Disjointness in Ergodic Theory, Minimal Sets, and a Problem in Diophantine Approximation[1]

By

HARRY FURSTENBERG
The Hebrew University, Jerusalem

0. Summary. The objects of ergodic theory—measure spaces with measure-preserving transformation groups—will be called *processes,* those of topological dynamics—compact metric spaces with groups of homeomorphisms—will be called *flows.* We shall be concerned with what may be termed the "arithmetic" of these classes of objects. One may form *products* of processes and of flows, and one may also speak of *factor processes* and *factor flows.* By analogy with the integers, we may say that two processes are *relatively prime* if they have no non-trivial factors in common. An alternative condition is that whenever the two processes appear as factors of a third process, then their product too appears as a factor. In our theories it is unknown whether these two conditions are equivalent. We choose the second of these conditions as the more useful and refer to it as *disjointness.*

Our first applications of the concept of disjointness are to the classification of processes and flows. It will appear that certain classes of processes (flows) may be characterized by the property of being disjoint from the members of other classes of processes (flows). For example the processes with entropy 0 are just those which are disjoint from all Bernoulli flows.

Another application of disjointness of processes is to the following filtering problem. If $\{x_n\}$ and $\{y_n\}$ represent two stationary stochastic processes, when can $\{x_n\}$ be filtered perfectly from $\{x_n + y_n\}$? We will find (Part I, §9) that a sufficient condition is the disjointness of the processes in question.

For flows the principal application of disjointness is to the study of properties of minimal sets (Part III). Consider the flow on the unit circle $K = \{z: |z| = 1\}$ that arises from the transformation $z \to z^2$. What can be said about the "size" of the minimal sets for this flow, that is, closed subsets of K invariant under $z \to z^2$, but not containing proper subsets with these properties. Uncountably many such minimal sets exist in K. Writing $z = \exp(2\pi i\ \Sigma a_n/2^n)$, $a_n = 0, 1$, we see that this amounts to studying the mini-

[1] This research was sponsored in part by the Air Force Office of Scientific Research and by a grant from the Sloan Foundation.

MATHEMATICAL SYSTEMS THEORY, Vol. 1, No. 1.
Published by Springer-Verlag New York Inc.

mal sets of the "shift flow" on binary sequences. Most of the known examples of minimal sets would lead one to conjecture that a minimal set Δ is "small" in the sense that its Hausdorff dimension vanishes. More recent examples, however, show that the Hausdorff dimension may be positive. We shall show that, nevertheless, Δ is small in the sense that it does not form a basis for K. That is, there always exist numbers in K which are not finite products of members of Δ.

Another application is to the following problem. Products of minimal flows need no longer be minimal; are there any special properties that are nonetheless valid for such flows? Our solution to this problem depends upon the fact that minimal flows are disjoint from a class of flows that we call $\mathscr{F}$-flows. In ordinary arithmetic if x and y are relatively prime and y is a factor of xz, then y is already a factor of z. The analogue of this fact for flows is given in Lemma III.1, and this provides the key to our analysis.

The same class $\mathscr{F}$ of flows, as well as the notion of disjointness, arises in connection with a problem in Diophantine approximation (Part IV). Which sets S of positive integers have the property that for every irrational α and $\epsilon > 0$, there exists an $s \in S$ with $|\alpha - x/s| < \epsilon/s$ for some integer x? A complete solution may be given in the case that S is a multiplicative semigroup of integers. In that case we find (Theorem IV.1) that S necessarily possesses this property unless it is contained in the set of powers of a single integer (in which case it does not).

The interdependence between the four parts of this paper has been kept to a minimum. As a result it is possible for a reader interested in topological dynamics to omit Part I which deals with processes and to read only Parts II and III. Although Part IV is presented as an application of the notion of disjointness (and this will be apparent after reading Part III), we have suppressed any explicit dependence on the preceding parts in order to make Theorem IV.1 more readily accessible.

Part I. Disjoint Processes

1. Factors and Products of Processes. Let $(\Omega, \mathscr{F}, \mu)$ denote a probability triple; that is, Ω is a space, $\mathscr{F}$ a σ-field of sets in Ω, and μ a probability measure defined on sets in $\mathscr{F}$. We say T is a measure-preserving transformation of Ω if $T^{-1} A \in \mathscr{F}$ whenever $A \in \mathscr{F}$ and $\mu(T^{-1}A) = \mu(A)$. In this case we say that the quadruple $(\Omega, \mathscr{F}, \mu, T)$ determines a *process*. For a precise definition we introduce a notion of equivalence of quadruples. We say that $(\Omega, \mathscr{F}, \mu, T)$ and $(\Omega', \mathscr{F}', \mu', T')$ are equivalent if to every class of sets modulo null sets in Ω there corresponds a similar class in Ω', and this correspondence carries $\mathscr{F}$ to $\mathscr{F}'$, μ to μ' and T to T'. A process is then an equivalence class of quadruples. In practice we generally choose a representative quadruple which we simply refer to as the process in question and write $X = (\Omega, \mathscr{F}, \mu, T)$. If the transformation T is invertible for some representation of a process, we shall speak of *bilateral* process. As a result of this notion of equivalence, the nature of the space Ω matters very little and only the σ-field $\mathscr{F}$ plays a significant role. For example, if $\Omega' = \Omega \times I$, $\mathscr{F}' = \{A \times I, A \in \mathscr{F}\}$, $\mu'(A \times I) = \mu(A)$ and $T'(\omega,t) = (T\omega,t)$, then, although the spaces Ω and Ω' are quite different, $(\Omega,\mathscr{F},\mu,T)$ and $(\Omega',\mathscr{F}',\mu',T')$ determine the same process. We shall always assume that the probability space $(\Omega,\mathscr{F},\mu)$ is separable (a countable subset of $\mathscr{F}$ is "dense" in $\mathscr{F}$). Every process may then be represented by a quadruple $(\Omega,\mathscr{F},\mu,T)$ where Ω is a compact metric space, $\mathscr{F}$ the σ-field of Borel sets (or its completion with respect to μ), μ a regular Borel measure, and T a continuous transformation. (See the Appendix to Part I for a proof of this assertion.) When this is the case we may suppress the σ-field $\mathscr{F}$ and write $X = (\Omega,\mu,T)$.

If X is a process then Ω_X, $\mathscr{F}_X$, and μ_X will denote a space, a σ-field and a measure such that X may be represented by $(\Omega_X,\mathscr{F}_X,\mu_X,T)$. The symbol T will denote throughout the transformation defining the process. Its domain will always be clear from the context. We shall also denote by T the operator that T induces on the class of functions on Ω_X. Namely, if $\psi(\xi)$ is defined for $\xi \in \Omega_X$, then $T\psi$ will denote the function defined by $T\psi(\xi) = \psi(T\xi)$.

If x_0 is a complex-valued measurable function on Ω, then the sequence of random variables $\{x_n(\omega)\}$ defined by $x_n(\omega) = x_0(T^n\omega)$, or $x_n = T^n x_0$, is a *stationary sequence*. That is, the probability of an event of the form $(x_{n+1}(\omega), x_{n+2}(\omega), \cdots, x_{n+r}(\omega)) \in A$ for a Borel set A in $\mathbf{C}^r$ is independent of n. Conversely, every stationary sequence may be realized in this way. A stationary sequence is said to be *defined for a process* if it arises in the above fashion. Such a sequence also *defines the process* if the σ-field $\mathscr{F}$ is the smallest field with respect to which all the variables $x_n(\omega)$ are measurable. It is not difficult to see that a stationary sequence defines a unique process (because of the equivalence we have introduced). Naturally, a stationary sequence $\{x_n\}$ defined for $-\infty < n < \infty$ defines a bilateral process.

Customarily, one refers to a sequence $\{x_n\}$ of random variables exhibiting the aforementioned stationarity property as a *stationary stochastic process.* For many purposes, one need not distinguish between two stationary sequences that can be defined on the same measure space, e.g., between $\{x_n\}$ and $\{y_n\}$, where $y_n = 2x_n + x_{n+1}$. For this reason we emphasize the notion of a process in which the variables have been suppressed.

Let X and Y be two processes. Suppose φ is a measurable map from Ω_X to Ω_Y satisfying $\mu_Y(A) = \mu_X(\varphi^{-1}A)$ for $A \in \mathscr{F}_Y$, and $T\varphi(\xi) = \varphi T(\xi)$ for $\xi \in \Omega_X$. We then say that Y is a *factor process* of X and write $X \xrightarrow{\varphi} Y$. We also say that φ is a *homomorphism* of X onto Y. Briefly, the condition on φ is that it be measure-preserving and commute with T. For example, if $\mathscr{F}$ is a T-invariant subfield of $\mathscr{F}$ and φ denotes the identity map of Ω to Ω, then $(\Omega,\mathscr{F}',\mu,T)$ is a factor process of $(\Omega,\mathscr{F},\mu,T)$. In fact, every factor process may be realized in this way. We point out that the existence of φ for one realization of X and Y does not imply its existence for every representation; we nonetheless consider Y a factor process of X. When Y is a factor of X, then Ω_X and Ω_Y may be chosen as compact metric spaces in such a way that

φ actually is a continuous onto map. (See the Appendix to Part I.)

The representation of a homomorphism φ in the form $(\Omega,\mathscr{F},\mu,T) \to (\Omega,\mathscr{F}',\mu,T)$ with $\mathscr{F}'$ a subfield of $\mathscr{F}$ shows immediately that any stationary sequence defined for a factor process of X is also defined for X. In general if we have $\varphi: X \to Y$ and y is a random variable defined on Ω_Y then $y \circ \varphi$ is a random variable defined on Ω_X. We speak of $y \circ \varphi$ as a variable defined for X which is *lifted* from Y.

The product of two processes X and Y is defined by letting $(\Omega_{X\times Y},\mathscr{F}_{X\times Y},\mu_{X\times Y})$ be the usual product of the underlying measure spaces of X and Y, and setting $T(\xi,\eta) = (T\xi,T\eta)$. $\Omega_{X\times Y}$ will then be the product of Ω_X and Ω_Y and we shall denote by π_X and π_Y the projections of $\Omega_{X\times Y}$ onto its two components. Note that $X \times Y \xrightarrow{\pi_X} X$ and $X \times Y \xrightarrow{\pi_Y} Y$ so that every pair of processes can always be realized as factor processes of a single process. In this realization the two processes are *independent* in the sense that any set of variables defined for X when lifted to $X \times Y$ is independent of any set of variables defined for Y and lifted to $X \times Y$.

2. Disjointness. Suppose $\{x_n\}$ and $\{y_n\}$ denote two stationary sequences. In general we cannot speak of the joint distributions between variables x_n and y_n until both sequences are defined simultaneously on the same measure space. In particular one may always find a space for which $\{x_n\}$ becomes independent of $\{y_n\}$. It sometimes occurs, however, that this is the only manner in which the two sequences may be combined to form a stationary composite sequence $\{x_n,y_n\}$. More precisely, let us say two stationary sequences $\{x_n\}$ and $\{x'_n\}$ are isomorphic if corresponding joint distributions are identical. We then find that there exist pairs of sequences $\{x_n\}$, $\{y_n\}$, such that if $\{x'_n,y'_n\}$ is a stationary sequence with $\{x'_n\}$ isomorphic to $\{x_n\}$ and $\{y'_n\}$ to $\{y_n\}$, then $\{x'_n\}$ must be independent of $\{y'_n\}$. Let us call this phenomenon *absolute independence.* This is a special case of *disjointness* of processes.

Definitition 1. *Two processes X and Y are disjoint if whenever we have homomorphisms $Z \xrightarrow{\alpha} X$, $Z \xrightarrow{\beta} Y$, then there exists a homomorphism $Z \xrightarrow{\gamma} X \times Y$ such that $\alpha = \pi_X\gamma$, $\beta = \pi_Y\gamma$. We denote disjointness by $X \perp Y$.*

An equivalent condition is that whenever $Z \xrightarrow{\alpha} X$, $Z \xrightarrow{\beta} Y$ then the fields $\alpha^{-1}\mathscr{F}_X$ and $\beta^{-1}\mathscr{F}_Y$ are independent subfields of $\mathscr{F}_Z$.

The necessity is obvious since $\pi_X^{-1}\mathscr{F}_X$ and $\pi_Y^{-1}\mathscr{F}_Y$ are independent in $X \times Y$. The sufficiency stems from the fact that if $\alpha^{-1}\mathscr{F}_X$ and $\beta^{-1}\mathscr{F}_Y$ are independent, then their composition is a field isomorphic to $\mathscr{F}_{X\times Y}$ and $(\Omega_Z,\alpha^{-1}\mathscr{F}_X \cup \beta^{-1}\mathscr{F}_Y,\mu_Z,T) \cong X \times Y$.

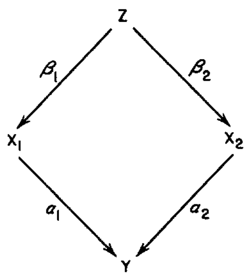
Returning to the sequences $\{x_n\}$ and $\{y_n\}$ that were previously considered, note that if X is the process defined by $\{x_n\}$ and Y the process defined by $\{y_n\}$, then X and Y are factors of the process defined by $\{x'_n,y'_n\}$. Hence if X and Y are disjoint, all the variables x'_n must be independent of all the variables y'_n. Thus, if X and Y are disjoint, the stationary sequences defined for X and Y respectively are absolutely independent. In fact,

LEMMA I.1. *X and Y are disjoint if and only if whenever* $\{x_n\}$ *and* $\{y_n\}$ *are stationary sequences defined for X and Y respectively, then* $\{x_n\}$ *and* $\{y_n\}$ *are absolutely independent.*

Proof: It suffices to show that this condition is sufficient. Let $Z \xrightarrow{\alpha} X$ and $Z \xrightarrow{\beta} Y$, let z_1 be a variable defined on Ω_Z and measurable with respect to $\alpha^{-1}\mathscr{F}_X$ and let z_2 be measurable with respect to $\beta^{-1}\mathscr{F}_Y$. The sequences $\{z_1(T^n\zeta)\}$, $\{z_2(T^n\zeta)\}$ are defined for X and Y respectively, so they are independent, and so $\alpha^{-1}\mathscr{F}_X$ and $\beta^{-1}\mathscr{F}_Y$ are independent fields. Notice that it suffices in fact to have absolute independence for stationary sequences that take on only the two values 0 and 1.

An important lemma is the following

LEMMA I.2. *Suppose as in the diagram we are given homomorphisms* $X_1 \xrightarrow{\alpha_1} Y$, $X_2 \xrightarrow{\alpha_2} Y$. *Then there exists a process Z with homomorphisms* $Z \xrightarrow{\beta_i} X_i$ *such that* $\alpha_1\beta_1 = \alpha_2\beta_2$.



Proof: We assume as we may that Ω_{X_i} and Ω_Y are compact metric spaces and that the maps α_i are continuous surjections. We write Ω_i for Ω_{X_i}, μ_i for μ_{X_i}. To define Z we set

$$\Omega = \{(\omega_1,\omega_2) \in \Omega_1 \times \Omega_2 \colon \alpha_1(\omega_1) = \alpha_2(\omega_2)\} .$$

Ω is a closed subset of $\Omega_1 \times \Omega_2$ and hence is compact. T is defined in Ω by $T(\omega_1,\omega_2) = (T\omega_1,T\omega_2)$. Define $\beta_i\colon \Omega \to \Omega_i$ by $\beta_i(\omega_1,\omega_2) = \omega_i$. Clearly $\alpha_1\beta_1 = \alpha_2\beta_2$ on Ω. It follows that if we find a T-invariant measure μ on Ω satisfying $\beta_i(\mu) = \mu_i$, then (Ω,μ,T) will define a process Z satisfying the requirements of the lemma.

Such a measure μ represents a linear functional L on $C(\Omega)$ satisfying

(i) $Lf \geqslant 0$ if $f \geqslant 0$
(ii) $L1 = 1$
(iii) $L(Tf - f) = 0$ where $Tf(\omega) = f(T\omega)$
(iv) $Lf = 0$ if $f(\omega_1, \omega_2) = F(\omega_i)$, where $\int F(\omega_i)d\mu_i(\omega_i) = 0$.

Conversely, if a functional L exists with these properties, it will determine a measure μ as required. Now conditions (i), (iii) and (iv) describe elements of $C(\Omega)$ that must lie in the half-space $L(f) \geqslant 0$. From the Hahn-Banach theorem it follows that L will exist provided these elements be in a proper closed convex subset of the space. This amounts to the condition that an inequality

$$(1) \qquad f(\omega_1,\omega_2) - f(T\omega_1,T\omega_2) + F_1(\omega_1) + F_2(\omega_2) > C_1 > 0$$

cannot be valid throughout Ω, if $\int F_i(\omega_i) d\mu_i(\omega_i) = 0$ and f, F_1, F_2 are continuous functions.

Suppose (1) were valid. Apply $T, T^2, \cdots, T^n$ to (1), and average the resulting inequalities. Since f is bounded we find, for sufficiently large n, there exist functions $\bar{F}_1$ and $\bar{F}_2$ with

$$(2) \qquad \bar{F}_1(\omega_1) + \bar{F}_2(\omega_2) > C_2 > 0,$$

where again $\int \bar{F}_i(\omega_i)\, d\mu_i(\omega_i) = 0$.

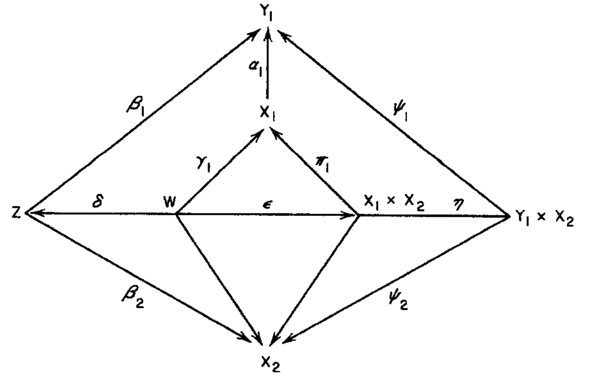
Let $\eta \in \Omega_Y$. For $i = 1, 2$ set $G_i(\eta) = \inf \bar{F}_1(\omega_i)$, where the infimum is taken over the set of ω_i satisfying $\alpha_i(\omega_i) = \eta$. G_i is lower semi-continuous and hence Borel measurable. Moreover $G_i(\alpha_i(\omega_i)) \leqslant F_i(\omega_i)$. Hence $\int G_i(\eta)\, d\mu_Y(\eta) \leqslant 0$. On the other hand, since (2) is valid for all pairs (ω_1,ω_2) with $\alpha_1(\omega_1) = \alpha_2(\omega_2)$, we see that $G_1(\eta) + G_2(\eta) \geqslant C_2 > 0$. Hence $\int [G_1(\eta) + G_2(\eta)]\, d\mu_Y(\eta) \geqslant C_2$, which is a contradiction. This proves the lemma.

The lemma implies the following result for stationary sequences. Let $\{x'_n,y'_n\}$ and $\{y''_n,z''_n\}$ denote composite stationary sequences with $\{y'_n\}$ isomorphic to $\{y''_n\}$. Then there exists a sequence $\{x_n,y_n,z_n\}$ with $\{x_n,y_n\}$ isomorphic to $\{x'_n,y'_n\}$ and $\{y_n,z_n\}$ isomorphic to $\{y''_n,z''_n\}$.

The main application of the lemma is in the proof of the following

PROPOSITION I.1. *For $i = 1, 2$ let $X_i \xrightarrow{\alpha_i} Y_i$ denote homomorphisms of processes. Then $X_1 \perp X_2$ implies $Y_1 \perp Y_2$.*

Proof: It suffices to prove that Y_1 and X_2 are disjoint, for then, by the same token, Y_1 and Y_2 will be disjoint. For this we consider the accompanying diagram. To show that $Y_1 \perp X_2$ we suppose Z is given with homomor-



phisms β_1, β_2. We wish to show that there exists a homomorphism $\theta: Z \to Y_1 \times X_2$ with $\beta_i = \psi_i\theta$, the ψ_i denoting projections. Apply the lemma to $X_1 \xrightarrow{\alpha} Y_1$ and $Z \xrightarrow{\beta} Y_1$; we obtain a process W with homomorphisms δ and γ_1 such that $\alpha_1\gamma_1 = \beta_1\delta$. Define $\gamma_2 = \beta_2\delta$. Consider the homomorphism $W \xrightarrow{\gamma_i} X_i$ and recall that $X_1 \perp X_2$. We thereby obtain a homomorphism ϵ with $\gamma_i = \pi_i\epsilon$, the π_i denoting projections. Finally the homomorphism η is defined by $\eta(\xi_1,\xi_2) = (\alpha_1(\xi_1),\xi_2)$. We then find that the 4 triangles and 2 quadrilaterals in the diagram are commutative. We now claim there is a unique homomorphism $\theta: Z \to Y_1 \times X_2$ that satisfies $\theta\delta = \eta\epsilon$. The uniqueness of θ follows from the fact that $\delta: \Omega_W \to \Omega_Z$ is onto; hence θ is determined by $\theta\delta$. For the existence of θ we must show that if $\omega, \omega' \in \Omega_W$ and $\delta(\omega) = \delta(\omega')$ then $\eta\epsilon(\omega) = \eta\epsilon(\omega')$. But any map φ into $\Omega_{Y_1} \times \Omega_{Y_2}$ is determined by $\psi_1\varphi$ and $\psi_2\varphi$, and $\psi_i\eta\epsilon = \beta_i\delta$ by the commutativity of the diagram. The same argument applied to measures rather than points shows that the map θ is in fact measure-preserving, i.e., a homomorphism. This gives us the desired homomorphism θ. Namely, as before, β_i are determined by $\beta_i\delta$, and $\beta_i\delta = \psi_i\eta\epsilon = \psi_i\theta\delta$ whence $\beta_i = \psi_i\theta$, as was to be shown.

An immediate consequence of this proposition is

PROPOSITION I.2. *If two processes are disjoint they can have no non-trivial factors in common.*

Proof: If $X_1 \perp X_2$ and $X_i \xrightarrow{\alpha_i} Y$, then Y is disjoint from itself. This implies that Y is trivial, i.e., that Ω_Y may be taken as a one-point space. For if $Y \perp Y$, then the identity homomorphism $Y \to Y$ induces a homomorphism $Y \to Y \times Y$ which maps Ω_Y onto the diagonal of $\Omega_Y \times \Omega_Y$. Then the measure $\mu_Y \times \mu_Y$ is concentrated on the diagonal of $\Omega_Y \times \Omega_Y$, which implies that μ_Y reduces to a single point mass and Y is trivial.

Accordingly, a necessary condition for disjointness of two processes is that they possess no common factors. It would be important to know if this condition is also sufficient, as it would facilitate a number of the results which we shall obtain. A result of this kind is in fact valid for other categories. For example, we may say that two groups G_1,G_2 are disjoint if whenever there exist epimorphisms $\alpha_i{:}G \to G_i$, then there exists an epimorphism $\gamma{:}G \to G_1 \times G_2$ with $\alpha_i = \pi_i\gamma$ as before. In this case it is quite easy to show that two groups are disjoint if and only if they have no common factor groups. Incidentally, here we must specify that the maps are onto, a stipulation that was not made for processes. The reason is that for processes it follows automatically that homomorphisms are onto from the requirement that the underlying map be measure-preserving.

The following two open problems are suggested by this analogy, as also by the analogy with ordinary arithmetic.

Problem A: If two processes have no common non-trivial factors, are they disjoint?

Problem B: Does $X \perp Y_1$, $X \perp Y_2$ imply $X \perp (Y_1 \times Y_2)$?

3. Classes of Processes. We present a list of some major classes of processes. The list is by no means exhaustive; we have given priority to those classes of processes to which the concept of disjointness may be fruitfully applied. In the sequel, if $\mathscr{C}$ denotes a class of processes, $\mathscr{C}^\perp$ will denote the class of all processes disjoint from all the processes of $\mathscr{C}$.

(A) Bernoulli Processes. X is a *Bernoulli process*, or $X \in \mathscr{B}$, if X is the process defined by a stationary sequence x_n whose variables are independent. A Bernoulli process may always be represented by forming the infinite product of a measure space $(\Lambda,\mathscr{G},v)$ with itself, $\Omega = \Lambda \times \Lambda \times \Lambda \times \cdots$ and setting $T(\lambda_1,\lambda_2,\lambda_3, \cdots) = (\lambda_2,\lambda_3,\lambda_4, \cdots)$. A bilateral Bernoulli process may be obtained similarly by setting $\Omega = \cdots \times \Lambda \times \Lambda \times \Lambda \times \cdots$.

(B) Pinsker Processes. The class $\mathscr{B}$ is contained in the class of Kolmogorov processes. We shall have nothing to say regarding these, so we refer the reader to [13] for their definition. (In Rokhlin's terminology, one speaks of Bernoulli *automorphisms* and Kolmogorov *automorphisms.*) This class in turn is contained in a class, also defined in [13], whose members we shall call *Pinsker processes.* In [13] they are referred to as automorphisms of *completely positive entropy.* Our definition, however, makes no reference to the notion of entropy.

Definition 2. *A stationary sequence* $\{x_n\}$ *is* deterministic *if* x_1 *is measurable with respect to the* σ*-field generated by the variables* $x_2,x_3,x_4, \cdots$.

In general a process will possess deterministic stationary sequences. For example, suppose $\{y_n, -\infty < n < \infty\}$ is a stationary sequence defined for a bilateral process X such that y_n takes on only the values 0 and 1. Let $x_n = \Sigma_0^\infty y_{n-j}2^{-j-1}$. Then x_n is the fractional part of $2x_{n+1}$ so that $\{x_n\}$ is deterministic. However, *finite-valued* deterministic stationary sequences need not always exist.

We now say that X is a *Pinsker process,* $X \in \mathscr{P}$, if a stationary finite-valued sequence defined for X must be trivial (i.e., $x_n =$ const with probability 1) if it is deterministic. It will develop presently that $\mathscr{B} \subset \mathscr{P}$, i.e., that Bernoulli processes have this property. In other words, if $\{x_n\}$ is a stationary sequence of independent variables and the sequence $\{y_n\}$ is obtained by setting $y_n = f(\cdots, x_n, x_{n+1}, \cdots)$ for some measurable function f with a finite range, then $\{y_n\}$ is not deterministic and a relationship of the form $y_n = F(y_{n+1}, y_{n+2}, \cdots)$ cannot subsist.

(C) Deterministic Processes. On the opposite end of the spectrum are processes with the property that *every* stationary sequence defined for them is deterministic. (If this is true for all finite-valued sequences, it is also true for all sequences.) We call such processes *deterministic* and designate the class of all deterministic processes by $\mathscr{D}$. Clearly $\mathscr{D} \cap \mathscr{P}$ consists of the trivial process.

(D) Ergodic Processes. $\mathscr{E}$ will designate the class of all ergodic processes. We recall that X is ergodic if the condition $TA \subset A$ for $A \in \mathscr{F}_X$ implies $\mu_X(A) = 0$ or 1. We recognize readily that $\mathscr{B} \subset \mathscr{E}$ and $\mathscr{P} \subset \mathscr{E}$.

(E) Mixing Processes. The following four properties of a process X are equivalent:

(i) The equation $Tx = \lambda x$ with x measurable over $\mathcal{F}_X$ has only the solutions $\lambda = 1$, $x = \text{const}$, or $x = 0$.

(ii) $X \times X \in \mathcal{E}$.

(iii) $X \times \mathcal{E} \subset \mathcal{E}$.

(iv) If $\psi,\varphi \in L^2(\Omega_X,\mathcal{F}_X,\mu_X)$ then for each $\epsilon > 0$, the sequence $\{n_k\}$ for which $|\langle\psi,T^{n_k}\varphi\rangle - \langle\psi,1\rangle\langle 1,\varphi\rangle| > \epsilon$ satisfies $n_k/k \to \infty$.

If X satisfies any of these properties it is called *weakly mixing*. We then say $X \in \mathcal{M}$.

The equivalence of (i), (ii) and (iv) is well known (see [9]) and clearly (iii) implies (ii). It is apparently not well known that the usual definition of weak mixing implies (iii), so let us give a proof that (iv) implies (iii).

PROPOSITION I.3. *If X satisfies* (iv) *and Y is ergodic, then $X \times Y$ is ergodic.*

Proof: Since (B) shows that $Z = X \times Y$ is ergodic, it suffices to show that if z is an L^2-variable defined for Z, then $(N+1)^{-1}\sum_0^N T^n z$ converges weakly to a constant. To show this for all such z it suffices to consider z of the form $z(\xi,\eta) = x(\xi)y(\eta)$. Furthermore, in proving weak convergence, it suffices to consider inner products with functions of the same form. Thus, what is to be shown is that

$$\frac{1}{N+1}\sum_0^N \langle T^n x,x'\rangle\langle T^n y,y'\rangle \to \langle x,1\rangle\langle 1,x'\rangle\langle y,1\rangle\langle 1,y'\rangle \ . \tag{3}$$

If we could replace $\langle T^n x,x'\rangle$ by $\langle x,1\rangle\langle 1,x'\rangle$ in the left side of (3), then the result would be immediate, since $(N+1)^{-1}\sum T^n y \to \langle y,1\rangle$ by ergodicity of Y. However, according to (iv), the error in this replacement tends to 0 as $N \to \infty$. This proves the proposition.

We mention in passing that $\mathcal{M}$ contains the class of *strongly mixing processes* which in turn contain those that are *strongly mixing of every order.* The definitions may be found in [13]. It is also known that the processes in $\mathcal{B}$ and $\mathcal{P}$ are strongly mixing of every order. In particular $\mathcal{B},\mathcal{P} \subset \mathcal{M}$. That $\mathcal{B} \subset \mathcal{M}$ may be seen as follows. It is clear that $\mathcal{B} \times \mathcal{B} \subset \mathcal{B}$. Since $\mathcal{B} \subset \mathcal{E}$, it follows that for $X \in \mathcal{B}$, $X \times X \in \mathcal{E}$ and so $X \in \mathcal{M}$ by (ii). It is also quite easy to show, using (i), that $\mathcal{P} \subset \mathcal{M}$ but this will also appear as a consequence then of other considerations.

(F) Kronecker Processes. Let Ω be a compact group and let $\tau \in \Omega$ be an element with the property that it is not contained in a proper closed subgroup of Ω. Let μ denote Haar measure on Ω. Then defining the transformation T on Ω by $T\omega = \tau\omega$, we obtain a process $X = (\Omega,\mu,\tau)$. A process obtained in this way will be called a *Kronecker* process and the class of these will be denoted $\mathcal{K}$. Kronecker processes are never weakly mixing. For, if χ is a character on Ω, then $T\chi(\omega) = \chi(\tau)\chi(\omega)$. On the other hand, a Kronecker process is necessarily ergodic. For, if $\sum c_\chi \chi$ is the expansion of an invariant L^2-function, then $c_\chi = \chi(\tau)c_\chi$. Hence $c_\chi \neq 0$ implies $\chi(\tau) = 1$. But $\{\omega:\chi(\omega) = 1\}$ is a closed subgroup, so that $c_\chi \neq 0$ implies $\chi \equiv 1$.

(G) Weyl Processes. Kronecker processes are special cases of *Weyl processes*. To define the latter we require the notion of a group extension.

Definition 3: Let G be a group of measure-preserving transformations of the underlying measure space $(\Omega_X, \mathcal{F}_X, \mu_X)$ of a process X. Denote the action of G by $\xi \to \xi g$, and assume that G commutes with T so that $(T\xi)g = T(\xi g)$. We suppose that $(\xi, g) \to \xi g$ determines a measurable map from $\Omega_X \times G$ to Ω_X. Let $\mathcal{F}_Y$ denote the σ-field of all G-invariant sets. Then if $Y = (\Omega_X, \mathcal{F}_Y, \mu_X)$, we say that X is a group extension of Y and we write $Y = X/G$.

To illustrate the notion, let Ω_X be a 2-dimensional torus: $\Omega_X = \{(\zeta_1, \zeta_2), |\zeta_1| = |\zeta_2| = 1\}$, μ_X Haar measure on Ω_X, and define T by

$$T(\zeta_1, \zeta_2) = (e^{i\alpha}\zeta_1, \varphi(\zeta_1)\zeta_2) ,$$

where φ denotes a measurable function from the circle to the circle. If we take G to be the group of rotations $(\zeta_1, \zeta_2) \to (\zeta_1, e^{i\beta}\zeta_2)$, then G commutes with T. We see that X is a group extension of the Kronecker process in the circle defined by rotation by α. Processes of this kind were studied in [4]. Note that Kronecker processes are also group extensions of the trivial process. Conversely, it is easily seen that an ergodic group extension of the trivial process is a Kronecker process.

We now define the class $\mathcal{W}$ of Weyl processes as the smallest class of processes satisfying the following three conditions:

(i) The trivial process belongs to $\mathcal{W}$.
(ii) A factor process of a process in $\mathcal{W}$ is again in $\mathcal{W}$.
(iii) An ergodic group extension of a process in $\mathcal{W}$ is in $\mathcal{W}$.

Our nomenclature owes its origin to the fact that a celebrated equidistribution theorem of Weyl (if $p(n)$ is a polynomial in n with an irrational coefficient, its values are equidistributed modulo 1) may be deduced by studying a particular Weyl process.

4. Entropy. For details regarding the contents of this section the reader is referred to [10]. We shall briefly summarize the basic results that we shall need. If $\mathcal{F}$ is a finite field of measurable sets in a measure space, it possesses a quantity of information $H(\mathcal{F})$ defined by

$$H(\mathcal{F}) = \Sigma p_i \log \frac{1}{p_i} ,$$

where the p_i are the probabilities of the atoms of $\mathcal{F}$. Given two fields we find that

$$H(\mathcal{F}_1), H(\mathcal{F}_2) \leq H(\mathcal{F}_1 \vee \mathcal{F}_2) \leq H(\mathcal{F}_1) + H(\mathcal{F}_2) . \tag{4}$$

If we define $H(\mathcal{F}_2/\mathcal{F}_1)$ by $H(\mathcal{F}_1 \vee \mathcal{F}_2) - H(\mathcal{F}_1)$, then

$$0 \leq H(\mathcal{F}_2/\mathcal{F}_1) \leq H(\mathcal{F}_2) \tag{5}$$

In (4) we can show that equality holds on the left for $H(\mathscr{F}_1)$ only if $\mathscr{F}_1 = \mathscr{F}_1 \vee \mathscr{F}_2$, i.e., if $\mathscr{F}_1 \subset \mathscr{F}_1$. Also equality holds on the right only if $\mathscr{F}_1$ and $\mathscr{F}_2$ are independent. Thus $H(\mathscr{F}_2/\mathscr{F}_1) = 0$ implies $\mathscr{F}_2 \subset \mathscr{F}_1$, and $H(\mathscr{F}_2/\mathscr{F}_1) = H(\mathscr{F}_2)$ implies $\mathscr{F}_2$ is independent of $\mathscr{F}_1$. The function $H(\mathscr{F}_2/\mathscr{F}_1)$ is monotonic in both variables, but increases as $\mathscr{F}_2$ increases while it decreases if $\mathscr{F}_1$ increases. Hence it may be defined for $\mathscr{F}_1$ an arbitrary field, as long as $\mathscr{F}_2$ is finite. With these stipulations $H(\mathscr{F}_2/\mathscr{F}_1)$ satisfies the foregoing conditions; it is an increasing function of $\mathscr{F}_2$, a decreasing function of $\mathscr{F}_1$, $H(\mathscr{F}_2/\mathscr{F}_1)$ vanishes only if $\mathscr{F}_2$ is contained (modulo null sets) in $\mathscr{F}_1$, and $H(\mathscr{F}_2/\mathscr{F}_1) = H(\mathscr{F}_2)$ only if $\mathscr{F}_2$ is independent of $\mathscr{F}_1$. We also have

$$(6) \qquad H(\mathscr{F}'_2/\mathscr{F}_1), H(\mathscr{F}''_2/\mathscr{F}_1) \leq H(\mathscr{F}'_2 \vee \mathscr{F}''_2/\mathscr{F}_1) \leq H(\mathscr{F}'_2/\mathscr{F}_1) + H(\mathscr{F}''_2/\mathscr{F}_1)$$

Given a set of random variables in a measure space, there is determined a least σ-field with respect to which they are all measurable. If $\mathscr{F}_1$ is the field determined in this manner by $\{x_\alpha\}$ and $\mathscr{F}_2$ the field determined by $\{y_\beta\}$, we shall write $H(\{y_\beta\}/\{x_\alpha\})$ in place of $H(\mathscr{F}_2/\mathscr{F}_1)$. So if y is a finite-valued random variable, then $H(y/\{x_\alpha\})$ is always defined and finite. It will vanish only if y is a function of the x_α, and it will equal $H(y)$ only if y is independent of the x_α.

Now let $\{x_n\}$ denote a stationary sequence of finite-valued random variables. We define the *entropy* of the sequence, by

$$(7) \qquad \epsilon(\{x_n\}) = H(x_1/x_2,x_3,x_4, \cdots)$$

Note that

$$\begin{aligned} H(x_1,x_2, \cdots ,x_n) &= H(x_2, \cdots ,x_n) + H(x_1/x_2, \cdots ,x_n) \\ &= \sum_1^{n-1} H(x_i/x_{i+1}, \cdots ,x_n) + H(x_n) \\ &= \sum_2^{n} H(x_1/x_2, \cdots ,x_j) + H(x_1) \end{aligned}$$

by stationarity. Dividing by n and using the monotonicity of $H(x_1/x_2, \cdots, x_j)$ we find that

$$(8) \qquad \epsilon(\{x_n\}) = \lim_{n \to \infty} \frac{H(x_1,x_2, \cdots ,x_n)}{n}$$

By (8) we see that

$$(9) \qquad \epsilon(\{x_n\}), \epsilon(\{y_n\}) \leq \epsilon(\{x_n,y_n\}) \leq \epsilon(\{x_n\}) + \epsilon(\{y_n\}) .$$

Now let X be a process. We define $\epsilon(X)$ as the l.u.b. of the entropies of all finite-valued stationary sequences defined for X.

The following result is due to Sinai [10]:

PROPOSITION I.4. *If X is the process defined by a finite-valued stationary sequence $\{x_n\}$, then $\epsilon(X) = \epsilon(\{x_n\})$.*

The notion of entropy gives a simple characterization of the classes $\mathscr{D}$ and $\mathscr{P}$:

PROPOSITION I.5. *$X \in \mathscr{D}$ if and only if $\epsilon(X) = 0$, and $X \in \mathscr{P}$ if and only if $\epsilon(Y) > 0$ for every non-trivial factor process Y of X.*

Proof: A deterministic finite-valued sequence clearly has entropy 0 by (7). Conversely, if a sequence has entropy 0, it must be deterministic since, as has been remarked, $H(x_1/\mathscr{F}) = 0$ only if x_1 is measurable with respect to $\mathscr{F}$. Since a Pinsker process possesses no deterministic finite-valued sequences, its entropy must be positive. But a factor process of a Pinsker process is clearly again a Pinsker process, so the proposition follows.

An important consequence of Proposition I.4 is that a deterministic finite-valued sequence defines a deterministic process. It is not *a priori* clear that if X is the process defined by a deterministic finite-valued sequence, then *every* sequence defined for X is deterministic. But, according to Proposition I.4, $\epsilon(X) = 0$, and so $\epsilon(\{y_n\}) = 0$ for every $\{y_n\}$ defined for X.

5. Entropy and Disjointness

THEOREM I.1. *Two processes with positive entropy cannot be disjoint.*

Proof: By Lemma I.1, it suffices to prove that two stationary sequences with positive entropy cannot be disjoint. Moreover, by the remark following Lemma I.1, it suffices to do this for stationary 2-valued sequences. So suppose $\{x_n\}$, $\{y_n\}$ have positive entropy and take on only the values 0, 1. Let us show that we can form a stationary sequence $\{x^*_n, y^*_n\}$ with $\{x^*_n\}$ isomorphic to $\{x_n\}$, $\{y^*_n\}$ isomorphic to $\{y_n\}$ but with $\{x^*_n\}$ not independent of $\{y^*_n\}$. There will be no loss of generality if we suppose that the sequences are indexed for $-\infty < n < \infty$.

Let Ω_2 denote the space of all binary sequences: $\omega \in \Omega_2$ if $\omega = (\cdots, \omega_{-2}, \omega_{-1}, \omega_0)$ with $\omega_i = 0, 1$. Let μ_X denote the measure on Ω_2 representing the distribution of the Ω_2-valued random variable $(\cdots, x_{-2}, x_{-1}, x_0)$, and let μ_Y denote the distribution of $(\cdots, y_{-2}, y_{-1}, y_0)$. Let (I,m) denote the underlying measure space of the Bernoulli process defined by a sequence $t_1, t_2, \cdots$ when the t_i are independent and uniformly distributed in $[0, 1]$. Form the measure space

$$(\Omega^*, \mu^*) = (\Omega_2 \times \Omega_2 \times I,\ \mu_X \times \mu_Y \times m)\ .$$

(The spaces being compact metric, we have suppressed the σ-field.)

The conditional probabilities $P(x_1 = 1/x_0, x_{-1}, \cdots)$, $P(y_1 = 1/y_0, y_{-1}, \cdots)$ are defined almost everywhere on (Ω_2, μ_X) and (Ω_2, μ_Y) respectively. We can therefore define the following function on $[0, 1] \times \Omega_2 \times \Omega_2$:

$$F(t,\xi,\eta) = \begin{cases} 1 & \text{if } t < P(x_1 = 1/x_0, x_{-1}, \cdots) \\ 0 & \text{if } t \geqslant P(x_1 = 1/x_0, x_{-1}, \cdots)\ , \end{cases} \tag{10}$$

where the conditional probability is to be evaluated at ξ. Similarly we set

$$G(t,\xi,\eta) = \begin{cases} 1 \text{ if } t < P(y_1 = 1/y_0,y_{-1}, \cdots) \\ 0 \text{ if } t \geq P(y_1 = 1/y_0,y_{-1}, \cdots) \end{cases} , \tag{11}$$

with the conditional probability evaluated at η.

We now define a sequence of variables on (Ω^*,μ^*) which we denote (x^*_n,y^*_n). We define the variables x^*_n,y^*_n for $n \leq 0$ by $x^*_n(\xi,\eta,\tau) = \xi_n$, $y^*_n(\xi,\eta,\tau) = \eta_n$. Also the variable t_n for $n > 0$ is defined by $t_n(\xi,\eta,\tau) = \tau_n$. To define x^*_n,y^*_n we proceed inductively. Namely, we set

$$\begin{aligned} x^*_{n+1} &= F(t_{n+1}; (\cdots ,x_{n-1},x_n), (\cdots ,y_{n-1},y_n)) \\ y^*_{n+1} &= G(t_{n+1}; (\cdots ,x_{n-1},x_n), (\cdots ,y_{n-1},y_n)) . \end{aligned} \tag{12}$$

We claim that by proceeding in this manner we obtain sequences $\{x^*_n\}$, $\{y^*_n\}$ isomorphic to $\{x_n\}$, $\{y_n\}$ respectively. Because of the stationarity of the definition (10), it suffices to prove that $\{\cdots ,x_{-1},x_0,x_1\}$ is isomorphic to $\{\cdots ,x^*_{-1},x^*_0,x^*_1\}$, and similarly for the y_n. To compute a typical expectation involving these variables, it suffices to consider functions of the form $x_1\psi(x_0,x_{-1}, \cdots)$ and $x^*_1\psi(x^*_0,x^*_{-1}, \cdots) = x^*_1\psi(x_0,x_{-1}, \cdots)$. The expectation of the former is $E[P(x_1 = 1/x_0,x_{-1}, \cdots)\psi(x_0,x_{-1}, \cdots)]$, and that of the latter is $E[F(t_1; \cdots ,x_{-1},x_0, \cdots ,y_{-1},y_0)\psi(x_0,x_{-1}, \cdots)]$. The variable t_1 is independent of all the other variables, and so the latter expectation becomes

$$E\left[\int_0^1 F(t; \cdots ,x_{-1},x_0, \cdots ,y_{-1},y_0)\psi(x_0,x_{-1}, \cdots)dt\right]$$
$$= E[P(x_1 =1/x_0,x_{-1}, \cdots)\psi(x_0,x_1, \cdots)] ,$$

by (10). The same argument holds for the y_n, and the isomorphism between the starred and unstarred sequences is established.

We next claim that $\{x^*_n\}$ and $\{y^*_n\}$ are not independent. In fact, were they independent then $E(x^*_1y^*_1/x_0,y_0,x_{-1},y_{-1}, \cdots)$ would be given as the product $E(x^*_1/x_0,x_{-1}, \cdots)E(y^*_1/y_0,y_{-1}, \cdots)$. But

$$E(x^*_1y^*_1/x_0,y_0,x_{-1},y_{-1}, \cdots) =$$

$$E[F(t; \cdots ,x_{-1},x_0, \cdots ,y_{-1},y_0)G(t; \cdots ,x_{-1},x_0, \cdots ,y_{-1},y_0)/x_0,y_0, \cdots] =$$

$$\min \{P(x_1 = 1/x_0,x_{-1}, \cdots), P(y_1 = 1/y_0,y_{-1}, \cdots)\} .$$

On the other hand, the product referred to is simply $P(x_1 = 1/x_0,x_{-1}, \cdots) \times P(y_1 = 1/y_0, y_{-1}, \cdots)$. Setting $u = P(x_1 = 1/x_0, x_{-1}, \cdots)$, $v = P(y_1 = 1/y_0, y_{-1}, \cdots)$ we obtain, almost everywhere, $uv = \min \{u,v\}$. Suppose, with positive probability, $0 < u < 1$, $0 < v < 1$; this is incompatible with $uv = \min \{u,v\}$.

Hence if u is neither 0 or 1, then v must be either 0 or 1. On the other hand, u and v are independent; hence either u takes on only the values 0, 1 or v takes on only these values. But if, say, $u = 0$, 1 with probability 1, then $\{x_n\}$ is deterministic. By hypothesis, $\epsilon(\{x_n\}) > 0$ and $\epsilon(\{y_n\}) > 0$, and so we conclude that $\{x^*_n\}$ and $\{y^*_n\}$ are not independent; hence $\{x_n\}$ and $\{y_n\}$ are not absolutely independent. This proves the theorem.

There are concrete examples of processes with positive entropy where the conclusion of the theorem is not evident. For example, let A and B be respectively m- and n-dimensional unimodular matrices with eigenvalues distinct from roots of unity. Then letting K^m and K^n denote tori of dimensions m and n, (K^m,A) and (K^n,B) determine processes of positive entropy with respect to Lebesgue measure (see [13]). Hence they are not disjoint, and there exists a measure on K^{m+n} invariant with respect to $A \oplus B$ which projects into m- and n-dimensional Lebesgue measure under the projections onto the first n and last n coordinates, but which is distinct from $(m + n)$-dimensional Lebesgue measure. It would be of interest to construct such a measure directly.

COROLLARY: *$\mathscr{B}^\perp \subset \mathscr{D}$ and $\mathscr{P}^\perp \subset \mathscr{D}$.*

Proof: If $Y \in \mathscr{B}$ or $Y \in \mathscr{P}$, $\epsilon(Y) > 0$. Hence if $X \perp Y$, then $\epsilon(X) = 0$, or $X \in \mathscr{D}$.

6. $\mathscr{B}^\perp = \mathscr{D}$. To identify $\mathscr{B}^\perp$ with $\mathscr{D}$ we must still show that $\mathscr{B}$ and $\mathscr{D}$ are disjoint, i.e., that every Bernoulli process is disjoint from every deterministic process. By Lemma I.1, it suffices to show that every finite-valued Bernoulli sequence is absolutely independent of every finite-valued deterministic sequence.

Let $\{x_n\}$ denote a finite-valued Bernoulli sequence, $\{y_n\}$ a finite-valued deterministic sequence. From §4 (9) we obtain

$$\epsilon(\{x_n\}) \leqslant \epsilon(\{x_n,y_n\}) \leqslant \epsilon(\{x_n\}) + \epsilon(\{y_n\}) = \epsilon(\{x_n\}) \ , \tag{13}$$

since $\epsilon(\{y_n\}) = 0$. Also $\epsilon(\{x_n\}) = H(x_1)$, the x_n being independent. So $\epsilon(\{x_n,y_n\}) = H(x_1)$. By §4 (7),

$$\epsilon(\{x_n,y_n\}) = H(x_1,y_1/x_2,y_2,x_3,y_3, \cdots) \ ,$$

so by §4 (6),

$$H(x_1) \leqslant H(x_1/x_2,y_2,x_3,y_3, \cdots) + H(y_1/x_2,y_2,x_3,y_3, \cdots) \ .$$

Now $H(y_1/y_2,y_3, \cdots) = 0$ and, *a fortiori*, $H(y_1/x_2,y_2, \cdots) = 0$. Hence

$$H(x_1) \leqslant H(x_1/x_2,y_2,x_3,y_3, \cdots) \leqslant H(x_1)$$

by §4 (5). But this implies that x_1 is independent of the field generated by $y_2,y_3, \cdots$ which is the field generated by all the y_n. Thus each individual

x_m is independent of $\{y_n\}$. The same argument applied to the sequence of variables $(x_1,x_2, \cdots ,x_r)$, $(x_{r+1},x_{r+2}, \cdots ,x_{2r})$, $\cdots$ shows that the entire sequence $\{x_n\}$ is independent of $\{y_n\}$. We have thereby proved

THEOREM I.2. $\mathscr{B}^\perp = \mathscr{D}$.

COROLLARY: $\mathscr{B} \subset \mathscr{P}$.

Proof: Since $\mathscr{B} \perp \mathscr{D}$ it follows that $\mathscr{B} \subset \mathscr{D}^\perp$. We claim that $\mathscr{D}^\perp \subset \mathscr{P}$. For, if X is disjoint from all deterministic processes, it cannot (by Proposition I.1) possess a deterministic factor process. So it cannot possess deterministic stationary finite-valued sequences since these define deterministic processes (see the end of §4). This proves the corollary.

According to Rokhlin [13], a result has been proved by Pinsker which is tantamount to the statement that $\mathscr{P} \perp \mathscr{D}$. This implies that $\mathscr{D}^\perp = \mathscr{P}$. For we saw that $\mathscr{D}^\perp \subset \mathscr{P}$, and $\mathscr{P} \perp \mathscr{D}$ implies that $\mathscr{P} \subset \mathscr{D}^\perp$. We thus have the relationships

$$\mathscr{B}^\perp = \mathscr{D},\ (\mathscr{B}^\perp)^\perp = \mathscr{P}\ . \tag{14}$$

We note that if the answer to Problem A in §2 is affirmative, then it is clear that $\mathscr{P} \perp \mathscr{D}$. For factors of Pinsker processes are Pinsker processes and factor of deterministic processes are deterministic, and we have already remarked that $\mathscr{P} \cap \mathscr{D}$ is trivial.

The result $(\mathscr{B}^\perp)^\perp = \mathscr{P}$ suggests another problem. For any class $\mathscr{C}$, $\mathscr{C}^{\perp\perp}$ not only contains $\mathscr{C}$ but it contains all factors of all processes in $\mathscr{C}$. This suggests

Problem C. Let $\mathscr{B}^*$ denote the class of all factors of Bernoulli processes. Then

$$\mathscr{B} \subset \mathscr{B}^* \subset \mathscr{P}\ .$$

Are either or both of these inclusions actually equalities?

The possibility that $\mathscr{B} = \mathscr{P}$ exists only in the bilateral case. The stationary Markov chain $x_1,x_2,x_3, \cdots$ with transition matrix $\begin{pmatrix} p & q \\ q & p \end{pmatrix}$, $p + q = 1$, $p \neq$ ½, determines a Pinsker process which is not a Bernoulli process. It has not been shown that the *bilateral* sequence $\cdots ,x_{-1},x_0,x_1, \cdots$ does not define a Bernoulli process.

7. Disjointness and Weakly Mixing Processes.

THEOREM I.3. $\mathscr{M} = \mathscr{K}^\perp \cap \mathscr{E}$.

Proof: We first show that an ergodic process which is disjoint from all Kronecker processes is weakly mixing. If X is not weakly mixing, there is a measurable non-trivial solution to $Tx = \lambda x$ on Ω_X. Inasmuch as T is measure-preserving, λx and x have the same distribution, so that $|\lambda| = 1$. By ergodicity $|x|$ is constant, and we may assume $|x| = 1$. Then x is a measurable map of Ω_X to the circle K. Let Y denote the process on K defined by the transformation $\zeta \to \lambda\zeta$ and whose measure is the image under x of μ_X. Then $X \to Y$. Since Y is ergodic, μ_Y is either Lebesgue measure or is concentrated at the vertices of a regular polygon. In either case Y is a Kronecker process and we cannot have $X \perp \mathscr{K}$.

To complete the proof of the theorem we must show that $\mathscr{M} \perp \mathscr{K}$. This is a consequence of the following stronger result:

THEOREM I.4. $\mathscr{M} \perp \mathscr{W}$.

Proof: In §3 (G) we defined $\mathscr{W}$ as the smallest class of processes satisfying (i), (ii), and (iii). To prove the theorem it will suffice to prove that $\mathscr{M}^{\perp}$ satisfies (i), (ii) and (iii) of §3 (G). Now (i) and (ii) are satisfied by any class of the form $\mathscr{C}^{\perp}$, so what remains is showing that $\mathscr{M}^{\perp}$ is closed under passage to ergodic group extensions. Suppose then that $Y = X/G$, that X is ergodic, and that $Y \perp \mathscr{M}$. We wish to show that $X \perp \mathscr{M}$. Choose $W \in \mathscr{M}$, and suppose we have $Z \xrightarrow{\alpha} X$, $Z \xrightarrow{\beta} W$. Define the map $(\alpha \times \beta)$: $\Omega_Z \to \Omega_X \times \Omega_W$ by

$$(\alpha \times \beta)(\zeta) = (\alpha(\zeta), \beta(\zeta)) \ .$$

Now $\pi_X(\alpha \times \beta) = \alpha$, $\pi_W(\alpha \times \beta) = \beta$; hence in order to prove that $X \perp W$, it suffices to show that $\alpha \times \beta$ defines a homomorphism of Z on $X \times W$. For this it is only necessary to show that $(\alpha \times \beta)(\mu_Z) = \mu_X \times \mu_W$. If we let $\mu^* = (\alpha \times \beta)(\mu_Z)$, then μ^* is a measure on $\Omega_X \times \Omega_W$ satisfying

$$\text{(a)} \qquad \int f(\xi,\omega)\, d\mu^*(\xi,\omega) = \int f(\xi,\omega)\, d\mu_X(\xi)\, d\mu_W(\omega),$$

where $f(\cdot,\omega)$ is measurable with respect to $\mathscr{F}_Y$, and

$$\text{(b)} \qquad T\mu^* = \mu^*.$$

The reason that (a) is satisfied is that $Y \perp W$. Hence μ^* when restricted to functions on $\Omega_Y \times \Omega_W$ reduces to a product measure. We wish to show that $\mu_X \times \mu_W$ is the only measure satisfying (a) and (b).

For any measure μ on $\Omega_X \times \Omega_W$ and $g \in G$, let μ_g denote the measure satisfying

$$\int f(\xi,\omega)\, d\mu_g(\xi,\omega) = \int f(\xi g,\omega)\, d\mu(\xi,\omega) \ .$$

Let $\psi(g)$ denote bounded measurable function on G, then μ_ψ will be defined by

$$(15) \qquad \int f(\xi,\omega)\, d\mu_\psi(\xi,\omega) = \int_G \int f(\xi g,\omega)\psi(g)\, d\mu(\xi,\omega)\, dg \,.$$

Clearly μ_ψ is absolutely continuous with respect to μ_1. Suppose μ satisfies (a) and (b). We shall show that μ_1 coincides with $\mu_X \times \mu_W$. By (15),

$$\int f(\xi,\omega)\, d\mu_1(\xi,\omega) = \int F(\xi,\omega)\, d\mu(\xi,\omega) \ ,$$

where $F(\xi,\omega) = \int_G f(\xi g,\omega)\, dg$. Hence $F(\xi g,\omega) = F(\xi,\omega)$ for $g \in G$ and so $F(\cdot,\omega)$ is measurable with respect to $\mathscr{F}_Y$. By (a) it follows that $\mu_1 = \mu_X \times \mu_W$. Now $W \in \mathscr{M}$ and $X \in \mathscr{E}$. Hence $X \times W \in \mathscr{E}$, and $(\Omega_X \times \Omega_W, \mu_1, T)$ is ergodic. As we have remarked, the measures μ_ψ are absolutely continuous with respect to μ_1; since they are also T-invariant by (b), it follows that $\mu_\psi = (\int \psi(g)\, dg)\mu_1$. This implies that almost all the measures μ_g are identical with $\mu_1 = \mu_X \times \mu_W$, and since the latter is G-invariant we conclude that $\mu = \mu_X \times \mu_W$. This completes the proof of the theorem.

COROLLARY 1. *Weyl processes are deterministic.*

Proof: By theorem I.2, $\mathscr{D} = \mathscr{B}^\perp$. Now $\mathscr{B} \subset \mathscr{M}$ implies that $\mathscr{M}^\perp \subset \mathscr{B}^\perp = \mathscr{D}$. Finally, $\mathscr{W} \subset \mathscr{M}^\perp$ implies $\mathscr{W} \subset \mathscr{D}$.

COROLLARY 2. *Pinsker processes are weakly mixing.*

Proof: In the proof of Theorem I.3 we saw that if an ergodic process is not mixing, it has a Kronecker factor. But $\mathscr{K} \subset \mathscr{W} \subset \mathscr{D}$, so such a process contains a deterministic factor. But a Pinsker process cannot contain a non-trivial deterministic factor; hence $\mathscr{P} \subset \mathscr{M}$.

8. Relations Among Classes. For the reader's convenience we assemble here the relationships between the various classes of processes we have introduced.

(a) $\mathscr{B} \subset \mathscr{P} \subset \mathscr{M} \subset \mathscr{E}$;

(b) $\mathscr{K} \subset \mathscr{W} \subset \mathscr{D}$;

(c) $\mathscr{B}^\perp = \mathscr{D}^\perp \subset \mathscr{P}$;

(d) $\mathscr{K}^\perp \cap \mathscr{E} = \mathscr{M}$; $\mathscr{M}^\perp \supset \mathscr{W}$;

(e) The classes $\mathscr{B}$, $\mathscr{M}$, $\mathscr{W}$, and $\mathscr{D}$ satisfy $\mathscr{C} \times \mathscr{C} \subset \mathscr{C}$;

(f) The classes $\mathscr{P}$, $\mathscr{M}$, $\mathscr{K}$, $\mathscr{W}$, $\mathscr{D}$, $\mathscr{E}$ are closed under passage to factors.

All of these have either been proven or are self-evident with the exception of

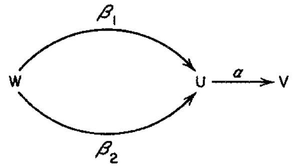
(i) $\mathscr{W} \times \mathscr{W} \subset \mathscr{W}$; (ii) $\mathscr{D} \times \mathscr{D} \subset \mathscr{D}$; (iii) A factor process of a Kronecker process is a Kronecker process.

(i) follows from the inductive definition of $\mathscr{W}$. Namely the class of processes X satisfying $X \times \mathscr{W} \subset \mathscr{W}$ clearly satisfies the conditions of §3(G). Hence $\mathscr{W}$ is contained in this class. (ii) follows by virtue of the fact that $\epsilon(X \times Y) = \epsilon(X) + \epsilon(Y)$. (iii) follows from the characterization of Kronecker processes as ergodic processes with discrete spectrum.

9. Disjointness and Filtering. The problem of filtering for stationary sequences may be described as follows. We suppose given a composite stationary sequence $\{x_n, y_n\}$ in the sense that all the joint distributions between all sets of variables are known. $\{x_n\}$ is to represent a "signal" and $\{y_n\}$ the "noise". The problem is to find a function of the sequence $\{x_n + y_n\}$ which comes as close as possible to a specified variable x_m. In particular we may ask under what conditions is each x_m itself a (measurable) function of $\{x_n + y_n\}$? In other words, when will it be the case that the σ-field spanned by the variables $\{x_n + y_n\}$ contains that spanned by $\{x_n\}$ (and hence also that spanned by $\{y_n\}$, since $y_n = (x_n + y_n) - x_n$)?

Let us reword this in the terminology of processes. Let U denote the process defined by the variables $\{x_n, y_n\}$ and V the process defined by $\{v_n\}$, where $v_n = x_n + y_n$. Ostensibly V is a factor process of U. We seek conditions under which the homomorphism $U \to V$ is actually an isomorphism.

PROPOSITION I.6. *Let $U \xrightarrow{\alpha} V$ denote a homomorphism of processes. If α is not an isomorphism, there exists a process W and there exist distinct homomorphisms $W \xrightarrow{\beta_1} U$, $W \xrightarrow{\beta_2} U$ such that $\alpha\beta_1 = \alpha\beta_2$.*



Proof: We assume (see §1) that Ω_U, Ω_V are compact metric spaces and that α is a continuous map of Ω_U onto Ω_V. Consider the conditional expectation $E(\cdot/\alpha^{-1}(\mathscr{F}_V))$ as an operator on bounded measurable functions on Ω_U. It is not hard to show that this operator has a "kernel". This means there is a measurable function μ_ω defined almost everywhere on Ω_U and whose values are themselves measures on Ω_U such that

$$E(f/\alpha^{-1}(\mathscr{F}_V))(\omega) = \int f(\omega')\, d\mu_\omega(\omega') \tag{16}$$

almost everywhere. The function has the following properties:

(a) μ_ω is a probability measure with support in the set of ω' satisfying $\alpha(\omega') = \alpha(\omega)$,

(b) $\mu_{T\omega} = T\mu_\omega$,

(c) (16) is satisfied whenever $f(\omega')$ is a continuous function on Ω_U.

For the construction of the function μ_ω we refer the reader to [4, p. 592].

Define Ω_W to be the closed subset in $\Omega_U \times \Omega_U$ of pairs of points satisfying $\alpha(\omega_1) = \alpha(\omega_2)$. We let β_1 and β_2 denote the two projections of Ω_W onto Ω_U. Clearly $\alpha\beta_1 = \alpha\beta_2$. T is defined in Ω_W in the usual manner. A measure μ_W is defined by the condition

$$\begin{aligned} &\int f(\omega_1,\omega_2)\, d\mu_W(\omega_1,\omega_2) \\ &\quad = \int \left[\int \int f(\omega_1,\omega_2)\, d\mu_\omega(\omega_1)\, d\mu_\omega(\omega_2) \right] d\mu_U(\omega) \end{aligned} \tag{17}$$

for every continuous function f on Ω_W. By (b) it follows that $T\mu_W = \mu_W$. Thus (Ω_W, μ_W, T) defines a process W. We claim that $\beta_i(W) = U$ for $i = 1, 2$. Since $\beta_i(\Omega_W) = \Omega_U$ we need only show that $\beta_i(\mu_W) = \mu_U$. Take $i = 1$. What has to be shown is that if $f = \varphi \circ \beta_1$ for $\varphi \in C(\Omega_U)$, then $\int f\, d\mu_W = \int \varphi\, d\mu_U$.

But by (17),

$$\int f\, d\mu_W = \int\left[\int \varphi(\omega_1)\; d\mu_\omega(\omega_1)\right] d\mu_U(\omega)$$
$$= E(E(\varphi/\alpha^{-1}(\mathscr{F}))) = E(\varphi) = \int \varphi\, d\mu_U .$$

Our proof will be complete if we can show that $\beta_1 \neq \beta_2$. Suppose then that $\beta_1 = \beta_2$. This means that the measure μ_W on $\Omega_U \times \Omega_U$ must actually concentrate on the diagonal of this product. This implies that

$$\int \varphi(\omega_1)\varphi(\omega_2)\; d\mu_W(\omega_1,\omega_2) = \int \varphi(\omega_1)^2\; d\mu_W(\omega_1,\omega_2) .$$

Applying (17) and (16), we obtain

$$E(E(\varphi/\alpha^{-1}(\mathscr{F}_V))^2) = E(\varphi^2) .$$

The projection $\varphi \to E(\varphi/\alpha^{-1}(\mathscr{F}_V))$ can be norm-preserving for all (or for a dense set of) $\varphi \in L^2(\Omega_U,\mu_U)$ only if it is the identity. This would imply that $\alpha^{-1}(\mathscr{F}_V) = \mathscr{F}_U$, which means that α is an isomorphism. This proves the proposition.

Let us now return to our filtering problem with U and V as before. Suppose the homomorphism $U \xrightarrow{\alpha} V$ is not an isomorphism; then we construct the process W with the homomorphisms $W \xrightarrow{\beta_i} U$ as in the projection. Set $\{x'_n,y'_n\} = \{x_n \circ \beta_1, y_n \circ \beta_1\}$, $\{x''_n,y''_n\} = \{x_n \circ \beta_2, y_n \circ \beta_2\}$. Since U is defined by $\{x_n,y_n\}$, the stipulation that β_1 and β_2 are distinct implies that we cannot have an identity $x'_n = x''_n$, $y'_n = y''_n$. On the other hand, the variables v_n on V satisfy $v_n \circ \alpha = x_n + y_n$, so that

$$x'_n + y'_n = x_n \circ \beta_1 + y_n \circ \beta_1 = v_n \circ \alpha\beta_1 = v_n\alpha\beta_2$$
$$= x_n \circ \beta_2 + y_n \circ \beta_2 = x''_n + y''_n .$$

We have thereby proved

PROPOSITION I.7. *Let $\{x_n,y_n\}$ be a stationary sequence with the property that if $\{x'_n,y'_n,x''_n,y''_n\}$ is a stationary sequence with $\{x'_n,y'_n\}$ and $\{x''_n,y''_n\}$ isomorphic to $\{x_n,y_n\}$ and with $x'_n + y'_n = x''_n + y''_n$, then $x'_n = x''_n$ and $y'_n = y''_n$. Then the σ-fields spanned by $\{x_n + y_n\}$ and $\{x_n,y_n\}$ are identical, so that $\{x_n\}$ is a "function" of $\{x_n + y_n\}$.*

We shall say in this case that the sequence $\{x_n,y_n\}$ *admits a perfect filter.* To apply the criterion of the proposition we shall need

LEMMA I.3. *Let u_1,u_2,v_1,v_2 denote four integrable random variables with each of the u_i independent of each of the v_j. Then $u_1 + v_1 = u_2 + v_2$ together with $E(u_1) = E(u_2)$ implies $u_1 = u_2$, $v_1 = v_2$.*

Proof: Each $u_i v_j$ is integrable because of the independence of the factors. We also have $E(u_1 - u_2) = E(v_1 - v_2) = 0$. But $(u_1 - u_2)^2 = (u_1 - u_2) \times (v_2 - v_1)$, so that

$$E((u_1 - u_2)^2) = E(u_1 - u_2)E(v_2 - v_1) = 0 .$$

Hence $u_1 = u_2$ and $v_1 = v_2$.

Remark: It would be of interest to know if the integrability stipulation may be omitted, replacing the equality of the expectations of u_1 and u_2 by equality of their distributions. That is to say, under these conditions together with the independence of each u_i of each v_j, does $u_1 + v_1 = u_2 + v_2$ imply $u_1 = u_2$, $v_1 = v_2$? A positive answer would be implied by an affirmative answer to the following "elementary" question.

Problem D: Let z_1, z_2, w_1, w_2 be four random variables with z_1 and z_2 having the same distribution and with w_1 and w_2 having the same distribution. Does the inequality $z_1 + w_1 \geqslant z_2 + w_2$ imply that equality holds with probability 1?

Setting $z_i = u_i v_1$ and $w_j = u_j v_2$, we see that the conditions of this problem are met, and $z_1 + w_1 = z_2 + w_2$ would imply $u_1 = u_2$, $v_1 = v_2$.

If on the other hand we stipulate that each u_i is independent of the *pair* (v_1, v_2), then the desired conclusion follows. For then $E(e^{itu_2}) = E(e^{itu_1}) \times E(e^{it(v_1 - v_2)})$ and for t sufficiently small, $E(e^{it(v_1 - v_2)}) = 1$. But this is known to imply that $v_1 - v_2 = 0$.

The main result of this section is

THEOREM I.5. *Let $\{x_n\}$, $\{y_n\}$ be two stationary sequences of integrable random variables, and suppose that the two sequences are absolutely independent (i.e., that the processes they determine are disjoint). Then $\{x_n, y_n\}$ admits a perfect filter.*

The proof is immediate, taking into account Proposition I.7 and Lemma I.3.

The question arises whether the integrability requirement is essential for the conclusion of the theorem. Whether or not it can be eliminated depends on the answer to the question raised in the foregoing Remark. The following is an example where the integrability requirement may be omitted.

PROPOSITION I.8. *If $\{y_n\}$ is defined for a Bernoulli process and $\{x_n\}$ defines a deterministic process, then $\{x_n, y_n\}$ admits a perfect filter.*

Proof: We must show that $x'_n + y'_n = x''_n + y''_n$ implies $x'_n = x''_n$. But $\{x'_n, x''_n\}$ is deterministic by §4 (9), and so $\{y'_n\}$ is independent of it (recall $\mathscr{B} \perp \mathscr{D}$). The argument preceding Theorem I.5 then shows that $x'_n = x''_n$.

In case $\{y_n\}$ is itself a Bernoulli sequence, i.e., if the y_n are independent, and if $\{x_n\}$ is deterministic, and if, furthermore, all the variables are integrable with $E(y_n) = 0$, then we can exhibit the "filter" explicitly, namely,

$$x_n = E(x_n + y_n / x_{n-1} + y_{n-1}, x_{n-2} + y_{n-2}, \cdots) . \tag{18}$$

For the conditioning σ-field in (18) is the same as that spanned by the variables $\{x_{n-1},y_{n-1},x_{n-2},y_{n-2},\cdots\}$ and

$$E(x_n/x_{n-1},y_{n-1},\cdots)+E(y_n/x_{n-1},y_{n-1},\cdots)=x_n+E(y_n)=x_n\ .$$

Here we have used the fact that x_n is measurable with respect to $\{x_{n-1},y_{n-1},\cdots\}$ and y_n is independent of this field (see §6).

10. Stochastic Sequences and Disjointness. In [3, Chap. 1] we developed the notion of a *stochastic sequence.* Briefly stated, it is a generalization of the notion of a *normal* number.

Definition 4. Let Λ be a compact metric space and $\xi=\{\xi_n, n=1,2,\cdots\}$ a Λ-valued sequence. We say ξ is a *stochastic* sequence if whenever f is a continuous function on some product Λ^k, the limit

$$\lim_{N\to\infty}\frac{1}{N}\sum_1^N f(\xi_{n+1},\xi_{n+2},\cdots,\xi_{n+k})$$

exists.

ξ is a point of the product space $\Omega=\Lambda\times\Lambda\times\Lambda\times\cdots$. It follows from the definition that if ξ is stochastic and $f\in C(\Omega)$, then

$$\lim_{N\to\infty}\frac{1}{N}\sum_1^N f(T^{n-1}\xi)=L(f)$$

exists, where T denotes the usual shift transformation. The linear functional $L(f)$ corresponds to a probability measure μ on Ω, and, since $L(Tf)=L(f)$, the triple (Ω,μ,T) defines a process X. The point $\xi\in\Omega$ is a *generic* point in accordance with the following definition.

Definition 5. Let X be a process and Ω_X a compact metric realization of the sample space of X. A point $\xi\in\Omega_X$ is *generic* if

$$\frac{1}{N}\sum_1^N f(T^{n-1}\xi)\to\int f(\omega)\,d\mu_X(\omega)$$

as $N\to\infty$, for every $f\in C(\Omega_X)$.

A stochastic sequence thus determines a process and a generic point for the process. Conversely, given a process X, a generic point $\xi\in\Omega_X$, and a Λ-valued continuous function $\lambda(\omega)$ on Ω_X, it is easily seen that

$$\xi_n=\lambda(T^{n-1}\xi)$$

represents a stochastic sequence.

Almost all points are generic in the case of an ergodic process. Hence almost all points lead to stochastic sequences. In fact, one can show that for any stationary sequence $\{x_n\}$, almost all $\{x_n(\omega)\}$ are stochastic sequences. All almost periodic sequences are stochastic. As with almost periodic sequences, a continuous function of a single stochastic sequence $\{\zeta_n\}=\{f(\xi_n,\xi_{n+1},\cdots)\}$ is again stochastic. The question arises whether functions of several stochastic sequences are again stochastic. This is not the case, as the following example shows.

Let $\eta_n = (-1)^\nu$, where ν is the greatest integer in log n. Then $\{\eta_n\}$ is not stochastic. However, if $\{\xi_n\}$ is generic for a Bernoulli process with independent variables x_n taking on the values ± 1 with probability ½, then it can be seen that $\{\xi_n\eta_n\} = \{\xi'_n\}$ is again generic for the same process. This means that the product $\{\xi_n\xi'_n\} = \{\eta_n\}$ of two stochastic sequences need not be stochastic. The same example shows that their sum need also not be stochastic, since $\eta_n = \frac{1}{2}((\xi_n + \xi'_n)^2 - 2)$.

We have however the following theorem.

THEOREM I.6. *Let $\{\xi_n\}$ and $\{\eta_n\}$ be stochastic sequences that are generic for disjoint processes X and Y respectively. Then the composite sequence $\{\xi_n,\eta_n\}$ is stochastic and is generic for $X \times Y$. In this case any sequence of the form*

$$\{\zeta_n\} = \{f(\xi_n,\xi_{n+1}, \cdots ; \eta_n,\eta_{n+1}, \cdots)\}$$

is stochastic if f is continuous.

Proof: The last statement of the theorem follows from the preceding one since functions of stochastic sequences are stochastic. To prove the theorem we must show that $(\xi,\eta) \in \Omega_X \times \Omega_Y$ is generic for $X \times Y$, i.e., that

$$\frac{1}{N}\sum_{1}^{N} f(T^{n-1}\xi,T^{n-1}\eta) \to \int f(\omega,\omega')\, d\mu_X(\omega)\, d\mu_Y(\omega) \tag{19}$$

as $N \to \infty$, for $f \in C(\Omega_X \times \Omega_Y)$.

If the sequence in (19) did not converge to the value indicated, a subsequence would converge to some other value. Refining the subsequence, we could obtain a limit not just for the f in question, but for a countable set of functions in $C(\Omega_X \times \Omega_Y)$. Choosing this countable set to be dense, we conclude that there is a sequence $\{N_k\}$ with

$$\frac{1}{N_k}\sum_{1}^{N_k} f(T^{n-1}\xi,T^{n-1}\eta) \to L(f)$$

for every $f \in C(\Omega_X \times \Omega_Y)$, and where the linear functional L does not coincide with that given by the right side of (19).

In any case $L(f)$ corresponds to a T-invariant probability measure μ_Z and defines a process $Z = (\Omega_X \times \Omega_Y,\mu_Z,T)$. Let π_X and π_Y denote the projections from $\Omega_X \times \Omega_Y$ to Ω_X and Ω_Y respectively. We claim that $\pi_X Z = X$, $\pi_Y Z = Y$. Now

$$\int \psi \circ \pi_X(\omega,\omega')\, d\mu_Z(\omega,\omega') = \lim_{N\to\infty} \frac{1}{N_k} \sum_{1}^{N_k} \psi \circ \pi_X(T^{n-1}\xi,T^{n-1}\eta)$$

$$= \lim_{k\to\infty} \frac{1}{N_k} \sum_{1}^{N_k} \psi(T^{n-1}\xi) = \int \psi(\omega)\, d\mu_X(\omega)$$

since ξ is generic for X. This means that $\pi_X(\mu_Z) = \mu_X$. Similarly $\pi_Y(\mu_Z) = \mu_Y$, and this proves that Z maps onto X and Y. Hence $Z = X \times Y$ so that $\mu_Z = \mu_X \times \mu_Y$. This contradicts our assumption that $L(f)$ is not given by $\mu_X \times \mu_Y$, and this proves the theorem.

As an application let us prove that if $.a_1a_2a_3\ldots$ is a normal number to the base r, then so is every arithmetic subsequence: $.a_la_{l+d}a_{l+2d}\ldots$. This contains the result of Niven and Zuckerman in [11]. The hypothesis of normality is equivalent to the requirement that $\{a_n\}$ is stochastic and generic for a Bernoulli process $\{x_n\}$, where $x_n = 0,1,\cdots,r-1$ with equal probabilities. The assertion that every arithmetic subsequence of difference d is again normal can be seen to be equivalent to the condition that the composite sequence $\{a_n,e^{2\pi in/d}\}$ is stochastic and generic for the process determined by $\{x_n,y_n\}$, where y_n is the Kronecker process defined by $y_{n+1} = e^{2\pi in/d}y_n$. But this is immediate since $\mathscr{B} \perp \mathscr{D} \supset \mathscr{K}$.

Appendix. Let X and Y denote processes and suppose there exists a homomorphism $X \xrightarrow{\alpha} Y$. We wish to show that X and Y can be realized as quadruples $(\Omega_X,\mathscr{F}_X,\mu_X,T)$ and $(\Omega_Y,\mathscr{F}_Y,\mu_Y,T)$ such that Ω_X and Ω_Y are compact metric spaces, $\mathscr{F}_X$ and $\mathscr{F}_Y$ are the Borel fields, μ_X and μ_Y are Borel measures and α corresponds to a continuous map of Ω_X onto Ω_Y.

We suppose to begin with that $X = (\Omega,\mathscr{F},\mu,T)$ and that $Y = (\Omega,\mathscr{F}',\mu,T)$ with $\mathscr{F}' \subset \mathscr{F}$. Choose a countable dense subset of $L^2(\Omega,\mathscr{F}',\mu)$ and extend it to a countable dense subset of $L^2(\Omega,\mathscr{F},\mu)$. Without loss of generality we may suppose that the functions chosen are bounded. Extending the set further, if necessary, we may suppose that it is invariant with respect to T. Let A denote the closure in $L^\infty(\Omega,\mathscr{F},\mu)$ of the algebra generated by these functions and their conjugates. Let A' denote the subalgebra corresponding to those functions measurable with respect to $\mathscr{F}'$. A and A' are commutative C^*-algebras and we may represent them as $A \cong C(\Omega_X)$, $A' \cong C(\Omega_Y)$, where Ω_X,Ω_Y are compact spaces, and Ω_Y is the image of Ω_X under a continuous map α. Since A and A' are separable, Ω_X and Ω_Y are metrizable. The measure μ induces a linear functional on $C(\Omega_X)$ which in turn determines a measure μ_X on Ω_X. In addition, the endomorphism T of A with itself induces a continuous map T of Ω_X onto itself, and similarly for A' and Ω_Y. It can be seen that μ_X is invariant under T and that with $\mu_Y = \alpha(\mu_X)$, (Ω_Y,μ_Y,T) defines a factor process of (Ω_X,μ_X,T). Now the correspondence between functions in A and continuous functions on Ω_X induces an isometry of a dense subset of $L^2(\Omega,\mathscr{F},\mu)$ with a dense subset of $L^2(\Omega_X,\mu_X)$. This extends to an isometry of the two L^2 spaces. Restricting this isometry to indicator functions of measurable sets in Ω we find that we have an equivalence of $(\Omega,\mathscr{F},\mu,T)$ and (Ω_X,μ_X,T). Similarly, $(\Omega,\mathscr{F}',\mu,T)$ is seen to be equivalent to (Ω_Y,μ_Y,T), and one sees easily that the original homomorphism restricting to sets in $\mathscr{F}'$ is given by the map α.

Part II. Disjoint Flows.

1. Basic definitions. A *flow* is a pair (Ω,T), where Ω is a compact metric space and T is a continuous map of Ω into Ω. When T is one-one and onto it generates a group of homeomorphisms of Ω, and we speak of a *bilateral* flow. If ω is a point of Ω, the sequence $\omega, T\omega, T^2\omega, \cdots$ is referred to as the *orbit* of ω.

Let $X = (\Omega_X, T)$, $Y = (\Omega_Y, T)$, and suppose that φ is a continuous map of Ω_X *onto* Ω_Y such that $T\varphi(\xi) = \varphi(T\xi)$, $\xi \in \Omega_X$. We then say that Y is a *factor flow* of X and we write $X \xrightarrow{\varphi} Y$.

It is also convenient to introduce the notion of a *subflow*. We say $Y = (\Omega_Y, T)$ is a subflow of $X = (\Omega_X, T)$ if Ω_Y is a closed subset of Ω_X and the transformation T on Ω_Y is the restriction of the corresponding transformation on Ω_X.

The *product* of two flows is defined by $\Omega_{X\times Y} = \Omega_X \times \Omega_Y$, $T(\xi,\eta) = (T\xi, T\eta)$. X and Y are both factors of $X \times Y$ and we will denote the projection homomorphisms by $X \times Y \xrightarrow{\pi_X} X$ and $X \times Y \xrightarrow{\pi_Y} Y$.

Definition II.1. *Two flows $X = (\Omega_X, T)$ and $Y = (\Omega_Y, T)$ are* disjoint *if whenever there exists a flow Z with $Z \xrightarrow{\alpha} X$, $Z \xrightarrow{\beta} Y$, then there also exists a homomorphism $Z \xrightarrow{\gamma} X \times Y$ with $\alpha = \pi_X\gamma$, $\beta = \pi_Y\gamma$. We then write $X \perp Y$.*

LEMMA II.1. *$X \perp Y$ if and only if the only closed subset $\Delta \subset \Omega_X \times \Omega_Y$ satisfying* (a) $T\Delta \subset \Delta$, (b) $\pi_X\Delta = \Omega_X$, (c) $\pi_Y\Delta = \Omega_Y$, *is $\Delta = \Omega_X \times \Omega_Y$.*

Proof: Suppose the condition is fulfilled and $Z \xrightarrow{\alpha} X$, $Z \xrightarrow{\beta} Y$. Form the map $\gamma: \Omega_Z \rightarrow \Omega_X \times \Omega_Y$ defined by $\gamma(\zeta) = (\alpha(\zeta), \beta(\zeta))$. We see at once that the set $\Delta = \gamma(\Omega_Z)$ satisfies (a), (b), and (c), and is closed. Hence $\gamma(\Omega_Z) = \Omega_X \times \Omega_Y$. But then γ defines a homomorphism of Z to $X \times Y$, and evidently $\pi_X\gamma = \alpha$, $\pi_Y\gamma = \beta$.

Conversely, assume that $X \perp Y$. Define Z by $\Omega_Z = \Delta$ for a set $\Delta \subset \Omega_X \times \Omega_Y$ satisfying (a), (b) and (c). Then $Z \xrightarrow{\pi_X} X$, $Z \xrightarrow{\pi_Y} Y$, and so γ must exist such that $Z \xrightarrow{\gamma} X \times Y$ with $\pi_X = \pi_X\gamma$, $\pi_Y = \pi_Y\gamma$. This means that $\gamma(\xi,\eta) = (\xi,\eta)$, and since γ must be onto, Δ must coincide with $\Omega_X \times \Omega_Y$.

The following lemma corresponds to Lemma I.2:

LEMMA II.2. *Let X_1, X_2, Y denote three flows and suppose there exist homomorphisms $X_1 \xrightarrow{\alpha_1} Y$, $X_2 \xrightarrow{\alpha_2} Y$. Then there exists a flow Z and maps $Z \xrightarrow{\beta_i} X_i$ with $\alpha_1\beta_1 = \alpha_2\beta_2$.*

The proof of this lemma is contained in the proof of the more difficult Lemma I.1 to which we refer the reader. We may also draw from this lemma the corresponding conclusion:

PROPOSITION II.1. *If $X_1 \xrightarrow{\alpha} Y_1$ and $X_2 \xrightarrow{\beta} Y_2$, then $X_1 \perp X_2$ implies $Y_1 \perp Y_2$.*

It is evident that isomorphic non-trivial flows cannot be disjoint. For, the diagonal of $\Omega_X \times \Omega_X$ satisfies (a), (b), and (c) in Lemma II.1, and the diagonal will not coincide with $\Omega_X \times \Omega_X$ unless Ω_X reduces to a single point. This implies

PROPOSITION II.2. *Two disjoint flows cannot have any common non-trivial factors.*

Whether the converse is true is again unknown.

We remark that while a flow cannot be disjoint from a factor flow, it may be disjoint from a subflow. We shall encounter examples of this in Part III.

If $X = (\Omega,T)$ is a flow, it is always possible to find a probability Borel measure μ on Ω which is invariant with respect to T. (Let ν be any probability measure and let μ be a weak limit point of $n^{-1}(\nu + \cdots + T^{n-1}\nu)$.) The triple (Ω,μ,T) then determines a process which we shall denote by (X,μ). In general the measure μ is not uniquely determined, so that in the above manner one may obtain a large family of processes *supported* by a flow X.

2. Sequences Defined for a Flow. Let $X = (\Omega,T)$ be a flow and suppose $f \in C(\Omega)$, that is, f is a continuous complex-valued function on Ω. For any $\omega \in \Omega$ we may form the sequence $\{f(T^n\omega), n = 0,1,2, \cdots\}$ which we then call a sequence *defined for the flow* X. More generally, f may be a continuous map of Ω into a compact metric space Λ, in which case we speak of a Λ-sequence. Every sequence with values in a compact metric space is associated with some flow. Namely, let $\xi = \{\xi(n)\}$, so that ξ is a point of the cartesian product $\Lambda \times \Lambda \times \Lambda \times \cdots$. Defining T on Λ^∞ as the operator $T(\lambda_1, \lambda_2, \cdots) = (\lambda_2,\lambda_3, \cdots)$, we find that (Λ^∞,T) defines a flow and $\{\xi(n)\}$ is a Λ-sequence defined for the flow.

It is evident that we need not necessarily choose $\Omega = \Lambda^\infty$ to obtain the sequence $\{\xi(n)\}$ as a sequence defined for the flow. Any T-invariant closed subset to which ξ belongs will do. In particular we can choose for Ω the closure of the orbit of ξ in (Λ^∞,T). This subflow (Ω,T) will be referred to as the flow *determined* by the sequence $\{\xi(n)\}$.

3. Classes of Flows. As in Part I we shall enumerate various classes of flows with an eye to determining disjointness relations between them. Among the classes we shall describe, the classes of minimal and distal flows have attracted attention in the literature ([2], [5], [6]). The others have not been studied extensively and their theory is only lightly touched upon even here.

(A) Bernoulli Flows. When the space Ω_X of a flow X has the form $\Omega_X = \Lambda \times \Lambda \times \Lambda \times \cdots$ (where Λ is a compact metric space) and T is the shift transformation, $T(\lambda_1,\lambda_2, \cdots) = (\lambda_2,\lambda_3, \cdots)$, then X is referred to as a *Bernoulli* flow. The class of Bernoulli flows will be denoted by $\mathscr{B}$. If Λ is finite, we say that X is a Bernoulli flow of *finite type.*

Every flow is isomorphic to a subflow of a Bernoulli flow. In fact, if $X = (\Omega,T)$, then Ω may be identified with the subset

$$\Omega' = \{(\omega_1,\omega_2,\omega_3, \cdots): \omega_{n+1} = T\omega_n\}$$

of $\Omega^\infty = \Omega \times \Omega \times \Omega \times \cdots$, and so X is a subflow of (Ω^∞,T).

Definition II.2. *A flow is of* finite type *if it is a subflow of a Bernoulli flow of finite type.*

(B) Ergodic Flows. We say that a flow X is *ergodic* if every T-invariant proper closed subset of Ω_X is nowhere dense. The family of ergodic flows will be denoted by $\mathscr{E}$. A straightforward argument shows that X is ergodic if and only if the set of points of Ω_X whose orbits are not dense in Ω_X form a set of the first category. Another criterion for ergodicity is that for any two open sets $A,B \subset \Omega_X$ there exists a power T^n of T with $T^nA \cap B$ non-empty. The analogy with ergodicity in the measure-theoretic setup is apparent. Moreover, suppose X supports an ergodic process (X,μ) such that the support of the measure μ is the entire space Ω_X. Then since a closed subset of Ω_X with μ-measure 0 must be nowhere dense, it follows that a T-invariant proper closed subset of Ω_X must be nowhere dense. Hence X is ergodic as a flow. This fact enables us to construct numerous ergodic flows. In particular, we recognize that Bernoulli flows are ergodic, a fact which is also otherwise evident.

The class $\mathscr{E}$ is not closed under multiplication but it is closed under passage to factors. To see this, suppose $X \xrightarrow{\alpha} Y$ and that A, B are open sets in Ω_Y. Then $\alpha^{-1}(A)$ and $\alpha^{-1}(B)$ are open in Ω_X, and if $T^n\alpha^{-1}(A) \cap \alpha^{-1}(B)$ is non-empty, so too is $T^nA \cap B$.

(C) Weakly Mixing Flows. We recall from Part I, §3 (E), that there are four equivalent definitions for the notion of a weakly mixing process. Two of these are related to ergodicity: $X \times X \in \mathscr{E}$ and $X \times \mathscr{E} \subset \mathscr{E}$. We choose the analogue of the first of these as the definition of a weakly mixing flow. Precisely, we say X is a *weakly mixing* flow if $X \times X$ is an ergodic flow. The class of weakly mixing flows will be denoted by $\mathscr{W}$. We do not know whether the product of a weakly mixing flow and any ergodic flow is ergodic. (Later we shall see that a special case of this is true: The product of a minimal flow and a weakly mixing flow is ergodic.) Nevertheless we have

PROPOSITION II.3. *If X is weakly mixing, then any power $X \times X \times X \times \cdots \times X$ is ergodic.*

Proof: X as a factor process of $X \times X$ is ergodic. Let $N(A,B)$ denote the set of positive integers n for which $T^nA \cap B$ is non-empty, where A and B denote open subsets of Ω_X. By the ergodicity of X, $N(A,B)$ is always non-empty. We shall show that, in fact, if A,B,C,D are open subsets of Ω_X, then there exist open sets $E,F \subset \Omega_X$ with $N(A,B) \cap N(C,D) \supset N(E,F)$. This is a consequence of the weak mixing property. Namely, since $X \times X$ is ergodic, there exists a $k \geqslant 0$ such that $T^k(A \times B) \cap (C \times D)$ is non-empty. That means that $(A \times B) \cap T^{-k}(C \times D) \neq \varnothing$, or $A \cap T^{-k}C \neq \varnothing$ and $B \cap T^{-k}D \neq \varnothing$. These sets are non-empty and open; call them E and F respectively. If $n \in N(E,F)$, then $A \cap T^{-k}C \cap T^n(B \cap T^{-k}D) \neq \varnothing$. But then $A \cap T^nB \neq \varnothing$, $T^{-k}C \cap T^n(T^{-k}D) \neq \varnothing$. The second of these implies $C \cap T^nD \neq \varnothing$ and so $N(E,F) \subset N(A,B) \cap N(C,D)$. As a result of this, $\bigcap_i N(A_i,B_i) \neq \varnothing$ for any finite collection of open sets $A_1, \cdots, A_n, B_1, \cdots, B_n$. This, however, yields $N(A_1 \times \cdots \times A_n, B_1 \times \cdots \times B_n) \neq \varnothing$ which implies that the n-fold product $X \times \cdots \times X$ is ergodic.

Suppose the flow X can be equipped with a T-invariant measure μ whose support is all of Ω_X, and such that the process (X,μ) is a weakly mixing process. Then $(X \times X,\mu \times \mu)$ is an ergodic process; hence $X \times X$ is an ergodic flow, and X itself is weakly mixing.

This condition again allows one to construct a wide class of examples of weakly mixing flows. For example, let G be a compact abelian group and suppose that T is an endomorphism of G. Denote by T^* the adjoint endomorphism of the dual group $\hat{G}$. Harmonic analysis on G shows (see [8]) that if m_G denotes Haar measure on G, then (G,m_G,T) is a weakly mixing process if and only if T^* has no finite orbit in $\hat{G}$. (If G is a torus, then T is represented by an integer matrix and the condition is that no root of unity occur among its eigenvalues.) With this condition, then, it follows that (G,T) is a weakly mixing flow.

(D) $\mathscr{F}$-flows. We say that a flow $X = (\Omega,T)$ belongs to the class $\mathscr{F}$ (for "fixed point"), or that X is an $\mathscr{F}$-flow, if it satisfies the two conditions:

(i) Each of the flows (Ω,T^m), $m = 1,2,3, \cdots$ is ergodic;
(ii) The totality of all fixed points of all the powers T^m (i.e., $\{\omega$: for some m, $T^m\omega = \omega\}$) is dense in Ω.

Note that $\mathscr{B} \subset \mathscr{F}$. For, if $(\Omega,T) \in \mathscr{B}$, then, as one sees readily, $(\Omega,T^m) \in \mathscr{B}$. Hence all (Ω,T^m) are ergodic. Secondly, a periodic sequence in $\Omega = \Lambda^\infty$ is a fixed point of some power of T, and the periodic sequences are dense in Ω.

Another class of examples of $\mathscr{F}$-flows are flows of the form (G,T), where, as in the preceding paragraph, T is an endomorphism of the abelian group G. This time, in addition to the condition that T^* have no finite orbit, we assume that G is a torus. By the first condition (G,T) is ergodic, and, since (G,T^m) satisfies the same condition, each (G,T^m) is ergodic. Suppose G is an r-dimensional torus, so that T is given by an $r \times r$ integer matrix $T \sim (a_{ij})$. Let (α_i) denote a point of G with rational coordinates. We can write $\alpha_i = p_i/q$ with $p_1, \cdots, p_r, q$ integers. (α_i) is a fixed point of T^m if $T^m(p_i) \equiv (p_i) \pmod q$. The condition that such an integer m exist is that det (a_{ij}) *be relatively prime to* q. (Then $T \in \mathrm{GL}(r,\mathbf{Z}/q\mathbf{Z})$ which is finite, and some power T^m is the identity.) The set of q which are relatively prime to det (a_{ij}) is infinite, which implies that in G the set of all fixed points of all powers of T is dense.

The simplest cases in this last category are obtained by taking $G = \mathbf{R}/\mathbf{Z}$, the circle group written additively, and setting $Tx = tx$ for some integer $t \neq 0$. The flow is almost identical with the Bernoulli flow obtained by setting $\Lambda = \{0,1, \cdots, t-1\}$, since

$$t \sum_1^\infty a_j t^{-j} = \sum a_{j+1} t^{-j} \pmod 1$$

where $a_j \in \Lambda$. The flows are not quite identical, since Λ^∞ is disconnected and $\mathbf{R}/\mathbf{Z}$ is connected.

(E) Minimal Flows. If a flow contains no proper subflows, it is called *minimal* (see [6]). The class of minimal flows will be denoted by $\mathscr{M}$. The closure of the orbit of a point in Ω_X is always a T-invariant subset. If X is minimal, it must coincide with Ω_X. Hence a flow is minimal if and only if every point has a dense orbit. In particular,

PROPOSITION II.4. *A minimal flow is ergodic.*

Every flow possesses minimal subflows. If (Ω,T) is a flow, the closed T-invariant subsets $\Delta \subset \Omega$ with the property that (Δ,T) is minimal are called *minimal sets.* Since every flow is a subflow of a Bernoulli flow, to study minimal flows, it suffices to study the minimal sets for all Bernoulli flows. We shall study a restricted version of this problem in Part III.

Products of minimal flows need not be minimal. In fact, if X is a non-trivial minimal flow, then $X \times X$ cannot be minimal inasmuch as the diagonal of $\Omega_X \times \Omega_X$ is T-invariant. On the other hand, a factor of a minimal flow is minimal. This follows immediately from the definition. It is also true that limits of inverse systems (or "inverse limits") of minimal flows are minimal. Suppose we have a system of minimal flows $\{X_\alpha\}$ indexed by a partially ordered set $\{\alpha\}$ and maps $X_\beta \xrightarrow{\pi_\alpha^\beta} X_\alpha$ for $\beta > \alpha$. Write $X_\alpha = (\Omega_\alpha,T)$; then $\lim X_\alpha = (\lim \Omega_\alpha,T)$, where $\lim \Omega_\alpha$ is the subset of $\prod \Omega_\alpha$ of points $\tilde{\omega} = \{\omega_\alpha\}$ satisfying $\pi_\alpha^\beta \omega_\beta = \omega_\alpha$. There is a natural map π_β: $\lim \Omega_\alpha \to \Omega_\beta$. One can show that a sequence $\{\tilde{\omega}_n\}$ in $\lim \Omega_\alpha$ is dense if and only if $\{\pi_\beta(\tilde{\omega}_n)\}$ is dense in Ω_β for each β. From this it follows that each orbit in $\lim X_\alpha$ is dense if and only if this is so in each X_α. Hence the inverse limit of minimal flows is minimal.

(F) Semi-simple Flows. Let $X = (\Omega,T)$ be a flow. If Ω is the union of all the minimal sets of X, then we say that X is *semi-simple.* In other words, X is semi-simple if $\Omega = \bigcup_\alpha \Omega_\alpha$ where the Ω_α are non-overlapping T-invariant closed sets such that (Ω_α,T) is minimal for each α. An example of a flow that is not semi-simple is $([0,1],T)$ where $Tx = x^2$. The only minimal sets for this flow are $\{0\}$, $\{1\}$. We denote the family of semi-simple flows by $\mathscr{S}$. In [6] this property is referred to as *pointwise almost periodicity* of T. As with minimality it is clear that factors of semi-simple flows are semi-simple.

PROPOSITION II.5. *If a flow is semi-simple and ergodic it is minimal. In symbols,* $\mathscr{S} \cap \mathscr{E} = \mathscr{M}$.

Proof: If Ω_X decomposes into more than one minimal set then no orbit can be dense.

One reason for studying semi-simple flows is that they give rise to a class of sequences that possess a property that we shall call *recurrence.* This property is a natural generalization of (Bohr) almost periodicity.

In the following Λ is a compact metric space.

Definition II.3. *Let $\{\xi(n)\}$ be a Λ-sequence defined for $n = 1,2,\cdots$. For any open set $V \subset \Lambda^k$, where k is some positive integer, form the sequence $n_1(V) < n_2(V) < \cdots < n_j(V) < \cdots$ (which may be empty) of values of n for which*

$$(\xi(n+1), \xi(n+2), \cdots, \xi(n+k)) \in V .$$

The sequence $\{\xi(n)\}$ *is* recurrent *if for each V, the sequence* $\{n_i(V)\}$ *is either empty or else it satisfies*

$$n_{i+1}(V) - n_i(V) < K(V) < \infty . \tag{1}$$

A sequence of integers satisfying (1) is said to be *relatively dense.* Loosely speaking, we say that a Λ-sequence $\{\xi(n)\}$ is recurrent if everything that occurs once in the sequence occurs for a relatively dense set of times.

We leave to the reader the proof of the following proposition.

PROPOSITION II.6. *A sequence defined for a semi-simple flow is recurrent. On the other hand, the flow defined by a recurrent sequence is minimal.*

(G) Distal Flows. We refer the reader to [2] and [5] for the details concerning distal flows. We shall content ourselves with a rapid outline of the theory.

A flow $X = (\Omega, T)$ is *distal* if the relationships $\lim T^{n_i}\xi = \zeta$, $\lim T^{n_i}\eta = \zeta$, for $\xi,\eta,\zeta \in \Omega$, imply $\xi = \eta$. Equivalently, X is distal if $\xi \neq \eta$ implies that $\inf_n D(T^n\xi, T^n\eta) > 0$, where $D(\cdot,\cdot)$ denotes a metric on Ω. We denote the class of distal flows by $\mathscr{D}$.

The powers $\{T^n,\ n \geqslant 0\}$ form a subsemigroup of the semigroup Ω^Ω of *all* maps (continuous or not) of $\Omega \to \Omega$. With the product topology, Ω^Ω becomes a compact topological space. Let Γ be the closure of $\{T^n\}$ in this compact space. In general Γ may be shown to be a semigroup. When the flow is distal, it is easy to see that $\gamma\gamma_1 = \gamma\gamma_2$, for $\gamma,\gamma_1,\gamma_2 \in \Gamma$, implies $\gamma_1 = \gamma_2$. For a compact semigroup, this condition implies it is a group. This shows, among other things, that T is invertible, $T^{-1} \in \Gamma$, and $\inf_{n\geqslant 0} D(T^n\xi, T^n\eta) = \inf_{-\infty<n<\infty} D(T^n\xi, T^n\eta)$. Hence a distal flow is bilateral. (In particular, the theory of [5] which is stated for groups of transformations applies.)

One consequence of the fact that Γ is a group is

PROPOSITION II.7. $\mathscr{D} \subset \mathscr{S}$**.**

Proof: Let $X = (\Omega, T) \in \mathscr{D}$. In each $\xi \in \Omega$, Γ_ξ is the closure of the orbit of ξ. Now suppose $\eta \in \Gamma_\xi$; then $\eta = \gamma\xi$ for $\gamma \in \Gamma$, and so $\xi = \gamma^{-1}\eta \in \Gamma_\eta$. It follows that Γ_ξ is a minimal set for (Ω, T). Since each $\xi \in \Gamma_\xi$, Ω is a union of minimal sets.

From the definition of distal flows we conclude easily that $\mathscr{D} \times \mathscr{D} \subset \mathscr{D}$. In particular $X \times X$ is semi-simple whenever X is distal. Conversely, if $X \times X$ is semi-simple, X must be distal. For the diagonal of $\Omega_X \times \Omega_X$ is T-invariant and can intersect a minimal set only if it contains the set. If we had $T^{n_i}(\xi,\eta) \to (\zeta,\zeta)$, then the minimal set to which (ξ,η) belongs must be contained in the diagonal; hence $\xi = \eta$. This shows that X is distal. Thus $\mathscr{D} = \{X\colon X \times X \in \mathscr{S}\}$. We shall presently establish still another characterization:

$$\mathscr{D} = \{X\colon X \times \mathscr{S} \subset \mathscr{S}\} .$$

(Note that these two characterizations relate $\mathscr{D}$ to $\mathscr{S}$ in the same way that the weakly mixing processes were related to ergodic processes in Part I, §3. Namely, $\mathscr{M} = \{X: X \times X \in \mathscr{E}\} = \{X: X \times \mathscr{E} \subset \mathscr{E}\}$.)

Factors of distal flows are distal. In fact, if $X \xrightarrow{\alpha} Y$ and $X \times X \in \mathscr{S}$, then $Y \times Y$ is a factor of a semi-simple flow, so it belongs to $\mathscr{S}$. Under certain circumstances we can argue in the opposite direction and conclude that if $X \xrightarrow{\alpha} Y$ and Y is distal, then X is distal. This is in the case of a *group extension.*

Definition II.4. *Let $X = (\Omega,T)$ be a flow and suppose G operates on Ω in such a way that $(\omega,g) \to \omega g$ defines a continuous map from $\Omega \times G \to \Omega$ with the property that if $\omega g = \omega$ for some $\omega \in \Omega$, then g is the identity. Assume, in addition, that the action of F commutes with that of T: $T(\omega g) = (T\omega)g$. Then the orbit space Ω/G determines a flow $Y = X/G = (\Omega/G,T)$, where $T(\omega G) = (T\omega)G$. Y is then a factor of X and we say that X is a* group extension *of Y.*

PROPOSITION II.8. *A group extension of a distal flow is distal.*

Proof: Suppose $T^{n_i}\xi_1 \to \zeta$, $T^{n_i}\xi_2 \to \zeta$, with $\xi_1,\xi_2,\zeta \in \Omega$. We consider two cases. Supoose first that for some g, $\xi_2 = \xi_1 g$. Then, by continuity, $\zeta = \zeta g$ and g must be the identity, which implies $\xi_1 = \xi_2$. If this is not the case, then $\xi_1 G \neq \xi_2 G$. Let α denote the map $\xi \to \xi G \in \Omega/G$. If Y is distal, then since $\alpha(\xi_1) \neq \alpha(\xi_2)$, $\lim T^{n_i}\alpha(\xi_1) \neq \lim T^{n_i}\alpha(\xi_2)$. But both limits are $\alpha(\zeta)$ since α is a continuous map. This contradiction proves that $\xi_1 = \xi_2$. Hence X is distal.

Another operation preserving distality is passage to inverse limits. This is immediate from the definition. In fact, an inverse limit of flows is a subflow of the product flow. But distality is preserved both for products and for subflows.

Recapitulating, the class $\mathscr{D}$ is closed under passage to products, subflows, factor flows, inverse limits, and group extensions. The main result of [5] is a restricted converse to this. We restate it here in a weaker form that is still sufficiently precise for our purpose. Note that since distal flows are semi-simple, they decompose into minimal flows, and the latter, as subflows, are again distal. In a certain sense, therefore, it suffices for the study of $\mathscr{D}$ to consider $\mathscr{D} \cap \mathscr{M}$.

PROPOSITION II.9. *The class $\mathscr{D} \cap \mathscr{M}$ of minimal distal flows coincides with the smallest class $\mathscr{C}$ of flows satisfying*

(a) *The trivial (one point) flow is in $\mathscr{C}$.*
(b) *Factors of flows in $\mathscr{C}$ are in $\mathscr{C}$.*
(c) *A group extension of a flow in $\mathscr{C}$ is in $\mathscr{C}$ provided it is minimal.*
(d) *Inverse limits of flows in $\mathscr{C}$ are in $\mathscr{C}$.*

As one application of this proposition let us prove:

PROPOSITION II.10. $\mathscr{D} \times \mathscr{S} \subset \mathscr{S}$.

Proof: Clearly it suffices to show that the product of a minimal distal flow and a semi-simple flow is semi-simple. So let $\mathscr{C}' = \{X: X \times \mathscr{S} \subset \mathscr{S}\}$.

One sees readily that $\mathscr{C}'$ satisfies conditions (a), (b) and (d). Let us show that it satisfies (c). This will prove the proposition, for it will show that $\mathscr{C}' \supset \mathscr{M} \cap \mathscr{D}$ and hence $\mathscr{D} \cap \mathscr{M} \times \mathscr{S} \subset \mathscr{S}$.

Suppose then that $Y = X/G$ and that $Y \times \mathscr{S} \subset \mathscr{S}$. Let Z be a particular flow in $\mathscr{S}$; we would like to show that $X \times Z$ is semi-simple. Now $X \times Z$ is clearly a group extension of $Y \times Z$, and by hypothesis, $Y \times Z$ is semi-simple. It suffices therefore to show that a group extension of a semi-simple flow is semi-simple. For this it is clearly sufficient to show that a group extension of a minimal flow is semi-simple. Suppose then that $U = V/G$, where U is minimal. Let Δ be a minimal set in Ω_V. One such minimal set certainly exists. Clearly each of the sets Δg is again minimal. Let α denote the map from Ω_V to Ω_U sending ω into ωG. Clearly $\alpha(\Delta)$ is a T-invariant closed subset of Ω_U. Hence by the minimality of U, $\alpha(\Delta) = \Omega_U$. But this means that, for every $\omega \in \Omega_V$, ωG intersects Δ. In other words

$$\Omega_V = \bigcup_{g \in G} \Delta g$$

and each point of Ω_V belongs to a minimal set. This proves the proposition.

(H) Kronecker Flows. Let G be a compact abelian group, τ an element of G whose powers are dense in G. Defining T by $Tg = \tau g$ we obtain a flow (G,T). A flow obtained in this manner is called a *Kronecker* flow, and the class of these is denoted by $\mathscr{K}$.

$\mathscr{K} \subset \mathscr{D}$. In fact a flow in $\mathscr{K}$ is a group extension of the trivial flow. Conversely, every minimal group extension of the trivial flow is a Kronecker flow.

A sequence defined for a Kronecker flow is easily seen to be (Bohr) almost periodic. Conversely, the flow defined by an almost periodic sequence is a Kronecker flow.

At this point let us summarize the relationships between the various classes of flows. We have

(i) $\mathscr{B} \subset \mathscr{W} \cap \mathscr{F} \subset \mathscr{W} \cup \mathscr{F} \subset \mathscr{E}$;
(ii) $\mathscr{K} \subset \mathscr{D} \cap \mathscr{M} \subset \mathscr{D} \cup \mathscr{M} \subset \mathscr{S}$;
(iii) $\mathscr{W} = \{X : X \times \mathscr{E} \subset \mathscr{E}\} = \{X : X \times X \in \mathscr{E}\}$;
(iv) $\mathscr{D} = \{X : X \times \mathscr{S} \subset \mathscr{S}\} = \{X : X \times X \in \mathscr{S}\}$;
(v) $\mathscr{S} \cap \mathscr{E} = \mathscr{M}$;
(vi) $\mathscr{D} \cap \mathscr{W} = \mathscr{M} \cap \mathscr{F} = \{\text{trivial flow}\}$.

We have proven all but (vi). Suppose X is distal and weakly mixing. Then $X \times X$ is both semi-simple and ergodic. By (v) $X \times X$ is minimal, which can only occur if X is trivial. Next suppose that X is both minimal and an $\mathscr{F}$-flow. As an $\mathscr{F}$-flow it possesses finite orbits; since there are dense, Ω_X must be finite. But then an N exists with $T^N\omega = \omega$ for each $\omega \in \Omega_X$, and (Ω_X, T^N) is not ergodic unless Ω_X reduces to a single point.

Finally, all the classes of flows but $\mathscr{B}$ are closed under passage to factor flows. In addition, $\mathscr{B}$, $\mathscr{W}$ and $\mathscr{D}$ are closed under formation of products.

4. Disjointness Relations. We begin with an analogue of Theorem I.1.

THEOREM II.1. *If two flows are disjoint, one of them must be minimal.*

Proof: Let $X = (\Omega_X, T)$, $Y = (\Omega_Y, T)$ and suppose $A \subset \Omega_X$ is T-invariant and closed, and $B \subset \Omega_Y$ is T-invariant and closed. If A and B are proper subsets, then $(A \times \Omega_Y) \cup (\Omega_X \times B)$ is a T-invariant closed subset of $\Omega_X \times \Omega_Y$ which projects onto Ω_X and Ω_Y under the projections π_X, π_Y.

The following result will be referred to repeatedly in Part III.

THEOREM II.2. $\mathscr{F} \perp \mathscr{M}$.

Proof: Let $X \in \mathscr{F}$, $Y \in \mathscr{M}$, and suppose $\Delta \subset \Omega_X \times \Omega_Y$ is a closed T-invariant set satisfying $\pi_X(\Delta) = \Omega_X$, $\pi_Y(\Delta) = \Omega_Y$. Consider the flows $Y_n = (\Omega_Y, T^n)$. Although $Y_1 = Y$ is minimal, it is possible that when $n > 1$, Y_n is not minimal. Choose a sequence of natural numbers $\{n_k\}$ such that (a) $n_k | n_{k+1}$ and (b) every integer m divides some n_k. An inductive procedure enables us to choose a sequence of subsets $\{\Omega_Y^k\}$ of Ω_Y satisfying: (a) $\Omega_Y^k \supset \Omega_Y^{k+1}$, (b) Ω_Y^k is a minimal set for Y_{n_k}. The significance of the condition $n_k | n_{k+1}$ is that a T^{n_k}-invariant set is also $T^{n_{k+1}}$-invariant. Hence, given Ω_Y^k which is $T^{n_{k+1}}$-invariant, it is possible to find a subset Ω_Y^{k+1} which is minimal $T^{n_{k+1}}$-invariant. Finally we set $\Omega_Y^\infty = \bigcap_{k=1}^{\infty} \Omega_Y^k$.

Fix k momentarily and form the sets

$$\Lambda_{ki} = \{\xi \in \Omega_X\text{: for some } \eta \in T^i\Omega_Y^k,\ (\xi,\eta) \in \Delta\} \tag{2}$$

where i ranges from 0 to $n_k - 1$. Since $T^{n_k}\Omega_Y^k \subset \Omega_Y^k$, $\bigcup_{i=0}^{n_k-1} T^i\Omega_Y^k$ is T-invariant, and hence it is all of Ω_Y, Y being minimal. It follows that $\bigcup_{i=0}^{n_k-1} \Lambda_{ki} = \Omega_X$. Since each Λ_{ki} is open it follows from this that at least one of them, say Λ_{kj}, has a non-empty interior. From (2), however, it follows that $T^{n_k}\Lambda_{kj} \subset \Lambda_{kj}$, and this contradicts the ergodicity of (Ω_X, T^{n_k}) (see §3 (D)) unless $\Lambda_{kj} = \Omega_X$. It follows that $\Lambda_{k0} = \Omega_X$. For, $T^{n_k-j}\Lambda_{kj} \subset \Lambda_{k0}$, so that $\Lambda_{k0} \supset T^{n_k-j}\Omega_X$. Now since X is an $\mathscr{F}$-flow $T(\Omega_X) = \Omega_X$, since all fixed points of powers of T lie in $T(\Omega_X)$. Hence $\Lambda_{k0} = \Omega_X$. In other words, for each $\xi \in \Omega_X$, there exists $\eta \in \Omega_Y^k$ with $(\xi,\eta) \in \Delta$. Since this is true for each k, we may conclude that, for a given $\xi \in \Omega_X$, there exists $\eta \in \Omega_Y^\infty$, with $(\xi,\eta) \in \Delta$.

Suppose in the foregoing we choose $\xi \in \Omega_X$ to be a fixed point of some T^m. Find $\eta \in \Omega_Y^\infty$ with $(\xi,\eta) \in \Delta$. For k sufficiently large, $m | n_k$, and $T^{n_k}\xi = \xi$. On the other hand $\{T^{nn_k}\eta\}$ represents the orbit of $\eta \in \Omega_Y^\infty \subset \Omega_Y^k$ for the flow (Ω_Y^k, T^{n_k}), and hence is dense in Ω_Y^k. Since Δ is invariant under T^{n_k}, we find that $(\xi,\eta') \in \Delta$ for every $\eta' \in \Omega_Y^k$. In particular, $(\xi,\eta') \in \Delta$ for every $\eta' \in \Omega_Y^\infty$.

Now fix $\eta' \in \Omega_Y^\infty$. We have seen that for each fixed point ξ of some T^m, $(\xi,\eta') \in \Delta$. But these are dense in Ω_X; hence, $\Omega_X \times \eta' \subset \Delta$. Now the set of η' with this property is closed and T-invariant. Since Y is minimal, it follows that $\Omega_X \times \Omega_Y \subset \Delta$ which is the conclusion sought after.

COROLLARY: $\mathscr{B}^\perp = \mathscr{F}^\perp = \mathscr{M}$.

Proof: We have just shown that $\mathscr{M} \subset \mathscr{F}^\perp$. On the other hand, since $\mathscr{F}$ contains non-minimal flows (in fact all are non-minimal by §3 (H) (vi)), $\mathscr{F}^\perp$ must consist of minimal flows. Hence $\mathscr{F}^\perp = \mathscr{M}$. For the same reason $\mathscr{B}^\perp \subset \mathscr{M}$. Since $\mathscr{B} \subset \mathscr{F}$, $\mathscr{B}^\perp \supset \mathscr{F}^\perp$; hence $\mathscr{B}^\perp = \mathscr{M}$.

5. Weakly Mixing Flows. There are two ways of defining weak mixing in terms of ergodicity. We chose the weaker definition:

$$\mathscr{W} = \{X: X \times X \in \mathscr{E}\} .$$

It is not known whether this implies the stronger property: $\mathscr{W} \times \mathscr{E} \subset \mathscr{E}$, i.e., that the product of a weakly mixing flow with any ergodic flow is ergodic. The following proposition shows that this is true at least for products of weakly mixing flows with minimal flows.

PROPOSITION II.11. $\mathscr{W} \times \mathscr{M} \subset \mathscr{E}$.

Proof: Let $X \in \mathscr{W}$, $Y \in \mathscr{M}$. To prove the ergodicity of $X \times Y$ we must show that for open sets $A, A' \subset \Omega_{X_n}$, $B, B' \subset \Omega_Y$, there exists an integer $n \geq 0$ with $T^n(A \times B) \cap (A' \times B') \neq \emptyset$. Equivalently n must be found with $T^nA \cap A' \neq \emptyset$ and $T^nB \cap B' \neq \emptyset$. By Proposition II.3, $X \times X \times \cdots \times X$ is ergodic for any finite product. This implies that for a finite family of open sets $A_1,A_2, \cdots ,A_m;A'_1,A'_2, \cdots ,A'_m$ in Ω_X, there exists n such that each $T^nA_i \cap A'_i \neq \emptyset$. In particular, for arbitrarily large m, there exists an n with $T^nA \cap A' \neq \emptyset$, $T^{n+1}A \cap A' \neq \emptyset, \cdots, T^{n+m}A \cap A' \neq \emptyset$. In other words, the sequence $\{n:T^nA \cap A' \neq \emptyset\}$ has arbitrarily large blocks of consecutive integers. To prove the proposition, it will suffice to show that $\{n:T^nB \cap B' \neq \emptyset\}$ is relatively dense (see Definition II.3). But, in fact, for any $\eta \in B$, $\{n:T^n\eta \in B'\}$ is relatively dense by virtue of the fact that $\{T^n\eta\}$ is recurrent (Definition II.3 and Proposition II.6). Here we have used the fact that $\{n:T^n\eta \in B'\}$ is non-empty, which is a consequence of the fact that $\{\eta:T^n\eta$ never belongs to $B'\}$ is a closed T-invariant subset of Ω_Y and Y is minimal.

THEOREM II.3. $\mathscr{W} \perp (\mathscr{D} \cap \mathscr{M})$.

Proof: According to Proposition II.9, it will be sufficient to prove that if

$$\mathscr{C} = \{X: X \perp \mathscr{W}\}$$

then (a) the trivial flow is in $\mathscr{C}$, (b) factors of flows in $\mathscr{C}$ are in $\mathscr{C}$, (c) a group extensions of a flow in $\mathscr{C}$ is in $\mathscr{C}$ if it is minimal, (d) inverse limits of flows in $\mathscr{C}$ are in $\mathscr{C}$. Now (a), (b), and (d) are always valid for the class of flows disjoint from an arbitrary class. To prove the theorem it therefore suffices to show that a group extension of a flow disjoint from $\mathscr{W}$ is itself disjoint from $\mathscr{W}$ if it is minimal.

Suppose then that $W \in \mathscr{W}$, that $Y \perp \mathscr{W}$ where $Y = X/G$, and that X is a minimal flow. We wish to show that $X \perp W$. Let $\Delta \subset \Omega_W \times \Omega_X$ with Δ closed, $T\Delta \subset \Delta$, $\pi_W(\Delta) = \Omega_W$, $\pi_X(\Delta) = \Omega_X$. Notice that G acts on $W \times X$ in a natural way with $(\omega,\xi,g) \rightarrow (\omega,\xi g)$, and this defines a continuous map of $\Omega_W \times$

$\Omega_X \times G \to \Omega_W \times \Omega_X$. Let α denote the natural map from Ω_X to $\Omega_Y = \Omega_X/G$ and let α' be the corresponding map from $\Omega_W \times \Omega_X$ to $\Omega_W \times \Omega_Y$. The set $\alpha'(\Delta)$ is a T-invariant closed subset of $\Omega_X \times \Omega_Y$ and $\pi_W(\alpha'(\Delta)) = \Omega_W$, $\pi_Y(\alpha'(\Delta)) = \alpha(\Omega_X) = \Omega_Y$. Since $Y \perp W$, we have $\alpha'(\Delta) = \Omega_W \times \Omega_Y$. From this it follows that for every $(\omega,\xi) \in \Omega_W \times \Omega_X$ there exists $g \in G$ with $(\omega,\xi g) \in \Delta$. This means that the product ΔG coincides with $\Omega_W \times \Omega_X$.

We shall show now that Δ itself coincides with $\Omega_W \times \Omega_X$, and this will complete the proof. Let V be a closed set in G with non-empty interior and consider the product ΔV. Since G is compact there exist $g_1, \cdots, g_r \in G$ with $\bigcup_{i=1}^{r} Vg_i = G$. Hence $\bigcup_{i=1}^{r} \Delta Vg_i = \Delta G = \Omega_W \times \Omega_X$. The sets ΔVg_i being closed, it follows that one of them, and therefore each of them, has non-empty interior. But ΔVg_i is T-invariant since Δ is T-invariant and T commutes with the action of G. However, we also know that $W \times X$ is ergodic by Proposition II.11; hence $\Delta Vg_i = \Omega_W \times \Omega_X$ and so $\Delta V = \Omega_W \times \Omega_X$. Now we can find a decreasing sequence of closed sets $\{V_n\}$ each with non-empty interior and with $\cap V_n = \{\text{identity}\}$. Since $\Delta V_n = \Omega_W \times \Omega_X$ we conclude that $\Delta = \Omega_W \times \Omega_X$. This proves the theorem.

We conclude with the statement of two open problems.

Problem F: Is it true that $\mathscr{W} \times \mathscr{E} \subset \mathscr{E}$?

Problem G: Describe the classes $\mathscr{M}^\perp$ and $\mathscr{D}^\perp$.

Part III. Properties of Minimal Sets.

1. Minimal Subsets of Groups. Let G be a compact abelian group and T an endomorphism of Ω such that $X = (G,T)$ is an $\mathscr{F}$-flow. In Part II, §3 (D), we showed that this is the case if G is a finite-dimensional torus, and T an endomorphism of G corresponding to an integer matrix with no roots of unity among its eigenvalues. It will also be the case if $G = \Lambda \times \Lambda \times \Lambda \times \cdots$, where Λ is a compact abelian group and T represents the shift operator.

With this assumption, $X \perp \mathscr{M}$ by Theorem II.2, and in particular, X is disjoint from its own minimal subflows. Using this, we shall be able to show that the minimal sets in G are "small" in a certain sense.

Definition III.1. *Let G be a topological group and T an endomorphism of G. A closed T-invariant subset A of G is said to be* restricted *if $AB = G$ for some closed T-invariant set $B \subset G$, implies $B = G$.*

Here AB is the collection of all products $\alpha\beta$, $\alpha \in A$, $\beta \in B$. Note that the notion of a restricted subset depends on the endomorphism T and the group structure of G.

Definition III.2. *A* basis *of a group G is a set B with the property that each element in G is a product of finitely many elements in B.*

LEMMA III.1. *Let G be a non-trivial compact metrizable group, and suppose that (G,T) is ergodic. Then a restricted set cannot be a basis of G.*

Proof: If A is a basis then $\cup A^n = G$. Then some A^n contains an open set (we assume A is closed) and since A is T-invariant, so is A^n. By ergodicity, $A^n = G$. But then $A^{n-1} = G$, $A^{n-2} = G, \cdots, A = G$, $\{e\} = G$, and so G is trivial if A is restricted.

The main result of this section is

THEOREM III.1. *If* (G,T) *is an* $\mathscr{F}$*-flow, then every minimal set for* (G,T) *is restricted.*

The proof of this theorem is based on the following lemma. This lemma is reminiscent of the familiar fact that if a number divides a product and is relatively prime to one of the factors, it necessarily divides the remaining factor.

LEMMA III.2. *Let* X *and* Y *be disjoint flows, and suppose there is a flow* Z *and a homomorphism* $X \times Z \xrightarrow{\pi} Y$. *Then for each* $\xi \in \Omega_X$, *the map* $\zeta \to \pi(\xi,\zeta)$ *takes* Ω_Z *onto* Ω_Y.

Proof: Let $W = X \times Z$. We have $W \xrightarrow{\pi_X^W} X$, $W \xrightarrow{\pi_Z^W} Z$, and $W \xrightarrow{\pi} Y$. Since $X \perp Y$, there must exist a homomorphism $W \xrightarrow{\gamma} X \times Y$ such that $\pi_X^W = \pi_X\gamma$, and $\pi = \pi_Y\gamma$. Then $\gamma(\xi,\zeta) = (\xi,\pi(\xi,\zeta))$. Now γ takes Ω_W onto $\Omega_X \times \Omega_Y$, so for each (ξ,η) there is a point $\omega \in \Omega_W$ with $\gamma(\omega) = (\xi,\eta)$. But this means that the equation $\eta = \pi(\xi,\zeta)$ has a solution for each pair (ξ,η).

Proof of Theorem III.1. Let M be a minimal set of the $\mathscr{F}$-flow (G,T). If M is not restricted, we may find a closed T-invariant set $B \subset G$ with $MB = G$ and $B \neq G$. Let $X = (M,T)$, $Y = (G,T)$, and $Z = (B,T)$. We have a map $\pi: M \times B \to G$ defined by $\pi(\mu,\beta) = \mu\beta$ which takes $M \times B$ onto G. Hence π defines a homomorphism of $X \times Z$ to Y. But $X \perp Y$ since X is minimal and Y is an $\mathscr{F}$-flow. By Lemma III.2, $\pi(\mu,B) = G$, or $\mu B = G$ for any $\mu \in M$. But G is a group, and this implies that $B = G$.

2. Relative Dimension, Hausdorff Dimension and Topological Entropy. Let $\Lambda = \{0,1,\cdots, a-1\}$, where a is a positive integer, set $\Omega = \Lambda \times \Lambda \times \Lambda \times \cdots$, and let T denote the shift transformation of Ω. Λ can be given a group structure in several ways; for each of these, Ω becomes a group and T an endomorphism of Ω. For each of these structures the results of the last section apply and a minimal set for (Ω,T) will be "restricted" with respect to the various group structures on Ω. The question arises whether it is possible to introduce a quantitative notion of the size of sets for which all minimal sets will be "small". There are several notions which suggest themselves, for all of which it is true that a "small" set is a restricted set. We shall, however, see that in spite of evidence to the contrary, minimal sets need not be small in any of these senses.

There is a correspondence between Ω and the unit interval which is almost one-one. Namely, we associate with the sequence $(\omega_1,\omega_2,\omega_3,\cdots)$ the real number $\sum_1^\infty \omega_n a^{-n}$. The operator T corresponds to the operator τ_a on $[0,1]$ which is defined by $\tau_a x = ax$ modulo 1. τ_a becomes continuous if we identify 0 and 1, i.e., if we take as its domain the additive group of reals modulo one. Calling this group K we obtain a flow (K,τ_a). For a set $A \subset \Omega$ we shall denote by A^* the corresponding set in K. A T-invariant set A in Ω is minimal for (Ω,T) if and only if A^* is a minimal set for (K,τ_a). In determining the "size" of a T-invariant set $A \subset \Omega$ we may consider interchangeably properties of A and properties of A^*.

One such measure of the size of a set in Ω or in K is given by its *Hausdorff dimension.* This exists for every subset B of K and we denote it by $D(B)$. Then $0 \leqslant D(B) \leqslant 1$. There is a related notion which is more useful for certain purposes and which we refer to as the *relative dimension.* Partition K into N equal intervals: $K = \bigcup_0^{N-1} [m/N,(m+1)/N]$, and let $\nu(B,N)$ be the number of these that contain points of B. If

$$\lim_{N\to\infty} \frac{\log \nu(B,N)}{\log N} \tag{1}$$

exists, we call it the *relative dimension* of B, and denote it by $d(B)$. It is easy to see that whenever $d(B)$ exists, it satisfies $d(B) \geqslant D(B)$. One establishes readily the relationship

$$\frac{\nu(B,N_2)}{\nu(B,N_1)} < 3 + \frac{N_2}{N_1},$$

which shows that if the limit (1) exists for a subsequence $\{N_k\}$ satisfying $N_{k+1}/N_k < M < \infty$, then the limit exists. In particular, it suffices to establish the existence of $\lim\limits_{n\to\infty} \dfrac{\log \nu(B,g^n)}{n}$ for some integer g.

Finally, there is a notion of size which is applicable to subsets of Ω. Let $A \subset \Omega$ and denote by $\mu(A,n)$ the number of Λ-valued n-tuples $(a_1,a_2,\cdots,a_n)$ which coincide with $(\omega_1,\omega_2,\cdots,\omega_n)$ for some point $\omega = (\omega_1,\omega_2,\cdots,\omega_n,\cdots) \in A$. If

$$\lim_{n\to\infty} \frac{\log \mu(A,n)}{n} \tag{2}$$

exists, it is called the *topological entropy* of A and denoted $\mathscr{E}(A)$. If A is T-invariant, then (A,T) is a flow. In this case $\mathscr{E}(A)$ always exists and it may be seen that $\mathscr{E}(A)$ coincides with the *topological entropy* of the flow (A,T) as defined in [1].

PROPOSITION III.1. *Let A be a T-invariant subset of Ω and let A^* denote the corresponding subset of K. Then $d(A^*)$ and $\mathscr{E}(A)$ exist and*

$$D(A^*) = d(A^*) = \frac{\mathscr{E}(A)}{\log a} \tag{3}$$

Remark: From this proposition it follows that the Hausdorff dimension of the classical Cantor set is $\log 2/\log 3$. For, the Cantor set corresponds to the set of sequences in Ω $(\Lambda = \{0,1,2\})$ in which 1 does not appear. This clearly has topological entropy $\log 2$.

Proof: By the T-invariance of A, it is clear that $\mu(A,n+m) \leqslant \mu(A,n) \times \mu(A,m)$. It is well known that this implies the existence of the limit in (2); hence $\mathscr{E}(A)$ exists. Consider now $\nu(A^*,a^n)$. The interval $[m/a^n,(m+1)/a^n]$ will contain a point of A^* if and only if the a-adic expansion of m, $m = \sum b_i a^{i-1}$, is such that $(b_1,b_2, \cdots, b_n)$ coincides with the initial n-block of a sequence in A. Hence $\nu(A^*,a^n) = \mu(A,n)$. Thus

$$\lim_{n\to\infty} \frac{\nu(A^*,a^n)}{n}$$

exists and coincides with $\mathscr{E}(A)$. By our remarks concerning relative dimension, we conclude that $d(A^*)$ exists and $d(A^*) = \mathscr{E}(A)/\log a$.

We must still show that the Hausdorff dimension of A^* coincides with $d(A^*)$. Since one always has $d(A^*) \geqslant D(A^*)$, we must show that $D(A^*) \geqslant d(A^*)$. Recall that $D(A^*) \geqslant \delta$ if, for every covering of A^* by intervals I, of sufficiently small length, $\sum |I_j|^\delta > c > 0$. We claim now that in applying this criterion, it suffices to consider coverings with intervals of the form $I_j = [m_j/a^{n_j},(m_j+1)/a^{n_j}]$. For, as one can see, it is possible to replace an arbitrary covering by one with intervals of this sort, thereby multiplying $\sum |I_j|^\delta$ by no more than $2a^\delta$. It will suffice therefore to show that if

$$A^* \subset \bigcup_{j=1}^{J} [m_j/a^{n_j},(m_j+1)/a^{n_j}]$$

and $\delta < d(A^*)$, then $\sum_1^J a^{-\delta n_j} \geqslant 1$.

We can restate this in terms of the set A. We denote by R' the collection of all n-tuples of elements of Λ, $R' = \bigcup_1^\infty \Lambda^n$. R' is a semigroup if we multiply by juxtaposition: $(a_1, \cdots, a_n)(b_1, \cdots, b_m) = (a_1, \cdots, a_n, b_1, \cdots, b_m)$. We denote by R the subset of R' consisting of n-tuples which occur as blocks in sequences of A. Thus $\mu(A,n)$ is the number of elements of R of length n. Notice that $\rho_1\rho_2 \in R$ implies that both ρ_1 and ρ_2 belong to R. We shall say that ρ' is *divisible* by ρ if $\rho' = \rho\rho_1$ for some $\rho_1 \in R'$. Also $l(\rho)$ will denote the length of ρ. With these preliminaries we may restate what must be proved as follows. If $\{\rho_i\}$ is a finite collection of elements of R such that each $\rho \in R$ of sufficiently great length is divisible by some ρ_i, and if $\delta < d(A^*)$, then $\sum a^{-\delta l(\rho_i)} \geqslant 1$.

Suppose to the contrary that $\sum a^{-\delta l(\rho_i)} < 1$. Then

$$\sum a^{-\delta l(\rho_{i_1}\rho_{i_2}\cdots\rho_{i_n})} < \infty,$$

where the sum is taken over the semigroup generated in R' by $\{\rho_i\}$. We now claim that there is a finite set of elements $\{\rho'_j\}$ such that every $\rho \in R$ can be expressed as a product $\rho = \rho_{i_1}\rho_{i_2} \cdots \rho_{i_n}\rho'_j$, for some sequence $\rho_{i_1}, \cdots, \rho_{i_n}$ in $\{\rho_i\}$ and for some ρ'_j. The reason is that each $\rho \in R$ is divisible by some ρ_i, $\rho = \rho_{i_1}\rho'$ with $\rho' \in R$, provided $l(\rho)$ is large enough. But from this it follows that

$$\sum_{R} a^{-\delta l(\rho)} < \infty .$$

In other words $\sum \mu(A,n)a^{-\delta n} < \infty$. However if $\delta < d(A^*) = \mathscr{E}(A)/\log a$, then for n sufficiently large, $\mu(A,n) > a^{\delta n}$ and the series in question must diverge. This proves the proposition.

3. Deterministic Flows. In [1] the analogy between "entropy" and "topological entropy" is developed. For a number of flows it may be shown that the flow X supports a measure μ such that the entropy of the process (X,μ) equals the topological entropy of X. In any case the former never exceeds the latter. The topological entropy of a product is the sum of the topological entropies of the factors, and the topological entropy of a flow is at least as great as that of any of its factors. This analogy suggests defining a class of "deterministic" flows by the condition that the topological entropy vanish. A deterministic flow has the property that any process supported by it is deterministic in the sense of Part I, §3. It may be that the converse is true as well.

Let A and B denote τ_a-invariant subsets of K, and let $A+B$ denote their sum, i.e., the set of all sums $\alpha+\beta$, $\alpha \in A$, $\alpha \in B$. Evidently, $\nu(A+B,N) \leqslant \nu(A,2N)\nu(B,2N)$. From this we find that $d(A+B) \leqslant d(A) + d(B)$.

PROPOSITION III.2. *If A is a τ_a-invariant closed subset of K such that (A,τ_a) is deterministic, then A is restricted.*

Proof: If (A,τ_a) is deterministic, then $d(A) = 0$. Hence $A+B=K$ implies $d(B) = 1$. Let B correspond to a T-invariant subset $B' \subset \Omega$. $d(B) = 1$ implies $\mathscr{E}(B') = \log a$. Now if B' is a proper subset of Ω, there is some n-tuple, for some n, of elements of Λ which does not occur in the sequences of B'. But then $\mu(B',n) \leqslant a^n - 1$ for some n, and so $\mu(B',nm) \leqslant (a^n-1)^m$, whence $\mathscr{E}(B') \leqslant \log(a^n-1)/n \log a < 1$. This proves the proposition.

There appears to be a certain amount of evidence to support the conjecture that minimal flows are deterministic. For one thing, a subset A of K is restricted if the flow (A,τ_a) is either minimal or deterministic. Secondly the class $\mathscr{B}^{\perp}$ of flows disjoint from Bernoulli flows coincides with $\mathscr{M}$, whereas the processes disjoint from Bernoulli processes are precisely the deterministic processes. Finally, the most familiar examples of minimal sets do correspond to deterministic flows. For instance, the recurrent sequences described by Robbins [12] lead to deterministic flows, and the minimal set described by Gottschalk [7] is deterministic. We shall see however that this evidence is misleading and that there exist minimal flows which are not deterministic, or, equivalently, minimal sets for (K,τ_a) with positive Hausdorff dimension.

THEOREM III.2. *There exist minimal sets with positive topological entropy.*

Proof: According to Proposition II.6, the flow defined by a recurrent sequence (see Definition II.3) is minimal. Moreover, if $\{\xi(n)\}$ is a Λ-valued sequence, and X is the flow it defines, it is possible to compute the topological entropy of X directly from $\{\xi(n)\}$. Namely, if H_n is the number of distinct n-tuples of symbols of the form $(\xi(l+1), \cdots, \xi(l+n))$ for some $l \geq 0$, then $\mu(A,n) = H_n$, A being the subset of Ω defined by the flow X. To construct a minimal flow with positive entropy we shall construct a recurrent sequence $\{\xi(n)\}$ for which H_n grows exponentially with n.

Decompose the natural numbers $J = \{1,2,3, \cdots\}$ into a disjoint union of arithmetic progressions: $J = \bigcup_1^\infty J_r$, $J_r = c_r + d_r J$, $J_r \cap J_s = \emptyset$ for $r \neq s$. For each r choose a number $a(r) \in \Lambda$, and define $\{\xi(n)\}$ by setting $\xi(n) = a(r)$ if $n \in J_r$. It is easily seen that the sequence $\{\xi(n)\}$ is recurrent. In fact, each m-tuple $(\xi(1), \xi(2), \cdots, \xi(m))$ recurs periodically.

The condition $J = \bigcup_1^\infty J_r$ implies that $\sum 1/d_r \leq 1$. We would expect equality, but one readily sees that the d_r may be chosen to increase as rapidly as we like. Moreover, it is not hard to see that the set of initial points $\{c_r\}$ may be made as dense as we like in J, simply by choosing $\{d_r\}$ to be increasing sufficiently rapidly. In particular, we may choose $\{J_r\}$ such that

$$\limsup c_r/r < 2 \,. \tag{4}$$

For such m there will be infinitely many r such that $c_{r+m} - c_{r+1} < 2m$. Denote this set of integers r by Q_m.

Now let $\{y_n\}$ denote a sequence of independent identically distributed random variables, each having for its value the symbols in Λ, all occurring with positive probability. For any choice of m symbols, $a_1, a_2, \cdots, a_m$, the probability is 0 that for each $r \in Q_m$, $(y_{r+1}, y_{r+2}, \cdots, y_{r+m}) \neq (a_1, a_2, \cdots, a_m)$. It follows that there exists a sample sequence $\{y_n\} = \{\eta_n\}$ such that for every m and every m-tuple $(a_1, a_2, \cdots, a_m)$, there is an $r \in Q_m$ with $(\eta_{r+1}, \eta_{r+2}, \cdots, \eta_{r+m}) = (a_1, a_2, \cdots, a_m)$.

With this sequence $\{\eta_n\}$ we form a recurrent sequence $\{\xi(n)\}$, setting $\xi(n) = \eta_r$ for $n \in J_r$. We find that, for every m-tuple $(a_1, a_2, \cdots, a_m)$, an r exists with $c_{r+m} - c_{r+1} < 2m$, and with $\xi(c_{r+j}) = a_j, j = 1, \cdots, m$. This means that there is a value of n such that in the sequence $\xi(n+1), \cdots, \xi(n+2m)$, the values $a_1, \cdots, a_m$ occur somewhere and in that order, though not necessarily consecutively (namely, take $n = c_{r+1}$). Now each fixed block $\xi(n+1), \cdots, \xi(n+2m)$ can give rise to at most $\binom{2m}{m} < 4^m$ m-tuples $(a_1, a_2, \cdots, a_n)$. It follows that the number H_{2m} for our sequence $\{\xi(n)\}$ exceeds $(a/4)^m$. As a result, if $a > 4$, the flow associated with $\{\xi(n)\}$ will have positive topological entropy. This proves our theorem.

We remark that this theorem has also been proven independently and in sharper form by F. Hahn and Y. Katznelson (as yet unpublished). They show that the flow X may be chosen to be strictly ergodic, and such that if μ is the unique invariant probability measure supported by X, then the entropy of the process (X,μ) is positive.

4. $\mathscr{M} \times \mathscr{M} \not\subset \mathscr{S}$. The class of deterministic flows introduced in the last section has the property that products of flows in the class are again in the class and factors of flows in the class are in it. The object of the remainder of this part is to construct a class with the same properties and containing the class of minimal flows. By Theorem III.2, the class of deterministic flows does not suffice for this purpose. The problem may also be described in another way. If we define a *deterministic sequence* as one which is defined for a deterministic flow, then it is easily seen that continuous functions of finitely many deterministic sequences $[\zeta(n) = f(\xi_1(n), \xi_2(n), \cdots, \xi_r(n))]$ are again deterministic sequences. This is a direct consequence of the closure properties of the class of deterministic flows. Since neither $\mathscr{M}$ nor $\mathscr{S}$ is closed under multiplication, there will exist functions of recurrent sequences which are not recurrent. The question arises, what properties are shared by sequences formed in this way?

The fact that products of minimal flows need not even be semi-simple, or equivalently, that a composite sequence $\{\xi_1(n), \xi_2(n)\}$ whose components are recurrent need not be recurrent, may be seen as follows. As before let Ω be the product $\Lambda \times \Lambda \times \Lambda \times \cdots$, with Λ a finite set.

PROPOSITION III.3. *Let $X = (A,T)$ be a subflow of (Ω,T) such that A is infinite. Then $X \times X$ is not semi-simple.*

Remark: In particular, if X is minimal and A is not finite, the hypotheses are fulfilled and $X \times X$ is a product of minimal flows which is not semi-simple. This proposition also shows that a flow of finite type cannot be distal unless it is periodic. For we recall from Part II, §3 (G), that X is distal if and only if $X \times X$ is semi-simple.

Proof: On Ω, and therefore on A, are defined coordinate functions x_n: $x_n(\omega_1, \omega_2, \cdots) = \omega_n \in \Lambda$. Suppose $X \times X$ were semi-simple. Then for any continuous function φ on $A \times A$, the sequence $\varphi(T^n\omega, T^n\omega')$ would be recurrent (Proposition II.6). Using this we shall show that on A, x_1 is a continuous function of $x_2, x_3, x_4, \cdots$. To show this it suffices to show that if $\omega_1, \omega_2 \in A$ and if $x_n(\omega_1) = x_n(\omega_2)$ for $n = 2,3,4, \cdots$, then $x_1(\omega_1) = x_1(\omega_2)$. But if $X \times X$ is semi-simple, then $x_n(\omega_1) - x_n(\omega_2)$ represents a recurrent sequence. If it vanishes for $n = 2,3,4, \cdots$, it must vanish for $n = 1$. Thus we may write $x_1 = F(x_2, x_3, x_4, \cdots)$ with F a continuous function on A. Now it is easily seen that since each x_i takes on only finitely many variables, $x_1 = F(x_2, x_3, \cdots, x_m)$. Since A is T-invariant, we have $x_n = F(x_{n+1}, x_{n+2}, \cdots, x_{n+m-1})$. However, there exist only finitely many sequences satisfying these conditions, and hence A must be finite. This proves the proposition.

An explicit example of a minimal set of finite type is

$$A = \text{set of sequences } \{\xi(n) = \operatorname{sgn} \sin(n\alpha + \theta)\}$$

where α is fixed and θ varies between 0 and 2π. When $n\alpha + \theta = \nu\pi$, ν an integer, we define sgn sin $(n\alpha + \theta)$ as either +1 or −1. That is, θ of the form $\nu\pi - k\alpha$ gives rise to two sequences $\{\xi(n)\}$, both of which are included in A. We leave it to the reader to verify that (A,T) is minimal. That the flow is not distal can be seen directly by considering limits of translates of $\{\xi^{\pm}(n)\}$ where

$$\xi^{\pm}(n) = \begin{cases} \operatorname{sgn} \sin n\alpha & n \neq 0 \\ \pm 1 & n = 0 . \end{cases}$$

The above example shows directly that sums and products of recurrent sequence are, in general, not recurrent. For $\xi^{+}(n) - \xi^{-}(n)$ and $\xi^{+}(n)\xi^{-}(n)$ are evidently not recurrent.

5. Binary Sequence Spaces. Our object in the next two sections is to construct a class of flows of finite type (Part II, §3 (A)), closed under products and passage to factor flows (that is, if X and Y are of finite type, $X \xrightarrow{\alpha} Y$ and X is in the class, then Y should also be in the class) and containing all minimal flows of finite type. It is clear that what is sought is a relatively restricted class of flows; the more restricted this class, the more information is to be had concerning products of minimal flows.

As yet we have only been successful in dealing with flows of finite type, but in all likelihood a similar situation will present itself in the general case.

In the present section, all the flows to be considered are subflows of (Ω_2,T), where Ω_2 is the space of all $\{0,1\}$-valued sequences $\xi = (\xi(1), \xi(2), \cdots, \xi(n), \cdots)$ and T is the shift transformation. It will be convenient to regard Ω_2 as a ring, addition and multiplication of sequences being carried out term by term, treating $\{0,1\}$ as the field with two elements. During the remainder of this section, the expression "invariant set" will refer to a closed, T-invariant subset of Ω_2. A minimal subset Ω_2 will also be understood to be an invariant set which is minimal for (Ω_2,T).

By Definition III.1, a subset A of Ω_2 is restricted (with respect to the additive structure of Ω_2) if $A + B = \Omega_2$ implies $B = \Omega_2$. Here A and B are invariant sets. It is clear that the sums of two restricted sets is restricted. We do not know if the product of restricted sets is restricted. Nevertheless, we have the following proposition which is a sharpening of Theorem III.1:

PROPOSITION III.4. *If A is a restricted subset of Ω_2 and M is a minimal set in Ω_2, then MA is restricted.*

Proof: Suppose B is an invariant set satisfying $MA + B = \Omega_2$. We show that $B = \Omega_2$. Consider the flows (M,T), (A,T), (B,T), and (Ω_2,T). By hypothesis, there is a map π of $M \times A \times B$ onto Ω_2 which defines a homomorphism of the product of the first three flows onto the last; namely $\pi(\mu,\alpha,\beta) = \mu\alpha + \beta$. Recalling that (M,T), being minimal, is disjoint from (Ω_2,T) which is an $\mathscr{F}$-flow, we are in a position to apply Lemma III.2. We conclude that for each $\mu \in M$, $\pi(\mu,A,B) = \Omega_2$, or, $\mu A + B = \Omega_2$. Fix μ and choose $\omega' \in \Omega_2$ in the form $\omega' = 1 - \mu + \mu\omega$ with ω an unspecified element of Ω_2. For each choice of ω there must exist $\alpha \in A$, $\beta \in B$ such that

$$\mu\alpha + \beta = 1 - \mu + \mu\omega . \tag{5}$$

Multiply both sides by $(1-\mu)(1-\beta)$ and recall that for $\xi \in \Omega_2$, $\xi^2 = \xi$, or $\xi(1-\xi) = 0$. We thus have

(6) $$(1-\mu)(1-\beta) = 0,\ (1-\beta) = \mu(1-\beta)\ .$$

Now multiply both sides of (5) by $(1-\beta)$:

$$(1-\beta)\mu\alpha = (1-\beta)\mu\omega$$

and since $(1-\beta)\mu = 1-\beta$,

$$(1-\beta)\alpha = (1-\beta)\omega$$

and

(7) $$\omega = \alpha + \beta(\omega - \alpha)\ .$$

(7) shows that every element in Ω_2 is a sum of an element of A and a multiple of an element of B:

$$\Omega_2 = A + B\Omega_2\ .$$

But $B\Omega_2$ is an invariant set and A_1 by hypothesis is restricted. We conclude that $B\Omega_2 = \Omega_2$. Now in Ω_2 the unit element 1 has a unique representation as a product: $1 = 1 \cdot 1$. It follows that $1 \in B$. We have shown that $MA + B = \Omega_2$ implies that $1 \in B$. But in addition, $MA + (B + C) = \Omega_2$ for any invariant set C. It follows that $1 \in B + C$. Take $C = \{1\}$; we conclude that $0 \in B$. Hence $0 \in B + C$ for any invariant set C. This means every invariant set C intersects $-B$. Thus B itself must intersect every invariant set, and hence it must contain every minimal set in Ω_2. So B contains every periodic sequence in Ω_2, and since these are dense, we infer that $B = \Omega_2$. This proves the proposition.

By repeated application of this proposition and the fact that sums of restricted sets are restricted, we conclude that any set in Ω_2 of the form

(8) $$A = \sum_{i=1}^{m} M_{i1}M_{i2} \cdots M_{il_i}$$

is restricted, the M_{ij} denoting minimal sets.

Definition III.3. *A sequence $\xi \in \Omega_2$ is* restricted *if it belongs to a restricted set. We denote by R the family of all restricted sequences.*

R is a T-invariant subset of Ω_2, but it is not closed. However if $\xi \in R$, then all limits of translates of ξ are in R. It is also clear that $R \neq \Omega_2$. To see this, observe that (Ω_2, T), being ergodic, has dense orbits. Hence there are sequences $\omega \in \Omega_2$ belonging to no invariant subsets other than Ω_2 itself. These points (which in fact constitute a residual set in Ω_2) are not restricted.

PROPOSITION III.5. *An invariant set $A \subset \Omega_2$ is restricted if and only if $A \subset R$.*

Proof: The necessity is clear, so we turn to the sufficiency. Suppose $A \subset R$ and $A + B = R$, A and B being invariant sets. As remarked above there exists $\omega \in \Omega_2$ whose orbit is dense in Ω_2. Write $\omega = \alpha + \beta$, $\alpha \in A$,

$\beta \in B$. Now $\alpha \in R$, hence $\alpha \in A'$, where A' is a restricted set. Hence $\omega \in A' + B$. The latter set is, however, invariant; hence $\Omega_2 = A' + B$ and so $B = \Omega_2$. This proves the proposition.

THEOREM III.3. *Let R_0 denote the subring of Ω_2 generated by all recurrent binary sequences. Then $R_0 \subset R$.*

Proof: Every element in R_0 is of form

$$\xi = \sum_{i=1}^{m} \xi_{i1}\xi_{i2} \cdots \xi_{il_i} ,$$

where the ξ_{ij} are recurrent sequences. Each ξ_{ij} belongs to a minimal set M_{ij} and hence ξ belongs to a set of the form (8). It follows that ξ is restricted.

R_0 is T-invariant and not closed. It also has the property that a limit of translates of an element of R_0 is again in R_0. Thus R_0 is a union of closed T-invariant sets.

COROLLARY. *R_0 is a set of the first category in Ω_2. R itself may not be a ring. By Proposition III.5, it is an additive subgroup. Also by Proposition III.4, it is a module over the ring R_0.*

6. R_0-Flows. The ring R_0 of binary sequences may be identified with a ring of subsets of the natural numbers. The latter ring we denote R^*_0. We shall make use of R^*_0 in studying general finite-valued sequences. Let Λ be a finite set and suppose $\{\xi(n)\}$ is a Λ-valued sequence.

Definition III.4. *A sequence ξ is R_0-*measurable *if the subsets of the form $\{n:\xi(n) \in \Delta\}$, Δ a subset of Λ, belong to R^*_0.*

THEOREM III.4. *(a) If a sequence is R_0-measurable, then so are all translates and limits of translates of the sequence; (b) Any function of finitely many R_0-measurable sequences ($\xi(n) = f(\xi_1(n), \cdots, \xi_m(n))$) is R_0-measurable; (c) Recurrent sequences are R_0-measurable.*

Proof: (a) and (b) follow by virtue of the fact that R^*_0 is a ring and R_0 has the invariance properties in question. For (c), suppose that $\{\xi(n)\}$ is recurrent. If we let $\eta(n) = 1$ or 0 according as $\xi(n) \in \Delta$ or not, then $\{\eta(n)\}$ is a function of $\{\xi(n)\}$ and hence itself recurrent. Hence $\eta \in R_0$ and $\{\xi(n)\}$ is R_0-measurable.

Finally we introduce the flows that correspond to R_0-measurable sequences.

Definition III.5. *Let X be a subflow of (Ω,T), where $\Omega = \Lambda \times \Lambda \times \Lambda \times \cdots$, and Λ is finite. If each sequence of Ω_X is R_0-measurable, then X is said to be an R_0-*flow. *The class of R_0-flows will be denoted by $\mathscr{R}_0$.*

THEOREM III.5. *(a) All semi-simple flows of finite type are in $\mathscr{R}_0$; (b) $\mathscr{R}_0 \times \mathscr{R}_0 \subset \mathscr{R}_0$; (c) Subflows of flows in $\mathscr{R}_0$ are in $\mathscr{R}_0$; (d) Factors of flows in $\mathscr{R}_0$ are in $\mathscr{R}_0$ provided they are of finite type; (e) Non-trivial Bernoulli flows do not occur in $\mathscr{R}_0$.*

Proof: (a), (b) and (c) are immediate from Theorem III.4. To prove (d) assume that $X \in \mathscr{R}_0$ and that $X \xrightarrow{\varphi} Y$, where Y is a flow of finite type. Writing the points of Ω_X and Ω_Y as sequences, $\alpha = (\alpha_1, \alpha_2, \alpha_3, \cdots) \in \Omega_X$, $\beta = (\beta_1, \beta_2, \beta_3, \cdots) \in \Omega_Y$, we have

$$\beta_n = \Phi(\alpha_n, \alpha_{n+1}, \cdots)$$

for some continuous function Φ on Ω_X. Since both the α_n and the β_m are finite-valued, Φ can depend only on finitely many coordinates:

$$\beta_n = \Phi(\alpha_n, \alpha_{n+1}, \cdots, \alpha_{n+r}) .$$

But then if α is R_0-measurable, so is β, and so $Y \in \mathscr{R}_0$.

Finally (e) follows from the fact that R_0 is a proper subset of Ω_2. Now the space of a Bernoulli flow consists of all Λ-sequences for some space Λ, and the only ring with respect to which all these sequences are measurable is the ring of all subsets of the natural numbers. This proves the theorem.

7. A Necessary Condition for an R_0-flow. Let A be an invariant subset of $\Omega = \Lambda \times \Lambda \times \Lambda \times \cdots$, where Λ is a finite set. Let Q be an infinite subset of the natural numbers. We say A is *free on Q* if for any choice of numbers $q_1, \cdots, q_r \in Q$, and $\lambda_1, \cdots, \lambda_r \in \Lambda$, we can find a sequence $\alpha \in A$ with $\alpha_{q_1} = \lambda_1, \cdots, \alpha_{q_r} = \lambda_r$.

THEOREM III.6. *If $(A,T) \in \mathscr{R}_0$ then A cannot be free on a relatively dense set (see Definition II.3).*

Proof: If A is free on Q and if $A' = \{\alpha' : \alpha'_n = \varphi(\alpha_n), \alpha \in A\}$ for some function $\varphi : \Lambda \to \Lambda'$, then A' is again free on Q. Choose φ as a two-valued function so that $A' \subset \Omega_2$. If $(A,T) \in \mathscr{R}_0$ then $(A',T) \in \mathscr{R}_0$ and so $A' \subset R_0$. We wish to show that if A' is free on Q, then Q cannot be relatively dense.

If $\beta \in \Omega_2$, then the set of n for which $\beta(n) = 1$ will be called the *support* of β. With A' as before, let B denote the set of all $\beta \in \Omega_2$ such that A' is free on the support of β. B is T-invariant, and because the notion of being free depended on conditions regarding finitely many coordinates at one time, B is seen to be closed. Suppose some $\beta \in B$ has for its support a relatively dense set. Clearly all translates of β have the same property (uniformly) and so do limits of these. The set $\overline{\{T^n\beta\}}$ then contains a minimal set M whose members have relatively dense support. In particular M does not degenerate to $\{0\}$. Let $\mu \in M$. Since A' is free on the support of μ, we can write $A'\mu = \Omega_2\mu$, or $\Omega_2 = A' + (1-\mu)\Omega_2$. If we set $M' = \{1 - \mu : \mu \in M\}$, then *a fortiori*, $\Omega_2 = A'M'\Omega_2$. Since A' is a restricted set, $M'\Omega_2 = \Omega_2$. Hence $1 \in M'$ and so $0 \in M$. But M is minimal and we assumed $M \neq \{0\}$. Hence no $\beta \in B$ has relatively dense support, and this proves the theorem.

Combining Theorems III.5 and III.6 we find that, beginning with semi-simple flows of finite type and forming subflows, factor flows of finite type and product flows we always obtain flows satisfying the conditions of the theorem. Namely, the space of the flow is a sequence space which cannot be free on a relatively dense set.

We note that deterministic flows also satisfy the conditions of Theorem III.6. For if a set A is free on a relatively dense set, it is clear that $\mu(A,n)$ increases exponentially with n.

We conclude this part by formulating two problems which we have not been able to solve:

Problem H. Is $R \subset \Omega_2$ a ring, and if so, does it coincide with R_0?

Problem I. Is the condition in Theorem III.6 sufficient as well as necessary for an R_0-flow?

Part IV. A Problem in Diophantine Approximation.

1. Minimal Sets on Tori. We let K^r denote the r-dimensional torus considered as an additive group: $K^r = \mathbf{R}^r/\mathbf{Z}^r$. The endomorphisms of K^r correspond to $r \times r$ integer matrices, and these form a semigroup $E(r)$. Each $\sigma \in E(r)$ determines a flow on K^r, and under certain mild conditions, (K^r,σ) will be an $\mathcal{F}$-flow. Theorem III.1 then gives us certain information about minimal sets of (K^r,σ): If M is minimal and B is an invariant set, then $M + B = K^r$ implies $B = K^r$.

In all this we have been dealing exclusively with a "one-parameter" semigroup of transformations. It is not difficult to see that the various notions we have used generalize to the situation where a more complicated semigroup acts on the space. Namely, we can allow flows (Ω,Σ) where Σ is now not a single transformation, but an abelian semigroup of transformations. The notions of minimal flows, semi-simple flows, $\mathcal{F}$-flows, and the notion of disjointness all generalize to this situation. Developing the analogy we may obtain a generalization of Theorem III.1 to abelian groups on which an abelian semigroup of endomorphisms acts. Specifically, we may obtain the following result whose usefulness will develop presently.

PROPOSITION IV.1. *Let Σ denote a commutative semigroup of endomorphisms of the r-torus K^r. We assume (i) that the adjoint semigroup of endomorphisms Σ^* of the dual group $\mathbf{Z}^r$ possesses no finite invariant subset in $\mathbf{Z}^r - \{0\}$, and (ii) that there exists a prime q with the property that all* det σ, $\sigma \in \Sigma$, *are relatively prime to q. Then if M and B are two closed Σ-invariant subsets of K^r and* M *is minimal with respect to these properties, $M + B = K^r$ implies $M = K^r$.*

Remark: We do not know whether condition (ii) is essential. Condition (i) is indispensable because, for example, if Σ were finite, the conclusion of the proposition would be false.

In proving this proposition, we shall, in fact, not proceed in the manner indicated, that is, we shall not redevelop the machinery alluded to in the case of flows in the wider sense. Since the results of this section may have some independent interest we have chosen to present the proof in a disguised form, avoiding the language of flows.

Proof of Proposition IV.1. Let $\Sigma^{(n)}$ denote the subsemigroup of $\sigma \in \Sigma$ for which $\sigma \equiv 1$ modulo q^n. Each $\sigma \in \Sigma$ has some power lying in $\Sigma^{(n)}$. For with respect to the finite ring $\mathbf{Z}/q^n\mathbf{Z}$, det σ is a unit for $\sigma \in \Sigma$. Hence, modulo

q^n, each $\sigma \in \Sigma$ is invertible, and Σ taken modulo q^n forms a group of matrices.

M is a minimal set with respect to Σ_j but it may not be minimal with respect to $\Sigma^{(n)}$. Let $M^{(n)}$ denote a subset of M which is minimal for $\Sigma^{(n)}$. Clearly we may choose the sequence $M^{(n)}$ with $M^{(n+1)} \subset M^{(n)}$. Then $M^\infty = \bigcap_1^\infty M^{(n)}$ is non-empty.

For a fixed n, let $\sigma_1, \cdots, \sigma_N$ denote a complete set of representatives in Σ of the group Σ modulo q^n. We may suppose that $\sigma_1 \in \Sigma^{(n)}$. We claim that $\bigcup_{i=1}^N \sigma_i M^{(n)} = M$. For this it suffices to show that the left-hand side is Σ-invariant, since it is a closed set contained in M. Let $\sigma \in \Sigma$; we must show that for each i, $\sigma\sigma_i M^{(n)} \subset \sigma_j M^{(n)}$ for some j. Choose j so that $\sigma_j = \sigma\sigma_i$ modulo q^n. Since both $\sigma\sigma_i M^{(n)}$ and $\sigma_j M^{(n)}$ are minimal Σ-invariant sets (at this point we have put the commutativity of Σ to use), it suffices to show that they intersect. For this it suffices to show that $\sigma\sigma_i\Sigma^{(n)}$ and $\sigma_j\Sigma^{(n)}$ intersect. Choose l such that $\sigma_j^l \in \Sigma^{(n)}$. Then $\sigma\sigma_i\sigma_j^l \in \sigma\sigma_i\Sigma^{(n)}$. Also $\sigma\sigma_i\sigma_j^{l-1} \equiv \sigma_j\sigma_j^{l-1} = \sigma_j^l \equiv 1 \pmod{q^n}$. Hence $\sigma\sigma_i\sigma_j^l \in \sigma_j\Sigma^{(n)}$ and $\sigma\sigma_i\Sigma^{(n)} \cap \sigma_j\Sigma^{(n)} \neq \emptyset$. This proves that $\bigcup_{i=1}^N \sigma_i M^{(n)} = M$.

Consider next $M^{(n)} + B$, which we denote $K^{(n)}$. From the foregoing we find that $\bigcup_{i=1}^N \sigma_i K^{(n)} = K^r$. Thus, some $\sigma_i K^{(n)}$ has a non-empty interior. Then $K^{(n)}$ itself must have a non-empty interior. Now $K^{(n)}$ is invariant under $\Sigma^{(n)}$, and in particular, it is invariant under each $\sigma \in \Sigma^{(n)}$. Now by [8], σ acts ergodically on (K^r,m), m being Lebesgue measure on K^r, if the eigenvalues of σ are distinct from roots of unity. However, hypothesis (i) ensures the existence in Σ of an endomorphism with no roots of unity among its eigenvalues, and an appropriate power of this endomorphism lies in $\Sigma^{(n)}$ and still has no roots of unity among its eigenvalues. Since $K^{(n)}$ has positive Lebesgue measure, it must have measure 1; hence $K^{(n)} = K^r$. We have thus shown that $M^{(n)} + B = K^r$.

Fix $\xi \in M^\infty$; that is, ξ_0 belongs to each $M^{(n)}$. Consider a point in K^r of the form

$$\theta = (a_1 q^{-n}, a_2 q^{-n}, \cdots, a_r q^{-n}) , \tag{1}$$

where $a_1, a_2, \cdots, a_r$ are integers. We may write $\theta = \mu + \beta$ with $\mu \in M^{(n)}$, $\beta \in B$. The closure of $\Sigma^{(n)}\mu$ is a $\Sigma^{(n)}$-invariant subset of $M^{(n)}$, and since the latter is minimal for $\Sigma^{(n)}$, it follows that $\Sigma^{(n)}\mu$ is dense in $M^{(n)}$. In particular, $\xi = \lim \sigma_j \mu$ for a sequence $\{\sigma_j\}$ in $\Sigma^{(n)}$. Since $\sigma_j \in \Sigma^{(n)}$, it is congruent to the identity modulo q^n, and so $\sigma_j\theta = \theta$ for each θ. Applying σ_j to both sides of $\theta = \mu + \beta$, we find that $\theta = \xi + \beta'$ for some $\beta' \in B$. Thus $\xi + B$ includes all points of the form (1). But the latter are dense in K. Hence $\xi + B = K^r$, whence $B = K^r$. This proves the proposition.

2. Semigroups of Endomorphisms of *K*. If the semigroup Σ of Proposition IV.1 is sufficiently large, one can show that a minimal set cannot be "small" in the sense required by that proposition unless it is finite. When $r = 1$, i.e., for the circle K, this is, in fact, usually the case. For $r = 1$, $E(r)$, the endomorphism semigroup of K, consists of multiplication by integers. Σ may therefore be identified with a subset of $\mathbf{Z}$. Σ always contains non-negative integers and we denote this subset by Σ^+.

Definition IV.1. *A multiplicative semigroup Σ of integers is* lacunary *if all the members of Σ^+ are powers of a single integer a. Otherwise, Σ is* non-lacunary.

For example, $\{a^p, a^{p+1}, a^{p+2}, \cdots\}$ is a lacunary semigroup. $\{2^n 3^m; n,m = 1,2,3, \cdots\}$ is a non-lacunary semigroup.

LEMMA IV.1. *Let Σ be a non-lacunary semigroup, and suppose $\Sigma^+ = \{s_1, s_2, s_3, \cdots\}$ with $s_i < s_{i+1}$. Then, as $n \to \infty$,*

$$\frac{s_{n+1}}{s_n} \to 1 \ . \tag{2}$$

Proof: This is a consequence of the following: an additive semigroup of positive real numbers is either contained in a discrete subgroup, or becomes more and more dense as the numbers tend to ∞. Let S be such a semigroup; $S - S$ is a group, and it is not hard to see that if $S - S$ is a discrete subgroup, then S itself is contained in a discrete subgroup. ($S - S \subset \mathbf{Z}u$ implies $S \subset \mathbf{Z}u + s'$, and $2(nu + s') = mu + s'$, whence $s' \in \mathbf{Z}u$; hence $S \subset \mathbf{Z}u$.) Now it is known that a subgroup of $\mathbf{R}$ is either discrete or dense. Hence we may suppose that $S - S$ is dense in $\mathbf{R}$. We may still suppose this to be the case for a subsemigroup (possibly S itself) of S which is countably generated. Assume therefore that S itself is countably generated, say, by $a_1, a_2, a_3, \cdots$. It may be seen that

$$S - S = \bigcup_{n=1}^{\infty} (S - n(a_1 + a_2 + \cdots + a_n)) \ , \tag{3}$$

where the summands to the right of (3) form an increasing sequence. Let $\epsilon > 0$. Suppose the sets $S - n(a_1 + a_2 + \cdots + a_n)$ omitted an interval of length greater than ϵ somewhere in $(0,\infty)$. Then $S - n(a_1 + a_2 + \cdots + a_n) - k_n a_1$ would omit an interval of length ϵ inside of $(-a_1, 0)$ if k_n is appropriately chosen. Now

$$S - n(a_1 + a_2 + \cdots + a_n) - k_n a_1 \supset S - n(a_1 + a_2 + \cdots + a_n)$$

and since the sets $S - n(a_1 + a_2 + \cdots + a_n)$ become successively more dense, this is not possible. Hence for n sufficiently large, $S - n(a_1 + a_2 + \cdots + a_n)$ is ϵ-dense in $(0,\infty)$, whence S is ϵ-dense in $(n(a_1 + \cdots + a_n), \infty)$. This proves the assertion regarding additive semigroups. Returning to Σ, we now find two alternatives for $\log \Sigma$. Suppose $\log \Sigma$ is contained in a discrete subgroup of $\mathbf{R}$; then for each pair i, j some power of s_i must coincide with some power of s_j. But this implies that Σ is lacunary. This leads to the conclusion of the lemma.

LEMMA IV.2. *Let Σ be a non-lacunary semigroup and let A be a closed Σ-invariant subset of K with the property that 0 is a non-isolated point of A. Then $A = K$.*

Proof: Consider a sequence of points $\alpha_n \in A$ with $\alpha_n \to 0$. Fix $\epsilon > 0$ and choose n_1 large enough so that $s_{n+1}/s_n < 1+\epsilon$ for $n > n_1$. Next choose n_2 so large that $s_{n_1}\delta_{n_2} < \epsilon$. For $n > n_1$

$$(s_{n+1} - s_n)\delta_{n_2} < \epsilon s_n \delta_{n_2} .$$

Let n range over the values for which $s_n\delta_{n_2} \in [\epsilon,1]$. The distance between successive values of $s_n\delta_{n_2}$ does not exceed ϵ, so that $\{s_n\delta_{n_2}\}$ is ϵ-dense in $[0,1]$. Since ϵ is arbitrary, $A = K$.

PROPOSITION IV.2. *If Σ is a non-lacunary semigroup of integers, the only minimal sets in K for Σ are finite sets (of rationals).*

Proof: Suppose that the proposition has been proved for semigroups with the property that some prime number q exists which is relatively prime to all numbers of the semigroup. It then follows in the general case. For if Σ is non-lacunary we can find $\Sigma' \subset \Sigma$ which is still non-lacunary and satisfies this condition. A minimal set for Σ contains a minimal set for Σ'. But if a minimal set for Σ contains a rational point, all its points must be rational with the same denominator. Thus we may suppose that Σ satisfies the condition in question and as a result Proposition IV.1 applies to minimal sets for Σ. Let M be a minimal set for Σ. $M - M$ is a Σ-invariant set. If M is infinite then $M - M$ contains 0 as a non-isolated point. But then, by the foregoing lemma, $M - M = K$. Now this contradicts Proposition IV.1. We conclude that M is finite. Since Σ is infinite, a Σ-invariant set can be finite only if it consists of rationals. This proves the proposition.

THEOREM IV.1. *If Σ is a non-lacunary semigroup of integers and α is an irrational, then $\Sigma\alpha$ is dense in K.*

Proof: Let A be the closure of $\Sigma\alpha$ in K. A is a Σ-invariant set and necessarily contains a minimal Σ-invariant set. Hence A contains a rational p/q. $\Sigma\alpha$ itself consists only of irrationals; hence p/q is a non-isolated point of A. Hence 0 is a non-isolated point of qA. Now qA is itself a Σ-invariant closed set, so by Lemma IV.2, $qA = K$. This implies that

$$\Delta \cup \left(\Delta + \frac{1}{q}\right) \cup \cdots \cup \left(\Delta + \frac{q-1}{q}\right) = K .$$

Since these sets are closed, Δ must have a non-empty interior. But then $s\Delta \subset \Delta$ for $s > 1$ implies $\Delta = K$. This proves the theorem.

The conclusion of the theorem is clearly false for lacunary semigroups. For, let a be a positive integer and set $\alpha = \sum_1^\infty a^{-k^2}$. Then $\{a^n\alpha\}$ modulo 1 has only the limit points $0, a^{-1}, a^{-2}, \cdots$.

If r is a positive integer, then the set of rth powers $\{n^r\}$ forms a non-lacunary semigroup. As a special case of Theorem IV.1, we have the result of Hardy and Littlewood: If α is irrational, $\{n^r\alpha\}$ is dense modulo 1. (This is most familiar as a special case of Weyl's theorem on equidistribution.) The semigroup $\{2^n3^m\}$ corresponds to a subset which is thinner than any of the $\{n^r\}$ and for which the same conclusion still holds.

We point out that the equidistribution conclusion cannot be made in the same generality. If the sequence $\{2^n 3^m\}$ is arranged in increasing order as $\{s_n\}$, then $\{s_n\alpha\}$ is not necessarily equidistributed modulo one when α is irrational. We mention without proof the fact that if $\alpha = \Sigma 6^{-n_k}$ and n_k is of sufficiently rapid growth, then $\{s_n\alpha\}$ is not equidistributed modulo one.

BIBLIOGRAPHY

[1] ADLER, R. L., KONHEIM, A. G. and MCANDREW, M. H., Topological entropy, *Trans. Amer. Math. Soc.* **114** (1965), 309–319.

[2] ELLIS, R., Distal transformation groups, *Pacific J. Math.* **8** (1958), 401–405.

[3] FURSTENBERG, H., *Stationary Processes and Prediction Theory*, Princeton Univ. Press, Princeton, 1960.

[4] FURSTENBERG, H., Strict ergodicity and transformations of the torus, *Amer. J. Math.* **83** (1961), 573–601.

[5] FURSTENBERG, H., The structure of distal flows, *Amer. J. Math.* **85** (1963), 477–515.

[6] GOTTSCHALK, W. H., and HEDLUND, G. A., *Topological Dynamics*, Amer. Math. Soc. Colloq. Publ., Vol. 83, 1961.

[7] GOTTSCHALK, W. H., An irreversible minimal set, in *Ergodic Theory*, Academic Press, New York, 1963.

[8] HALMOS, P. R., On automorphisms of compact groups, *Bull. Amer. Math. Soc.* **49** (1943), 619–624.

[9] HALMOS, P. R., *Lectures on Ergodic Theory*, Publ. Math. Soc. Japan, Tokyo, 1956.

[10] HALMOS, P. R., *Entropy in Ergodic Theory*, Univ. Chicago Press, Chicago, 1959.

[11] NIVEN, I., and ZUCKERMAN, H. S., On the definition of normal number, *Pacific J. Math.* **1** (1951), 103–109.

[12] ROBBINS, H. E., On a class of recurrent sequences, *Bull. Amer. Math. Soc.* **43** (1937), 413–417.

[13] ROKHLIN, V. A., New progress in the theory of transformations with invariant measure, *Russian Math. Surveys* **15** (1960), 1–22.

[14] WU, T. S., On homomorphisms in topological dynamics, (to appear).

(Received 15 July 1966)