L < \infty$, such that in any interval (a, b) with $b - a > L$ there is an integer $n \in (a, b)$ with $T^n x \in V$.

For a uniformly recurrent point, any sufficiently long segment of its orbit comes arbitrarily close to the point. There is a close connection between uniform recurrence and the notion of a minimal dynamical system.

DEFINITION. A system (X, T) is *minimal* if there does not exist a nonempty closed T -invariant subset $Y \subsetneq X$.

Equivalently, a system (X, T) is minimal if every (forward) orbit $\{T^n x, n \geq 0\}, x \in X$, is dense in X . For if (X, T) is minimal, the closed T -invariant set $\overline{\{T^n x, n \geq 0\}}$ must coincide with X . Conversely, if every forward orbit is dense, so is every T -invariant set, and the system is minimal.

THEOREM 1.3. *If (X, T) is minimal, then each $x \in X$ is uniformly recurrent for (X, T) . Conversely, if a point $x \in X$ is uniformly recurrent, then x belongs to a closed T -invariant set $Y \subset X$ for which the subsystem (Y, T) is minimal.*

According to the second part of the theorem, if every point of X is uniformly recurrent, then X is the union of *minimal sets*, i.e., minimal closed T -invariant subsets of X . Note that by definition, two different minimal sets are necessarily disjoint.

PROOF OF THE THEOREM. Assume (X, T) is minimal and let V be any open set in X . The set $\bigcup_{n=0}^{\infty} T^{-n}V$ is an open set whose complement is a T -invariant closed set. Since it is nonempty, we must have $\bigcup_{n=0}^{\infty} T^{-n}V = X$. Since X is compact, we will already have $\bigcup_{n=0}^L T^{-n}V = X$ for some L . Hence for any $x \in X$ and any n , one of the points $T^n x, T^{n+1}x, \dots, T^{n+L}x$ belongs to V . This shows that x is uniformly recurrent.

Conversely, suppose that $x \in X$ is uniformly recurrent. We wish to show that x is contained in a minimal set of X . Suppose Z is a T -invariant closed subset of the forward orbit closure of x , $Y = \{T^n x, n \geq 0\}$. If $x \in Z$ then Z coincides with Y . If $x \notin Z$, then $x \in V = X \setminus Z$. By uniform recurrence, for any n , one of the points $T^n x, T^{n+1}x, \dots, T^{n+L}x$ belongs to V . For any $z \in Z$ there are points $T^n x$ arbitrarily close to z . It follows that one of the points $z, Tz, \dots, T^L z$ belongs to V .

But this is a contradiction since these points are all in Z and $Z \cap V = \emptyset$. It follows that $Z = Y$ and so Y is minimal. This completes the proof.

The proof of the theorem shows that in the case of a minimal system, if V is an open set, there is an L so that for each $x \in X$, one of the points $x, Tx, \dots, T^L x$ meets V . It follows that for any $\epsilon > 0$ there is some M so that each orbit segment $x, Tx, \dots, T^M x$ comes within ϵ of each point of the space. Using now the second assertion of the theorem, this applies in particular to the orbit closure of a uniformly recurrent point. This gives us the following characterization of uniform recurrence.

COROLLARY. *A point is uniformly recurrent iff every sufficiently long segment of its orbit comes arbitrarily close to every point in its orbit.*

The proof of Theorem 1.2 shows that any system (X, T) with X compact possesses a minimal subsystem. On the basis of the foregoing theorems we now conclude:

THEOREM 1.4. *If X is a compact metric space, T a homeomorphism of X , there exists a point $x \in X$ which is uniformly recurrent for (X, T) .*

2. Symbolic systems. In the course of our discussion we shall extend and refine the recurrence theorems of the preceding section. To obtain results of a number-theoretical nature we shall apply these theorems to a particular kind

of nonconventional dynamical system—a *symbolic system*. Let $\Lambda = \{a, b, c, \dots\}$ be a finite set and form the space

$\Omega = \Lambda^{\mathbf{Z}}$ of all sequences with entries from Λ :

$$x \in \Omega \Leftrightarrow x = \{\dots, x(-2), x(-1), x(0), x(1), x(2), \dots\}.$$

Ω can be made into a compact metric space, taking as metric, for example,

$$(2.1) \quad d(x, x') = \inf \left\{ \frac{1}{k+1} \mid x(i) = x'(i) \text{ for } |i| < k \right\}.$$

Here $d(x, x') = 1$ if $x(0) \neq x'(0)$, and $d(x, x') < 1$ otherwise. We define the *shift* homeomorphism $T: \Omega \rightarrow \Omega$ by $Tx(n) = x(n+1)$. If X is any closed T -invariant subset of Ω we call (X, T) a *symbolic dynamical system*.

Given a particular function $\xi: \mathbf{Z} \rightarrow \Lambda$ so that ξ is a point in Ω , we form the smallest shift invariant closed set $X \subset \Omega$ containing the point ξ . The idea which we shall use repeatedly is that the dynamic properties of (X, T) reflect certain aspects of the behavior of ξ . More specifically, recurrence properties for X will imply the existence of certain patterns in the function $\xi(n)$.

To illustrate the idea, consider the implication of the following extension of Theorem 1.2 which will be proved in §6.

THEOREM 2.1. *Let X be a compact metric space, and let $T: X \rightarrow X$ be a continuous map. Then for any integer $l \geq 1$, there exists a point $x \in X$ and a sequence $n_k \rightarrow \infty$ with $T^{n_k}x \rightarrow x$, $T^{2n_k}x \rightarrow x$, \dots , $T^{ln_k}x \rightarrow x$.*

Let (X, T) be any symbolic dynamical system, $X \subset \Omega = \Lambda^{\mathbf{Z}}$. According to Theorem 2.1, there exists $x \in X$ and an $n > 0$ with the points $x, T^n x, T^{2n} x, \dots, T^{ln} x$ within distance < 1 of one another. Now for two points of Ω , $d(x, x') < 1$ implies $x(0) = x'(0)$. It follows that $x(0) = T^n x(0) = T^{2n} x(0) = \dots = T^{ln} x(0)$. Bearing in mind that $T^k x(0) = x(k)$ we have

PROPOSITION 2.2. *Let Λ be a finite set. If X is any shift invariant closed subset of $\Omega = \Lambda^{\mathbf{Z}}$ and $l \geq 1$, there exists a point $x \in X$ and an $n \geq 1$ with $x(0) = x(n) = x(2n) = \dots = x(ln)$.*

If we take $\Lambda = \{1, 2, \dots, r\}$, then a point $\xi \in \Lambda^{\mathbf{Z}}$ corresponds to a partition of \mathbf{Z} into r sets, $\mathbf{Z} = \cup C_i$, where $C_i = \{n: \xi(n) = i\}$. Let $X \subset \Omega$ be the closure of the set of all translates of $\xi: X = \overline{\{T^n \xi, n \in \mathbf{Z}\}}$. Apply Proposition 2.2 to X , and we find there is an $x \in X$ with $x(0) = x(n) = \dots = x(ln)$. Since $X = \overline{\{T^n \xi, n \in \mathbf{Z}\}}$, in view of the definition of the metric on Ω , we can find m so that $T^n \xi$ and x agree on the interval $(-ln, ln)$. It then follows that $\xi(m) = \xi(m+n) = \xi(m+2n) = \dots = \xi(m+ln)$. Say the common value is j . Then C_j contains the arithmetic progression $m, m+n, m+2n, \dots, m+ln$. We have thus established van der Waerden's theorem on arithmetic progressions.

THEOREM 2.3. *Let $\mathbf{Z} = \cup_{i=1}^r C_i$ be a partition of the integers into r subsets. Then one of the sets C_j contains an arithmetic progression of length $l+1$.*

Our next example shows how measure-theoretic ideas can be used. The following refinement of Poincaré's recurrence theorem will be proved in the next section.

THEOREM 2.4. *Let $(X, \mathfrak{B}, \mu, T)$ be a measure-preserving system, and let $A \in \mathfrak{B}$ with $\mu(A) > 0$. Then there exists an integer of the form $n = m^2$ such that $\mu(A \cap T^{-n}A) > 0$.*

Let us now return to our symbolic dynamical system, taking this time $\Lambda = \{0, 1\}$. Suppose that $X \subset \Lambda^{\mathbb{Z}}$ is T -invariant and closed and let $A_1 = \{x \in X; x(0) = 1\}$. Assume now that we can find a T -invariant Borel measure μ on X for which $\mu(A_1) > 0$. We can then form the measure-preserving system $(X, \mathfrak{B}, \mu, T)$, with \mathfrak{B} the Borel sets of X and T the shift. Applying Theorem 2.4 to the system with $A = A_1$, we conclude that there exists some point $x \in A_1 \cap T^{-m^2}A_1$. Then $x(0) = 1$ and $x(m^2) = T^{m^2}x(0) = 1$.

Now let $S \subset \mathbb{Z}$ be a subset of *positive upper density*. For the present purposes this will be taken to mean that there is a sequence of intervals $[a_k, b_k] \subset \mathbb{Z}$ with $b_k - a_k \rightarrow \infty$ and

$$\frac{|S \cap [a_k, b_k]|}{b_k - a_k + 1} > \delta$$

for some $\delta > 0$ and all k . (For any finite set Q , we denote by $|Q|$ the number of elements in Q .) Now set $\xi(n) = 1$ for $n \in S$ and $\xi(n) = 0$. Again let $X = \{T^n \xi, n \in \mathbb{Z}\}$. We claim that we can indeed find some T -invariant measure on X with $\mu(A_1) > 0$.

LEMMA 2.5. *Let $X \subset \{0, 1\}^{\mathbb{Z}}$ contain a point ξ satisfying $\xi(n) = 1$ for a set of n of positive upper density. Then there exists a T -invariant probability measure μ on the Borel sets of X satisfying $\mu(A_1) > 0$.*

PROOF. A Borel measure μ on the compact metric space X determines a linear functional on $C(X)$, the continuous functions in X , by

$$L(f) = \int f d\mu.$$

If μ is a probability measure then L satisfies

- (i) $L(f) \geq 0$ for $f \geq 0$,
- (ii) $L(1) = 1$.

Conversely, given L satisfying these conditions there corresponds to L a probability measure on Borel sets of X . If μ is a T -invariant Borel measure then the corresponding functional satisfies

- (iii) $L(f \circ T) = L(f)$.

Conversely, if L satisfies (i)–(iii) it defines a T -invariant measure μ .

To obtain a linear functional on $C(X)$ satisfying (i)–(iii) we take the point $\xi \in X$ and set

$$L_k(f) = \frac{1}{b_k - a_k + 1} \sum_{n=a_k}^{b_k} f(T^n \xi).$$

Let $\{f_1, f_2, \dots, f_n, \dots\}$ be a countable dense set of functions in $C(X)$. Since $\{L_k(f_l), k = 1, 2, 3, \dots\}$ is bounded for each l , we can choose a subsequence

that converges. By a diagonal procedure we can choose a subsequence $\{k_n\}$ so that

$$\lim_{n \rightarrow \infty} L_{k_n}(f_l)$$

exists for each l . Finally, since $\{f_l\}$ is dense in $C(X)$, we find that the corresponding limit exists for each $f \in C(X)$. Thus replacing $\{a_k, b_k\}$ by a subsequence we may assume that

$$L(f) = \lim_{k \rightarrow \infty} \frac{1}{b_k - a_k + 1} \sum_{n=a_k}^{b_k} f(T^n \xi)$$

exists for all $f \in C(X)$. It is readily verified that L satisfies (i)–(iii).

We next put to use the fact that $\xi(n) = 1$ on a set of integers having positive upper density. The function

$$\varphi(A) = x(0)$$

is a continuous function on X . We have

$$\begin{aligned} L(\varphi) &= \lim_{k \rightarrow \infty} \frac{1}{b_k - a_k + 1} \sum_{n=a_k}^{b_k} T^n \xi(0) \\ &= \lim \frac{|S \cap [a_k, b_k]|}{b_k - a_k + 1} \geq \delta > 0. \end{aligned}$$

Now if L corresponds to the measure μ on X , then

$$\mu(A_1) = \int \varphi \, d\mu = L(\varphi) > 0.$$

This proves the lemma.

The converse of the lemma is also true. In fact, by the ergodic theorem,

$$\lim_{N \rightarrow \infty} \frac{1}{N + 1} \sum \varphi(T^n \eta) = g(\eta)$$

exists for almost all $\eta \in X$ with respect to μ , and $\int g(\eta) d\mu(\eta) = \int \varphi \, d\mu = \mu(A_1)$. So $g(\eta) > 0$ for some $\eta \in X$ and so $\eta(n) = 1$ on a set of positive density, where now the limit

$$\lim_{N \rightarrow \infty} \frac{|S \cap [0, N]|}{N + 1}$$

exists and is positive.

Note the following corollary of the proof of Lemma 2.5.

COROLLARY. *If X is a compact metric space and T is a continuous map of X to itself, there exists some probability measure μ on Borel sets of X which is invariant with respect to T .*

Combining Lemma 2.5 and Theorem 2.4 we obtain the following.

THEOREM 2.6. *Let S be a subset of \mathbf{Z} having positive upper density. Then there exist $s, t \in S$ with $s - t = m^2$, m an integer.*

PROOF. As before let $\xi(n) = 1$ for $n \in S$, $\xi(n) = 0$ otherwise, and let $X = \overline{\{T^n, n \in \mathbf{Z}\}} \subset \{0, 1\}^{\mathbf{Z}}$. Let $A_1 = \{x \in X: x(0) = 1\}$ and let μ be a

T -invariant measure on X with $\mu(A_1) > 0$. Apply Theorem 2.4 to $(X, \mathfrak{B}, \mu, T)$ and the set $A_1 \subset X$. If $\mu(A_1 \cap T^{-n}A_1) > 0$ with $n = m^2$, then, in particular, $A_1 \cap T^{-n}A_1 \neq \emptyset$, and so there exists $\eta \in A_1$ with $T^{m^2}\eta \in A_1$. In other words there is a point $\eta \in X$ with $\eta(0) = \eta(m^2) = 1$. Now η is in the orbit closure of ξ , and we can find a neighborhood of η so small that any point in that neighborhood will have the same coordinates as η in the interval between $-m^2$ and m . It follows that for some s ,

$$T^s\xi(0) = T^s\xi(m^2) = 1$$

or

$$\xi(s) = \xi(s + m^2) = 1.$$

Hence $s \in S$ and $s + m^2 \in S$. This concludes the proof.

The next two theorems describe in a more general manner how recurrence results are used to prove the existence of patterns in subsets of integers. These theorems extend Theorems 2.3 and 2.6 respectively, and the proofs are entirely analogous. In the formulation of these two theorems the term *configuration* will be synonymous with a finite subset of \mathbf{Z} , except that two sets that are translates of one another are said to define the same configuration.

THEOREM 2.7. *Let Q be a family of configurations having the property that for any compact metric dynamical system (X, T) and any $\epsilon > 0$, there exists a point $x_0 \in X$ and a configuration $(n_1, n_2, \dots, n_l) \in Q$ such that $T^{n_1}x_0, T^{n_2}x_0, \dots, T^{n_l}x_0$ are all within ϵ of one another. Then if $\mathbf{Z} = C_1 \cup C_2 \cup \dots \cup C_r$ is any partition of \mathbf{Z} into finitely many sets, one of the subsets C_j contains some configuration in Q .*

In particular the hypothesis is fulfilled when Q consists of arithmetic progressions of a given length, according to Theorem 2.1. Theorem 2.3 is thus a special case of the foregoing. It should be remarked that a slight modification of the argument in proving Theorem 2.7 shows that the same result is true for partitions of the natural numbers $\mathbf{N} = C_1 \cup C_2 \cup \dots \cup C_r$.

THEOREM 2.8. *Let Q be a family of configurations having the property that for any measure-preserving system $(X, \mathfrak{B}, \mu, T)$ and any set $A \in \mathfrak{B}$ with $\mu(A) > 0$, there exists a subset $A' \in \mathfrak{B}$ with $\mu(A') > 0$ and a configuration $(n_1, n_2, \dots, n_l) \in Q$ such that $T^{n_1}A' \subset A, \dots, T^{n_l}A' \subset A$. Then if S is any subset of \mathbf{Z} of positive upper density, S necessarily contains some configuration in Q .*

Theorem 2.6 is the special case of this where $Q = \{(0, m^2), m \in \mathbf{Z}\}$. In §7 we will discuss the applicability of Theorem 2.8 to the case $Q =$ arithmetic progression of length l .

We remark that if Q is a family of configurations satisfying the condition of Theorem 2.8, it also satisfies the condition of Theorem 2.7. For by the corollary to Lemma 2.5, if (X, T) is a compact metric dynamical system, we can endow X with a T -invariant probability measure μ and thereby obtain a measure-preserving system. For any $\epsilon > 0$ we can find a set of diameter ϵ

having positive measure. Let A be such a set and let (n_1, n_2, \dots, n_l) be a configuration in Q . If A' satisfies $T^{n_1}A' \subset A, T^{n_2}A' \subset A, \dots, T^{n_l}A' \subset A$, then for any $x_0 \in A'$, the points $T^{n_1}x_0, T^{n_2}x_0, \dots, T^{n_l}x_0$ are within ε of one another.

We can also see that the conclusion of Theorem 2.8 is stronger than that of Theorem 2.7. For if $Z = \cup C_i$ is a finite partition, some C_j has positive upper density. So any configuration that occurs in any set of positive upper density also occurs in some subset of any finite partition of Z .

3. Sets of recurrence. The simplest nontrivial configurations that one can look for in subsets of Z are two-point configurations. These may be assumed to have the form $(0, n)$. In this section we shall inquire which families Q of configurations $\{(0, n)\}$ satisfy the condition of Theorem 2.8. In other words, for which sets $R \subset \mathbb{N}$ can we assert that whenever $(X, \mathfrak{B}, \mu, T)$ is a measure-preserving system, A a subset in \mathfrak{B} with $\mu(A) > 0$, then there exists $A' \subset A, \mu(A') > 0$, with $T^rA' \subset A$ for some $r \in R$? A set R with this property will be called a *set of recurrence*. Our considerations in this section are measure-theoretic.

We shall say that a set $R \subset \mathbb{N}$ is an *infinite difference set* if R consists of all differences $s_j - s_i, i < j$ where $\{s_1 < s_2 < \dots < s_n < \dots\}$ is some sequence in \mathbb{N} . Our first general result is the following.

THEOREM 3.1. *Every infinite difference set is a set of recurrence.*

The proof is identical to the proof of Poincaré's recurrence theorem. Let $R = \{s_j - s_i, i < j\}$ and let $(X, \mathfrak{B}, \mu, T)$ be a measure-preserving system, $A \in \mathfrak{B}$ with $\mu(A) > 0$. If $\mu(T^{-(s_j - s_i)}A \cap A) = 0$ for all $i < j$. Then

$$\mu(T^{-s_j}A \cap T^{-s_i}A) = 0,$$

and the sets $T^{-s_i}A$ are essentially disjoint sets all having the same positive measure $\mu(A)$. So we must have $\mu(T^{-(s_j - s_i)}A \cap A) > 0$ for some $i < j$ and we can take $A' = T^{-(s_j - s_i)}A \cap A$.

Choosing the sequence $\{s_n\}$ rapidly increasing shows that sets of recurrence can be extremely sparse. On the other hand a set of recurrence cannot be too sparse.

THEOREM 3.2. *If $R = \{r_1 < r_2 < \dots < r_n < \dots\}$ is lacunary, i.e., if $r_{n+1}/r_n > q > 1$, then R is not a set of recurrence.*

Before proving Theorem 3.2 we prove the following general lemma.

LEMMA 3.3. *If $R_1, R_2 \subset \mathbb{N}$ and neither R_1 and R_2 is a set of recurrence, then $R = R_1 \cup R_2$ is not a set of recurrence.*

PROOF. Since R_1 and R_2 are not sets of recurrence we can find systems $(X_i, \mathfrak{B}_i, \mu_i, T_i)$ and sets $A_i \in \mathfrak{B}_i$ with $\mu_i(A_i) > 0$ satisfying $\mu_i(A_i \cap T_i^{-r}A_i) = 0$ for $r \in R_i, i = 1, 2$. Form $X = X_1 \times X_2, \mathfrak{B} = \mathfrak{B}_1 \times \mathfrak{B}_2, \mu = \mu_1 \times \mu_2$ and let $T = T_1 \times T_2$. It is now clear that

$$\mu((A_1 \times A_2) \cap T^{-r}(A_1 \times A_2)) = 0$$

for $r \in R$. Thus R is not a set of recurrence.

On the basis of this lemma it suffices to establish Theorem 3.2 for arbitrarily large q .

LEMMA 3.4. *If $R = \{r_1 < r_2 < \dots < r_n < \dots\}$ satisfies $r_{n+1}/r_n \geq q > 4$, then R is not a set of recurrence.*

PROOF. We shall determine a number α so that the measure-preserving system consisting of translation by α in the group \mathbf{R}/\mathbf{Z} of reals modulo 1 provides a counterexample to the recurrence of R . Consider the sets $\Lambda_n \subset \mathbf{R}/\mathbf{Z}$ where $\Lambda_n = \{t: r_n t \in \Delta\}$ and Δ consists of numbers whose fractional part is between $1/3$ and $2/3$. Λ_n consists of intervals of length $(3r_n)^{-1}$ repeated periodically with period r_n^{-1} . Since $r_{n+1} > 4r_n$ it follows that each Λ_n interval contains some Λ_{n+1} interval. Hence $\cap \Lambda_n \neq \emptyset$. Take $\alpha \in \cap \Lambda_n$. We let $X = \mathbf{R}/\mathbf{Z}$ with Lebesgue measure and let T be a translation by α . Finally let A be any interval of length $< 1/6$. $T^n A = A + r_n \alpha$, and since the fractional part of the $r_n \alpha$ is between $1/3$ and $2/3$, $(A + r_n \alpha) \cap A = \emptyset$. So $T^{-n} A \cap A = \emptyset$ and R is not a set of recurrence. This completes the proof.

The next theorem describes sets of recurrence of polynomial growth. Theorem 2.4 is a special case of this theorem.

THEOREM 3.5. *Let $p(t)$ be a polynomial with integer coefficients and with $p(0) = 0$. The set $\{p(n), n = 1, 2, 3, \dots\}$ is a set of recurrence.*

PROOF. We consider the Hilbert space $L^2(X, \mathfrak{B}, \mu)$ and the unitary operator defined on it by

$$(Uf)(x) = f(Tx).$$

We use $\langle f, g \rangle$ to denote the inner product in $L^2(X, \mathfrak{B}, \mu)$. According to the spectral theorem, if $f \in L^2(X, \mathfrak{B}, \mu)$, there is a nonnegative measure ω on $[0, 1]$ such that

$$\langle U^n f, f \rangle = \int_0^1 e^{2\pi i n \theta} d\omega(\theta).$$

Now let $A \in \mathfrak{B}$ with $\mu(A) > 0$ and let $f = 1_A$ denote the characteristic function of the set A . Suppose that we had $\mu(T^{-n} A \cap A) = 0$ whenever $n = p(m), m \in \mathbf{N}$. Then

$$(3.1) \quad \int_0^1 e^{2\pi i p(m)\theta} d\omega(\theta) = 0$$

for every $m \in \mathbf{N}$. We now use Weyl's theorem on equidistribution [9] which implies that

$$(3.2) \quad \frac{1}{N} \sum_{m=1}^N e^{2\pi i p(m)\theta} \rightarrow 0$$

as $N \rightarrow \infty$ for every irrational θ . Write $\omega = \omega_i + \omega_r$ where ω_r is the atomic part of ω supported on rational points of $[0, 1]$, and ω_i is what is left of ω . Then (3.1) and (3.2) imply

$$(3.3) \quad \int_0^1 \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{m=1}^N e^{2\pi i p(m)\theta} d\omega_r(\theta) = 0,$$

or, ω_r being atomic,

$$(3.4) \quad \sum_{\theta \in \mathbb{Q}/\mathbb{Z}} \left\{ \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{m=1}^n e^{2\pi i p(m)\theta} \right\} \omega_r\{\theta\} = 0.$$

Now, since the polynomial $p(t)$ has no constant term, $a|p(am)$ for integers a, m . Since (3.1) is equally valid if m is replaced by a multiple of m , we can rewrite (3.4) replacing $p(m)$ by $p(k!m)$, k an arbitrary number. However,

$$e^{2\pi i p(k!m)\theta} = 1$$

for each rational θ whose denominator is $\leq k$. Hence

$$\sum_{\theta} \omega_r\{\theta\} = \lim_{k \rightarrow \infty} \sum_{\theta \in \mathbb{Q}/\mathbb{Z}} \left\{ \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{m=1}^N e^{2\pi i p(k!m)\theta} \right\} \omega_r\{\theta\} = 0$$

by (3.4). Since ω_r is nonnegative, $\omega_r = 0$, and, in particular $\omega_r\{0\} = 0$. But then

(3.5)

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \langle U^n f, f \rangle = \int_0^1 \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N e^{2\pi i n\theta} d\omega(\theta) = \omega\{0\} = 0.$$

On the other hand, by the mean ergodic theorem,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N U^n f = \bar{f}$$

exists in $L^2(X, \mathfrak{B}, \mu)$, \bar{f} is U -invariant and $\int \bar{f} d\mu = \int f d\mu$. By (3.5), $\langle \bar{f}, f \rangle = 0$. Hence $\langle \bar{f}, U^n f \rangle = \langle U^n \bar{f}, U^n f \rangle = \langle \bar{f}, f \rangle = 0$, and by averaging again over n , $\langle \bar{f}, \bar{f} \rangle = 0$, or $\bar{f} = 0$. But then

$$\mu(A) = \int 1_A d\mu = \int \bar{f} d\mu = 0,$$

a contradiction. This proves the theorem.

We may now combine this result with Theorem 2.8. Let \mathcal{Q} be the family of configurations $\{(0, p(m))\}$, $p(t)$ being a polynomial with integer coefficients and with no constant term. We obtain the following number-theoretic result.

THEOREM 3.6. *Let S be a subset of \mathbb{Z} of positive upper density. Let $p(t)$ be a polynomial with integer coefficients and with no constant term. Then one can solve the equation*

$$x - y = p(z)$$

in integers x, y, z with $x, y \in S$ and $z \geq 1$.

This result has been obtained independently by Conze and Sárközy.

4. Proximity and a lemma of Schur. We now turn to a more elaborate study of recurrence in the topological framework. In this section the notion of uniform recurrence (§1) plays a central role. According to Theorem 1.4, every compact dynamical system possesses uniformly recurrent points. In a sense to be made explicit, every orbit of a compact dynamical system keeps coming closer and closer to the orbit of some uniformly recurrent point. To form a

picture of what happens, imagine that X contains a unique minimal closed T -invariant subset which reduces to a point $\{y\}$. The point y is a fixed point and certainly uniformly recurrent. Since every orbit closure in X contains some minimal subset and in our case there is a unique such set, we conclude that every orbit comes arbitrarily close to y . Once the orbit is close to y , by continuity, it will spend a certain interval within a small neighborhood of y . In particular, for each $\varepsilon > 0$, there will be arbitrarily long intervals of n for which $d(T^n x, y) < \varepsilon$. This is a special case of what is called *proximality*.

DEFINITION. Two points $x, y \in X$ are *proximal* for a dynamical system (X, T) if

$$\liminf_{n \rightarrow \infty} d(T^n x, T^n y) = 0.$$

Note that in this definition $n \rightarrow \infty$ and not $|n| \rightarrow \infty$. We might have referred to this as “forward” proximality. It is this form which will be useful and we call attention to the lack of symmetry of past and future.

A far reaching extension of Theorem 1.4 is the following result of J. Auslander and R. Ellis.

THEOREM 4.1. *If (X, T) is a dynamical system, X a compact metric space, and if x is any point in X , there exists a uniformly recurrent point $y \in X$ such that x and y are proximal.*

We refer the reader to [3] or [5] for the proof of this result. We remark that it would be easy to show that any point x is proximal to some minimal set Y . What is more delicate is showing that x is proximal to a particular point $y \in Y$. Recall that by Theorem 1.3, every point in a minimal set is uniformly recurrent.

We now consider what the implications of Theorem 4.1 are for symbolic systems. Suppose then that $X \subset \Lambda^{\mathbb{Z}}$, Λ a finite set, with X a closed translation invariant subset, and let T be defined on X as the shift $Tx(n) = x(n + 1)$. We first inquire what uniform recurrence means for a point $x \in X$. Call a finite sequence of elements of Λ a *word*. We say that a *word occurs in a point* $x \in \Lambda^{\mathbb{Z}}$ if for some n , the sequence $x(n), x(n + 1), \dots, x(n + l - 1)$ coincides with the given word. Similarly we speak of a short word occurring in a long word.

PROPOSITION 4.2. *A point $x_0 \in X \subset \Lambda^{\mathbb{Z}}$ is uniformly recurrent for (X, T) iff every word that occurs in x_0 occurs in every sufficiently large word occurring in x_0 .*

The proof is a straightforward adaptation of uniform recurrence to the present situation. The metric on X being given by (2.1), a neighborhood of a particular point $x_0 \in X$ consists of all points for which a given word occurs as $(x(-k), x(-k + 1), \dots, x(k - 1), x(k))$ for a particular k . A translate of x_0 appears in this neighborhood if the word in question occurs later on in x_0 . The assertion of the proposition is then quite clear.

We now treat proximality.

PROPOSITION 4.3. *Two points $x_0, y_0 \in X$ are proximal iff arbitrarily long words occur at the same position arbitrarily far along the sequences $\{x_0(n)\}$ and $\{y_0(n)\}$.*

In other words, for some $n_l \rightarrow \infty$, we have

$$x(n_l) = y(n_l), x(n_l + 1) = y(n_l + 1), \dots, x(n_l + l - 1) = y(n_l + l - 1).$$

Once again, this follows directly from the definition of proximality. If the condition in the proposition holds then

$$d(T^{n_{2l+1}+l}x, T^{n_{2l+1}+l}y) < \frac{1}{l+1}.$$

The converse direction is also clear.

Let us now apply Theorem 4.1 to the full symbolic system $(\Lambda^{\mathbf{Z}}, T)$. Every $x \in \Lambda^{\mathbf{Z}}$ is therefore proximal to a point $y \in \Lambda^{\mathbf{Z}}$ which is uniformly recurrent. We will use this to conclude the following.

PROPOSITION 4.4. *If x is any point of $\Lambda^{\mathbf{Z}}$, there exist integers $p_1, p_2 > 0$ with $x(p_1 + p_2) = x(p_1) = x(p_2)$.*

PROOF. Let $y \in \Lambda^{\mathbf{Z}}$ be a uniformly recurrent point of $\Lambda^{\mathbf{Z}}$ proximal to x and let $a = y(0)$. Since y is uniformly recurrent, the word consisting of the single symbol a occurs inside of every sufficiently long word of y . Since y and x agree on arbitrarily long blocks of positive values of n , we can find $p_1 > 0$ with $x(p_1) = y(p_1) = a$. Consider next the word $(y(0), y(1), \dots, y(p_1))$. This word also occurs inside every sufficiently long word in y , and there will be a word of this length occurring in x at the same position. Then for appropriate $p_2 > 0$ we will have

$$(x(p_2), x(p_2 + 1), \dots, x(p_2 + p_1)) = (y(0), y(1), \dots, y(p_1)).$$

Hence $x(p_2) = y(0) = y(p_1) = x(p_2 + p_1)$. This proves the proposition.

Proposition 4.4 gives us the following result of Schur [6].

THEOREM 4.5. *If the natural numbers $\{1, 2, 3, \dots\}$ are partitioned into finitely many sets, then one of these sets contains two numbers p_1, p_2 together with their sum $p_1 + p_2$.*

To see this suppose $\mathbf{N} = C_1 \cup C_2 \cup \dots \cup C_r$. Define a point $x \in \{1, 2, \dots, r\}^{\mathbf{Z}}$ by setting

$$x(n) = \begin{cases} i & \text{if } n > 0 \text{ and } n \in C_i, \\ \text{anything} & \text{if } n \leq 0. \end{cases}$$

According to the proposition we can find $p_1, p_2 > 0$ with $x(p_1) = x(p_2) = x(p_1 + p_2)$. If j is this common value then $p_1, p_2, p_1 + p_2 \in C_j$.

Note that unlike the configurations described in §2, the configuration $(p_1, p_2, p_1 + p_2)$ is not translation invariant. It thus happens that if we decompose \mathbf{N} into odd and even numbers, one set contains solutions to $a + b = c$ whereas the other set does not. In particular the kind of configuration described here need not occur in every subset $S \subset \mathbf{N}$ of positive upper density.

Schur was interested in this combinatorial result because it relates to Fermat's last theorem. Theorem 4.5 leads readily to the following "finite" version.

THEOREM 4.5'. *There exists an integer-valued function $N(k) \geq 1$, such that if the interval of integers $[1, N(k)]$ is partitioned into k sets $C_1 \cup C_2 \cup \dots \cup C_k$, then some C_j contains a solution a, b, c to $a + b = c$.*

Now let q be any natural number and let p be a prime with $p > N(q)$ and with $q|p - 1$. The numbers $\{1, 2, \dots, p - 1\}$ may be partitioned into q cosets modulo the multiplicative subgroup of q th powers mod p . Theorem 4.5' asserts that we can find a, b, c inside the same coset satisfying $a + b = c$. This means that the congruence

$$aX^q + aY^q \equiv aZ^q \pmod{p}$$

has a nontrivial solution and therefore the congruence

$$(4.1) \quad X^q + Y^q \equiv Z^q \pmod{p}$$

has a solution with $XYZ \not\equiv 0 \pmod{p}$. It follows that for any q , Fermat's assertion of the unsolvability of $X^q + Y^q = Z^q$ is unlikely to be established by showing that the corresponding congruence has no solution, since (4.1) always has solutions for p sufficiently large. This is Schur's result.

To obtain further consequences of Theorem 4.1 we introduce the notion of an *IP-set* of integers.

DEFINITION. A subset of \mathbf{N} is called an *IP-set* if it consists of a sequence of integers (not necessarily distinct) p_1, p_2, p_3, \dots , together with all sums of these for distinct indices, i.e., all $p_{i_1} + p_{i_2} + \dots + p_{i_k}$ with $i_1 < i_2 < \dots < i_k$.

The notion of an IP-set is somewhat weaker than that of a semigroup. We shall find that IP-sets occur in connection with recurrence.

PROPOSITION 4.6. *Let (X, T) be a dynamical system, X compact metric, let $x, y \in X$ with x proximal to y and y a uniformly recurrent point. Then for any $\epsilon > 0$, there exists an IP-set $S = \{p_{i_1} + p_{i_2} + \dots + p_{i_k}\}$ with $d(T^n x, y) < \epsilon$ for $n \in S$.*

PROOF. We prove first that for any $\delta > 0$ we can find p such that

$$(4.2) \quad d(T^p x, y) < \delta \quad \text{and} \quad d(T^p y, y) < \delta.$$

By uniform recurrence of y there is an N so that for any $n \in \mathbf{Z}$, $d(T^{n+i} y, y) < \delta/2$ for some $i = 0, 1, \dots, N$. Let δ' be so small that

$$d(y_1, y_2) < \delta' \Rightarrow d(T^i y_1, T^i y_2) < \delta/2 \quad \text{for } i = 0, 1, \dots, N.$$

By proximality of x and y there is an n with $d(T^n x, T^n y) < \delta'$. For some $i = 0, 1, \dots, N$ we will have $d(T^{n+i} y, y) < \delta/2$ and at the same time $d(T^{n+i} x, T^{n+i} y) < \delta/2$. Then for $p = n + i$ we have both inequalities of (4.2).

Now suppose we have found p_1, p_2, \dots, p_n so that each $p = p_{i_1} + p_{i_2} + \dots + p_{i_k}$, $i_1 < i_2 < \dots < i_k$, satisfies

$$(4.3) \quad d(T^p x, y) < \epsilon, \quad d(T^p y, y) < \epsilon.$$

In these inequalities we can replace y by anything sufficiently close to it. Let $\delta > 0$ be chosen so that for any $z \in X$ with $d(z, y) < \delta$ we have $d(T^p z, y) < \epsilon$ for the aforementioned finite set of p . Determine p_{n+1} so that

$$d(T^{p_{n+1}}x, y) < \delta, \quad d(T^{p_{n+1}}y, y) < \delta.$$

We then have

$$d(T^{p+p_{n+1}}x, y) < \delta, \quad d(T^{p+p_{n+1}}y, y) < \delta.$$

We can then establish (4.3) for an IP-set of p and this proves the proposition.

Note that an IP-set in \mathbb{N} always contains a solution to $a + b = c$. It was conjectured by Graham and Rothschild that Schur's lemma could be extended to state that in any finite partition of the natural numbers, one of the subsets contains an IP-set. This result was proved by N. Hindman. We now show how this follows from Proposition 4.6.

THEOREM 4.7. *If the natural numbers $\{1, 2, 3, \dots\}$ are partitioned into finitely many sets, then one of these sets contains an IP-set.*

PROOF. Let $\mathbb{N} = C_1 \cup C_2 \cup \dots \cup C_r$ be the partition in question and define $x \in \{1, 2, \dots, r\}^{\mathbb{Z}}$ by setting

$$x(n) = \begin{cases} j & \text{if } n > 0 \text{ and } n \in C_j, \\ \text{anything} & \text{for } n \leq 0. \end{cases}$$

Let $y \in \{1, 2, \dots, r\}^{\mathbb{Z}}$ be a uniformly recurrent point proximal to x . Apply Proposition 4.6 to x and y taking $\epsilon = 1$. There exists accordingly an IP-set S with $d(T^p x, y) < 1$ for $p \in S$. Let $j = y(0)$. We have $x(p) = y(0)$ for $p \in S$; i.e., $S \subset C_j$. This proves the theorem.

A still more general result is given in the next theorem which was also proved by Hindman. In [3] and [5] it is shown how the more general result follows from the special case.

THEOREM 4.8. *If an IP-set of integers is partitioned into finitely many subsets, then one of these subsets contains an IP-set.*

5. Recurrent sets and strong recurrence. Let (X, T) be a dynamical system with X a compact metric space and let A be a closed subset of X . We shall say that *the set A is recurrent* if for any $\epsilon > 0$ there exist points $x, y \in A$ and an integer $n > 0$ with $d(T^n x, y) < \epsilon$. With some additional assumptions on A we can conclude that *some point of A is recurrent for (X, T)* and possibly even that *every point of A is recurrent*.

DEFINITION. A subset $A \subset X$ is *homogeneous* for the system (X, T) if there exists a group G of homeomorphisms of X commuting with T , $ST = TS$ for $S \in G$, such that for each $S \in G$, $SA = A$, and moreover that G acts transitively on A ; i.e., for $x, x' \in A$ there exists $S \in G$ with $Sx = x'$. A is *weakly homogeneous* if the condition that G is transitive on A is replaced by the condition that G acts minimally (so that every G -orbit in A is dense in A).

We shall prove the following theorem.

THEOREM 5.1. *If A is a homogeneous recurrent set for (X, T) then every point of A is recurrent. If A is weakly homogeneous and recurrent, then there exists a dense set of recurrent points in A .*

It is obvious that an automorphism of (X, T) , i.e., a homeomorphism $S: X \rightarrow X$ which commutes with T , takes recurrent points to recurrent points:

$$T^n x \rightarrow x \Rightarrow T^n Sx \rightarrow Sx.$$

Hence both statements of the theorem will follow once it is established that at least one point of A is recurrent. We prove this under the hypothesis of weak homogeneity of A .

LEMMA 5.2. *Suppose A is a weakly homogeneous recurrent set. Then for any $\epsilon > 0$ and for any $y \in A$, there exists $x \in A$ with $d(T^n x, y) < \epsilon$ for some $n > 0$.*

PROOF. By hypothesis we can find sequences $\{x_n\}, \{y_n\}$ with $d(T^n x_n, y_n) \rightarrow 0$. Passing to a subsequence we can assume that we have $y_n \rightarrow y'$. Then y' has the property that for every $\epsilon > 0$ there exists $x \in A$ and $n > 0$ with $d(T^n x, y') < \epsilon$. Now the set of points $y \in A$ with this property is clearly closed. It is also invariant under G . Since G acts minimally on A and the property holds for some point of A , it holds for every point of A .

LEMMA 5.3. *Suppose A has the property that for each $y \in A$ and $\epsilon > 0$ there exists $x \in A$ and $n > 0$ with $d(T^n x, y) < \epsilon$. Then for any $\epsilon > 0$ there exists a point $z \in A$ with $d(T^n z, z) < \epsilon$ for some $n > 0$.*

PROOF. Let $\epsilon > 0$ be given. We shall define inductively a sequence of points z_0, z_1, z_2, \dots in A , one of which will satisfy $d(T^n z, z) < \epsilon$. Set $\epsilon_1 = \epsilon/2$. Choose z_0 arbitrarily in A and let $z_1 \in A$ and $n_1 > 0$ be such that

$$(5.1) \quad d(T^{n_1} z_1, z_0) < \epsilon_1.$$

Now choose $\epsilon_2 < \epsilon_1$ so that $d(z, z_1) < \epsilon_2$ implies that the inequality (5.1) remains valid when z_1 is replaced by z . Then if we find $z_2 \in A$ and $n_2 > 0$ with

$$d(T^{n_2} z_2, z_1) < \epsilon_2,$$

we will also have

$$d(T^{n_1+n_2} z_2, z_0) < \epsilon_1.$$

Proceed inductively in this way, obtaining an array of inequalities

$$(5.2) \quad d(T^{n_j+n_{j-1}+\dots+n_1} z_j, z_0) < \epsilon_{j+1}$$

whenever $i < j$. The successive ϵ_k are chosen so that $d(z, z_j) < \epsilon_{j+1}$ implies that all the inequalities of (5.2) for $i < j$, j fixed, are still valid when z_j is replaced by z . Having determined ϵ_{j+1} we find $z_{j+1} \in A$ and $n_{j+1} > 0$ so that

$$d(T^{n_j+n_{j+1}} z_{j+1}, z_j) < \epsilon_{j+1}.$$

We then obtain the inequalities of (5.2) for $j + 1$ instead of j . The hypothesis of the lemma enables us to proceed indefinitely. So we can suppose that (5.2)

is valid for all $i < j < \infty$. Now for some i, j we shall have $d(z_i, z_j) < \varepsilon_1$, so that

$$d(T^{n_i+n_{i+1}+\dots+n_j}z_i, z_j) < \varepsilon_{i+1} + \varepsilon_1 < \varepsilon,$$

and this proves the lemma.

PROOF OF THE THEOREM. As we have seen it suffices to show that if A is a weakly homogeneous recurrent set, it contains a recurrent point. Form the function

$$F(x) = \inf_{n > 1} d(T^n x, x).$$

Clearly a point is recurrent if $F(x) = 0$. Now $F(x)$ is readily seen to be upper semicontinuous. It follows that it has a point of continuity when restricted to A . Let z be such a point. Assume $F(z) > 0$. We can find a relatively open set $V \subset A$ with $F(x) > \delta > 0$ for all $x \in V$. Now $\cup_{S \in G} S(V)$ is an open G -invariant subset of A and by minimality, this must be all of A . By compactness of A some finite set of $S(V)$ covers A : $A = \cup S_i(V)$. Now for each $S \in G$, $F(Sx) = \inf d(ST^n x, Sx)$, and it follows that for $\delta > 0$ there exists $\varepsilon > 0$ so that $F(x) < \varepsilon \Rightarrow F(Sx) < \delta$. Letting $S = S_i^{-1}$ and taking ε' as the minimum of the corresponding ε , we will have $F(x) \geq \varepsilon'$ for all $x \in A$. For if $F(x) < \varepsilon'$ with $x \in S_i(V)$ then $F(S_i^{-1}x) < \delta$ contradicting the choice of δ . This shows that $F(z) = 0$ at a point of continuity z of F . This proves that A possesses, in fact, a residual set of recurrent points. This completes the proof.

We illustrate the theorem by considering "group extensions" of a given dynamical system.

DEFINITION. Let (X, T) be a dynamical system and let G be a compact group. Let $\psi: X \rightarrow G$ be a continuous function and define \tilde{T} on $\tilde{X} = X \times G$ by

$$\tilde{T}(x, g) = (Tx, \psi(x)g).$$

Then (\tilde{X}, \tilde{T}) is said to be a *group extension* of (X, T) .

Note that the group G acts on \tilde{X} by setting

$$(x, g)^{g'} = (x, gg'),$$

and these transformations commute with T . The fibers $x \times G \subset \tilde{X}$ are invariant under the actions of G and they form homogeneous closed sets. Finally $x \times G$ is a recurrent set of (\tilde{X}, \tilde{T}) iff x is a recurrent point of (X, T) . Applying Theorem 5.1 we have

THEOREM 5.3. *If (\tilde{X}, \tilde{T}) is a group extension of (X, T) then a point $(x, g) \in \tilde{X}$ is recurrent for (\tilde{X}, \tilde{T}) iff x is recurrent for (X, T) .*

We illustrate this in the following proposition.

PROPOSITION 5.4. *Let \mathbf{T} denote the circle group $\mathbf{T} = \mathbf{R}/\mathbf{Z}$. Let $h_0 \in \mathbf{T}$ and for $i = 1, 2, \dots, d - 1$, let h_i be a continuous map from \mathbf{T}^i into \mathbf{T} . Define a*

transformation of $\mathbf{T}^d \rightarrow \mathbf{T}^d$ by

$$(5.3) \quad \begin{aligned} T(\theta_1, \theta_2, \dots, \theta_d) \\ = (\theta_1 + h_0, \theta_2 + h_1(\theta_1), \dots, \theta_d + h_{d-1}(\theta_1, \theta_2, \dots, \theta_{d-1})). \end{aligned}$$

Then every point of \mathbf{T}^d is a recurrent point.

PROOF. By induction on d . Notice that for $d + 1$ the system in question is a group extension of the system for d . For $d = 0$ we take the system to be a single point which is naturally recurrent.

The toral systems described by (5.3) have another property which we shall proceed to study. We call this property *strong recurrence*.

To define this notion we recall that the product of two systems (X, T) and (X', T') is the system $(X \times X', T \times T')$ where $(T \times T')(x, x') = (Tx, T'x')$.

DEFINITION. A point $x \in X$ is *strongly recurrent* for (X, T) if for any system (X', T') and any point $x' \in X'$ which is recurrent for it, the pair (x, x') is recurrent for the product system $(X \times X', T \times T')$.

We will see presently what are the implications of strong recurrence. In particular it will be seen to imply uniform recurrence. In the meantime we have the following example.

PROPOSITION 5.5. *Each point of the system described in Proposition 5.4 is strongly recurrent.*

PROOF. For in this case $(X \times X', T \times T')$ is obtained by a succession of group extensions of (X', T') , so by Theorem 5.3, (x, x') is recurrent for $(X \times X', T \times T')$ iff x' is recurrent for (X', T') .

For the remainder of this section we will be concerned with the notion of strong recurrence. In the next section we return to the notion of recurrent sets.

The next result relates strong recurrence to IP-sets.

THEOREM 5.6. *A point x is a strongly recurrent point of a dynamical system (X, T) iff for any neighborhood V of x and for any IP-set $S \subset \mathbf{N}$, there exists $q \in S$ with $T^q x \in V$.*

For the proof of the theorem we will use the following proposition.

PROPOSITION 5.7. *If $x_0 \in X$ is a recurrent point for (X, T) and V is a neighborhood of x_0 , then the set of $n \in \mathbf{N}$ for which $T^n x_0 \in V$ contains an IP-set. Conversely, if S is an IP-set in \mathbf{N} , there is a system (X, T) , a recurrent point $x_0 \in X$, and a neighborhood V of x_0 such that $S \supset \{n: T^n x_0 \in V\}$.*

PROOF. Assume $x_0 \in X$ is a recurrent point for (X, T) and V is a neighborhood of x_0 . For some number, say p_1 , we have $T^{p_1} x_0 \in V$. Then $x \in V \cap T^{-p_1} V = V_1$, and V_1 constitutes a neighborhood of x . Let p_2 be such that $T^{p_2} x \in V_1$; then $x \in V_1 \cap T^{-p_2} V_1 = V_2$. Continue in this way obtaining a succession of neighborhoods V_k of x_0 with $V_{k+1} = V_k \cap T^{-p_{k+1}} V_k$. At each stage the corresponding power p_{k+1} is defined by the condition $T^{p_{k+1}} x_0 \in V_k$.

We then have

$$V_k = \bigcap_{i_1 < i_2 < \dots < i_n} T^{-(p_{i_1} + p_{i_2} + \dots + p_{i_k})} V.$$

It follows immediately that $\{n: T^n x_0 \in V\}$ contains the IP-set generated by $\{p_i\}$.

For the converse direction suppose $S \subset \mathbb{N}$ is an IP-set; say, $S = \{p_{i_1} + p_{i_2} + \dots + p_{i_k}; i_1 < i_2 < \dots < i_k\}$ with p_1, p_2, \dots a sequence in \mathbb{N} . If we set

$$q_1 = p_{r_1} + p_{r_1+1} + \dots + p_{s_1}, \quad q_2 = p_{r_2} + p_{r_2+1} + \dots + p_{s_2} \\ \dots, q_n = p_{r_n} + p_{r_n+1} + \dots + p_{s_n}, \dots$$

where $\{r_n\}$ and $\{s_n\}$ are two sequences satisfying $r_1 < s_1 < r_2 < s_2 < \dots$, then the IP-set generated by the $q_n: \{q_{i_1} + q_{i_2} + \dots + q_{i_l}; i_1 < i_2 < \dots < i_l\}$ is a subset of S . In this manner we can arrive at an IP-set whose generators satisfy $q_n > 3q_{n-1}$. Let R be this set together with 0. Now let $X = \{0, 1\}^{\mathbb{Z}}$ and let $x_0 = 1_R \in X$. T will be the shift. Let $V = \{x \in X: x(0) = 1\}$. Clearly $R = \{n: T^n x_0 \in V\}$. To complete the proof we show that x_0 is a recurrent point. We claim that $T^{q_n} x_0 \rightarrow x_0$. Namely, let $m \in \mathbb{Z}$. We show that if $m \notin R$, then for n large, $m + q_n \notin R$ and if $m \in R$ then for n large $m + q_n \in R$. The second statement is evident. For the first statement, notice that on account of the rate of growth of q_n ,

$$q_{i_k} < q_{i_1} + \dots + q_{i_k} < \frac{3}{2} q_{i_k}.$$

So for any m , if $m + q_n \in R$ for large n , then since $m + q_n \sim q_n$, $m + q_n = q_{i_1} + q_{i_2} + \dots + q_{i_k}$ for large n implies $n = i_k$, so that $m = q_{i_1} + \dots + q_{i_{k-1}}$.

This completes the proof.

To prove the theorem we suppose x is strongly recurrent, that V is a neighborhood of x and that S is an IP-set. Let x' be a recurrent point for a system (X', T') and V' a neighborhood of x' for which $\{n: T'^n x' \in V'\} \subset S$. Then (x, x') is recurrent for $(X \times X', T \times T')$, and for some n , $(T \times T')^n(x, x') \in V \times V'$. But then $T'^n x' \in V'$ so that $n \in S$; hence $T^n x \in V$ for some $n \in S$. Conversely, suppose x has the property in question and that x' is recurrent for (X', T') . Let W be a neighborhood of (x, x') in $X \times X'$. $W \subset V \times V'$ where V, V' are neighborhoods of x, x' respectively. Since $T'^n x' \in V'$ along some IP-set of n , and for one of these $T^n x \in V$, by hypothesis, we will have $(T \times T')^n(x, x') \in W$ so that (x, x') is recurrent.

We shall use this characterization of strong recurrence in order to prove that strong recurrence implies uniform recurrence. First a lemma.

LEMMA 5.8. *If a subset $S \subset \mathbb{N}$ contains arbitrarily long intervals, then S contains an IP-set.*

PROOF. We have $[a_k, b_k] \subset S$ with $b_k - a_k \rightarrow \infty$. Suppose p_1, p_2, \dots, p_n have been found in S so that every sum of the form $p_{i_1} + p_{i_2} + \dots + p_{i_k}$, $i_1 < i_2 < \dots < i_k \leq n$, belongs to S . We choose $p_{n+1} = a_j$ for some j so large that $b_j - a_j > p_1 + p_2 + \dots + p_n$. Clearly every sum $p_{i_1} + p_{i_2} + \dots + p_{i_k}$, $i_1 < i_2 < \dots < i_k \leq n + 1$, now belongs to S .

THEOREM 5.9. *Strong recurrence implies uniform recurrence.*

PROOF. Suppose $x \in X$ is not uniformly recurrent for the system (X, T) . Then for some neighborhood V of x , the set $\{n; T^n x \in V\}$ has arbitrarily long gaps. By Lemma 5.8, there is an IP-set of n for which $T^n x \notin V$. So by Theorem 5.6, x is not strongly recurrent.

The converse of this is not true. In demonstrating this we shall use the following lemma.

LEMMA 5.10. *If $x_1 \in X_1$ is a strongly recurrent point for (X_1, T_1) and $x_2 \in X_2$ is strongly recurrent for (X_2, T_2) , then (x_1, x_2) is strongly recurrent for $(X_1 \times X_2, T_1 \times T_2)$.*

PROOF. Let x' be recurrent for some third system (X', T') . By strong recurrence (x_2, x') is recurrent for $(X_2 \times X', T_2 \times T')$ and by strong recurrence of x_1 , (x_1, x_2, x') is recurrent for $(X_1 \times X_2 \times X', T_1 \times T_2 \times T')$. But this shows that (x_1, x_2) is strongly recurrent.

The horocycle flow for a compact 2-dimensional surface of constant negative curvature is known to be minimal as a one-parameter flow $(X, T_t; -\infty < t < \infty)$. It follows that every point of X is uniformly recurrent for the flow and so it is uniformly recurrent for (X, T_1) . It is also known that the horocycle flow is weakly mixing ([1]), so that the product system $(X \times X, T_1 \times T_1)$ possesses dense orbits. Now a point $(x, y) \in X \times X$ having a dense orbit cannot be uniformly recurrent, for that would imply, by Theorem 1.3, that $(X \times X, T_1 \times T_1)$ is minimal, and this cannot be the case since the diagonal $\{(x, x); x \in X\}$ is an invariant subset. Hence (x, y) is not uniformly recurrent, and so it is not strongly recurrent. But by Lemma 5.10 this means that one of the points x and y is not strongly recurrent. We thus see that uniform recurrence does not imply strong recurrence.

The same example shows that Lemma 5.10 is not valid if strong recurrence is replaced by uniform recurrence. For each point is uniformly recurrent for the horocycle system and yet there are pairs (x, y) which are not uniformly recurrent for the product system.

The next theorem gives two more characterizations of strong recurrence. The proof can be found in [3].

THEOREM 5.11. *Each of the following conditions for a point $x \in X$ is equivalent to strong recurrence for (X, T) .*

(a) *x is uniformly recurrent and it is not proximal to any point $y \neq x$ in its orbit closure.*

(b) *For any uniformly recurrent point x' for a system (X', T') the point (x, x') is uniformly recurrent for $(X \times X', T \times T')$.*

The characterization (a) of strong recurrence relates the notion to that of *point distality*. Two points of a system are *distal* if they are not proximal. A minimal system with a point that is distal from every other point is *point distal*. There is a rather deep structure theorem of Veech and Ellis which shows how one can construct all point distal metric dynamical systems [8]. The idea of the theorem is that certain extensions of point distal systems

preserve this property. For example, the considerations leading to Theorem 5.3 show that if x is strongly recurrent for (X, T) and (\tilde{X}, \tilde{T}) is a group extension of (X, T) , then each (x, g) is strongly recurrent for (\tilde{X}, \tilde{T}) . This means that the passage from (X, T) to (\tilde{X}, \tilde{T}) preserves point distality. The structure theorem in question asserts that by a possibly transfinite sequence of "allowable" extensions of this kind one can arrive at an arbitrary point distal system starting with the trivial one-point system. The implication of this is that strong recurrence is a relatively rare phenomenon.

We conclude this section with some observations comparing recurrence properties for the systems (X, T) and (X, T^m) . First note that as a special case of Proposition 5.5, the periodic system $(Y_m, T_m): Y_m = \{0, 1, \dots, m-1\}$, $T_m y = y + 1 \pmod{m}$, has every point strongly recurrent. From this it follows that if a point x is recurrent for (X, T) , $(x, 0)$ is recurrent for $(X \times Y_m, T \times T_m)$, and from this it follows that x is recurrent for (X, T^m) .

By a similar argument, using Theorem 5.11(b), we deduce that if x is a uniformly recurrent point for (X, T) then it is also uniformly recurrent for (X, T^m) .

We can also show that a strongly recurrent point of (X, T) is strongly recurrent for (X, T^m) . For by Theorem 5.6, x is strongly recurrent for (X, T^m) iff for every neighborhood V of x and every IP-set $S \subset \mathbb{N}$, some $T^n x \in V$ for $n \in S$. But if S is an IP-set so is mS ; so it follows that some $T^{mn} x \in V$ for $n \in S$. Now this implies that x is strongly recurrent for (X, T^m) .

One can also show that for each of the notions of recurrence a recurrent point for (X, T^m) is also recurrent for (X, T) . We leave it to the reader to verify this.

Combining the foregoing remarks with Lemma 5.10 we obtain the following.

PROPOSITION 5.11. *If x is a strongly recurrent point of (X, T) then $(x, x, \dots, x) \in X^r$ is a strongly recurrent point of $(X^r, T \times T^2 \times \dots \times T^r)$. In particular it is recurrent and so there exists a sequence $n_k \rightarrow \infty$ with $T^{n_k} x \rightarrow x, T^{2n_k} x \rightarrow x, \dots, T^{rn_k} x \rightarrow x$.*

We will see in the next section that the latter property always holds for some point of any system (X, T) . On the other hand we do not know if there always exists a point x such that (x, x, \dots, x) is a uniformly recurrent point for $T \times T^2 \times \dots \times T^r$.

6. Multiple recurrence. We turn now to the proof of Theorem 2.1 which, as we saw in §2, implies the van der Waerden theorem on arithmetic progressions. We prove the following more general theorem.

THEOREM 6.1. *Suppose X is a compact metric space and T_1, T_2, \dots, T_k are commuting maps of X into itself. Then there exists a point $x \in X$ and a sequence $n_k \rightarrow \infty$ with $T_1^{n_k} x \rightarrow x, T_2^{n_k} x \rightarrow x, \dots, T_k^{n_k} x \rightarrow x$.*

The idea of the proof is to show that some point of the diagonal $\Delta = \{(x, x, \dots, x) \in X^l; x \in X\}$ is recurrent for $T_1 \times T_2 \times \dots \times T_l$. In order

to do this we make use of the notion of recurrent sets discussed in the preceding section.

We shall assume that the maps T_i are invertible. It is easy to pass from this case to the general case. The proof of the theorem will proceed by induction on l . The case $l = 1$ is Birkhoff's theorem (Theorem 1.2). Suppose then that the result has been established for $l - 1$ and we have l invertible maps T_1, T_2, \dots, T_l on X . Let G_1 be the group of transformations of X generated by T_1, T_2, \dots, T_l , and let G_l consist of the transformations of $X' = X \times X \times \dots \times X$ of the form $\tilde{S} = S \times S \times \dots \times S, S \in G_1$. Note that G_l takes the diagonal Δ into itself, and in fact that action of G_l on Δ simply mirrors the action of G_1 on X . Now we can assume that G_1 acts minimally on X ; otherwise simply restrict the discussion to a minimal closed G_1 -invariant subset of X . If G_1 acts minimally on X , G_l acts minimally on Δ . Let $\tilde{T} = T_1 \times T_2 \times \dots \times T_l$. \tilde{T} commutes with all the transformations of G_l ; hence Δ is a weakly homogeneous subset for the system (X', \tilde{T}) . If we show that Δ is a recurrent set for \tilde{T} , then by Theorem 5.1, it contains a recurrent point and this will prove the theorem.

Now use the induction hypothesis applied to the $l - 1$ transformations $S_1 = T_1 T_l^{-1}, S_2 = T_2 T_l^{-1}, \dots, S_{l-1} = T_{l-1} T_l^{-1}$. Accordingly, there exists a point $x \in X$ and a sequence $n_k \rightarrow \infty$ with

$$S_1^{n_k} x \rightarrow x, S_2^{n_k} x \rightarrow x, \dots, S_{l-1}^{n_k} x \rightarrow x.$$

This means that for any $\epsilon > 0$ we can find n with the two points

$$(x, x, \dots, x) \in X', \quad (T_1 \times T_2 \times \dots \times T_l)^n (T_l^{-n} x, T_l^{-n} x, \dots, T_l^{-n} x) \in X'$$

within distance ϵ of one another. Since the points (x, x, \dots, x) and $(T_l^{-n} x, T_l^{-n} x, \dots, T_l^{-n} x)$ are in Δ , it follows that Δ is recurrent for $\tilde{T} = T_1 \times T_2 \times \dots \times T_l$. This now completes the proof of the theorem.

Theorem 6.1 leads to a multidimensional extension of van der Waerden's theorem. Namely suppose $Z^h = C_1 \cup C_2 \cup \dots \cup C_r$ is a partition of the h -dimensional lattice Z^h . Consider the space $\Omega = \{1, 2, \dots, r\}^{Z^h}$ and let $\xi \in \Omega$ be the point $\xi(v) = i \Leftrightarrow v \in C_i$. Here v denotes a typical point of Z^h . Let $\{e_1, \dots, e_l\}$ be any finite subset of Z^h , and let $T_i: \Omega \rightarrow \Omega$ be defined by $T_i \omega(v) = \omega(v + e_i), i = 1, \dots, l$. Let $X \subset \Omega$ be the smallest closed subset of Ω containing ξ and invariant under T_1, T_2, \dots, T_l . The T_i commute and so by Theorem 6.1, there exists $\eta \in X$ and $n \in \mathbb{N}$ with the points $T_1^n \eta, T_2^n \eta, \dots, T_l^n \eta$ within distance ϵ of one another. In particular ϵ can be chosen so that this implies that the values of $T_1^n \eta(0), T_2^n \eta(0), \dots, T_l^n \eta(0)$ coincide. This means that $\eta(ne_1) = \eta(ne_2) = \dots = \eta(ne_l)$. Using the fact that η belongs to the closure of translates of ξ we find that there is a vector $v \in Z^h$ with

$$\xi(v + ne_1) = \xi(v + ne_2) = \dots = \xi(v + ne_l).$$

If j is this common value, then $v + ne_i \in C_j$ for $i = 1, 2, \dots, l$. We have thereby proved the following result, due to Gruenwald.

THEOREM 6.2. *Let F be any finite configuration in \mathbf{Z}^h and suppose $\mathbf{Z}^h = C_1 \cup C_2 \cup \dots \cup C_r$ is a finite partition. Then some C_j contains a configuration "similar" to F , i.e., a set of the form $\nu + nF$ for some $\nu \in \mathbf{Z}^h$ and $n \in \mathbf{N}$.*

7. Szemerédi's theorem. It is natural to expect that just as Theorem 6.1 extends Birkhoff's theorem (Theorem 1.2) to the case of commuting transformations, there should exist a theorem extending the Poincaré recurrence theorem to "multiple recurrence". This is indeed the case and one has the following result.

THEOREM 7.1. *Let (X, \mathfrak{B}, μ) be a measure space and let T_1, T_2, \dots, T_l be commuting measure-preserving transformations of (X, \mathfrak{B}, μ) . Let $A \in \mathfrak{B}$ with $\mu(A) > 0$. Then there exists $n \in \mathbf{N}$ with*

$$\mu(A \cap T_1^{-n}A \cap T_2^{-n}A \cap \dots \cap T_l^{-n}A) > 0.$$

A special case of Theorem 7.1 consists of setting $T_i = T^i$, $i = 1, 2, \dots, l$ where (X, \mathfrak{B}, T) is a measure-preserving system. It follows that the set of arithmetic progressions of length l form a set Q of configurations satisfying the hypothesis of Theorem 2.8. According to Theorem 2.8 we can conclude that any subset of \mathbf{Z} of positive upper density contains arithmetic progressions of length l . This is Szemerédi's theorem [7]:

THEOREM 7.2. *If $S \subset \mathbf{Z}$ is a subset of positive upper density, then S contains arbitrarily long arithmetic progressions.*

As in the case of van der Waerden's theorem, the multiple recurrence theorem for arbitrary commuting transformation of a measure space implies (indeed, is equivalent to) a multidimensional extension of Szemerédi's theorem. Namely we have

THEOREM 7.3. *Let S be a subset of \mathbf{Z}^h of positive upper density. If F is any finite configuration in \mathbf{Z}^h , S contains a configuration "similar" to F , i.e., $S \supset \nu + nF$ for some $\nu \in \mathbf{Z}^h$ and $n \in \mathbf{N}$.*

Theorem 7.3 is an example of a combinatorial result which was first proved (and, so far, this is the only proof) by ergodic-theoretic means. For a complete discussion of this as well as the proof of Theorem 7.1 the reader is referred to [3] and [4].

In contrast to the proof of Theorem 6.1 which is the topological analogue of Theorem 7.1, the proof of the measure-theoretic result involves a careful study of the structure of the measure space with regard to the action of T_1, T_2, \dots, T_l . In proving Theorem 6.1 we presented an argument that covers all cases. In proving Theorem 7.1 it seems necessary to distinguish between *mixing actions* of the T_i and some form of *almost periodic* action. At one extreme the expression

$$\mu(A \cap T_1^{-n}A \cap \dots \cap T_l^{-n}A)$$

tends to a constant, $\mu(A)^{l+1}$. At the other extreme this expression behaves like an almost periodic function returning "almost periodically" to the value of the expressions for $n = 0$, namely $\mu(A)$. More precisely one forms the group Γ

of transformations generated by T_1, T_2, \dots, T_l . The measure space (X, \mathfrak{B}, μ) is then decomposed into a succession of factors such that at each stage the group Γ operates and Γ decomposes into two subgroups, one operating in a *relatively mixing* manner, and the other in a *relatively compact* manner. It is rather curious, and not yet understood, why in the topological situation there is no need to carry out such a careful analysis of the structure of the action, whereas this appears, so far, unavoidable in the measure-theoretic case.

REFERENCES

1. L. Auslander, L. Green and F. Hahn, *Flows on homogeneous spaces*, Ann. of Math. Studies, No. 53, Princeton Univ. Press, Princeton, N. J., 1963.
2. G. D. Birkhoff, *Dynamical systems*, Amer. Math. Soc. Colloq. Publ., vol. 9, Amer. Math. Soc., Providence, R. I., 1927.
3. H. Furstenberg, *Recurrence in ergodic theory and combinatorial number theory*, Princeton Univ. Press, Princeton, N. J., 1981.
4. H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for commuting transformations*, J. d'Analyse Math. **34** (1978), 275–291.
5. H. Furstenberg and B. Weiss, *Topological dynamics and combinatorial number theory*, J. d'Analyse Math. **34** (1978), 61–85.
6. I. Schur, *Über die Kongruenz $x^m + y^m \equiv z^m \pmod{p}$* , Jahresber. Deutsch. Math.-Verein. **25** (1916), 114–117.
7. E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245.
8. W. A. Veech, *Topological dynamics*, Bull. Amer. Math. Soc. **85** (1977), 775–830.
9. H. Weyl, *Über die Gleichverteilung von Zahlen mod Eins*, Math. Ann. **77** (1916), 313–352.

INSTITUTE OF MATHEMATICS, HEBREW UNIVERSITY OF JERUSALEM, JERUSALEM, ISRAEL