

# MEASURE-PRESERVING SYSTEMS

KARL PETERSEN

## OUTLINE

### Glossary

1. Glossary	2
2. Definition of the subject and its importance	3
3. Introduction: The dynamical viewpoint	3
3.1. Two examples	3
3.2. A range of actions	4
4. Where do measure-preserving systems come from?	5
4.1. Systems in equilibrium	5
4.2. Stationary stochastic processes	6
5. Construction of measures	7
5.1. The Carathéodory construction	7
5.2. Measures on shift spaces	10
5.3. The Kolmogorov Consistency Theorem	11
6. Invariant measures on topological dynamical systems	12
6.1. Existence of invariant measures	12
6.2. Ergodicity and unique ergodicity	13
7. Finding finite invariant measures equivalent to a quasi-invariant measure	14
8. Finding $\sigma$ -finite invariant measures equivalent to a quasi-invariant measure	16
8.1. First necessary and sufficient conditions	16
8.2. Conservativity and recurrence	17

---

Received by the editors May 3, 2007.

1991 *Mathematics Subject Classification.* 37A30, 37A05.

8.3.	Using first-return maps, and counterexamples to existence	19
8.4.	Invariant measures for maps of the interval or circle	20
9.	Some mathematical background	20
9.1.	Lebesgue spaces	20
9.2.	Conditional expectation	25
9.3.	The Spectral Theorem	26
10.	Future directions	28
11.	Bibliography	29
	Primary Literature	29
	Books and Reviews	30

## 1. GLOSSARY

*Dynamical system.* A set acted upon by an algebraic object. Elements of the set represent all possible states or configurations, and the action represents all possible changes.

*Ergodic.* A measure-preserving system is ergodic if it is essentially indecomposable, in the sense that given any invariant measurable set, either the set or its complement has measure 0.

*Lebesgue space.* A measure space that is isomorphic with the usual Lebesgue measure space of a subinterval of the set of real numbers, possibly together with countably or finitely many point masses.

*Measure.* An assignment of sizes to sets. A measure that takes values only between 0 and 1 assigns probabilities to events.

*Stochastic process.* A family of random variables (measurable functions). Such an object represents a family of measurements whose outcomes may be subject to chance.

*Subshift, shift space.* A closed shift-invariant subset of the set of infinite sequences with entries from a finite alphabet.

## 2. DEFINITION OF THE SUBJECT AND ITS IMPORTANCE

Measure-preserving systems model processes in equilibrium by transformations on probability spaces or, more generally, measure spaces. They are the basic objects of study in ergodic theory, a central part of dynamical systems theory. These systems arise from science and technology as well as from mathematics itself, so applications are found in a wide range of areas, such as statistical physics, information theory, celestial mechanics, number theory, population dynamics, economics, and biology.

## 3. INTRODUCTION: THE DYNAMICAL VIEWPOINT

Sometimes introducing a dynamical viewpoint into an apparently static situation can help to make progress on apparently difficult problems. For example, equations can be solved and functions optimized by reformulating a given situation as a fixed point problem, which is then addressed by iterating an appropriate mapping. Besides practical applications, this strategy also appears in theoretical settings, for example modern proofs of the Implicit Function Theorem. Moreover, the introduction of the ideas of change and motion leads to new concepts, new methods, and even new kinds of questions. One looks at actions and orbits and instead of always seeking exact solutions begins perhaps to ask questions of a qualitative or probabilistic nature: what is the general behavior of the system, what happens for most initial conditions, what properties of systems are typical within a given class of systems, and so on. Much of the credit for introducing this viewpoint should go to Henri Poincaré [29].

**3.1. Two examples.** Consider two particular examples, one simple and the other not so simple. Decimal or base 2 expansions of numbers in the unit interval raise many natural questions about frequencies of digits and blocks. Instead of regarding the base 2 expansion  $x = .x_0x_1\dots$  of a fixed  $x \in [0, 1]$  as being given, we can regard it as arising from a dynamical process. Define  $T : [0, 1] \rightarrow [0, 1]$  by  $Tx = 2x \bmod 1$  (the fractional part of  $2x$ ) and let  $\mathcal{P} = \{P_0 = [0, 1/2), P_1 = [1/2, 1]\}$  be a partition of  $[0, 1]$  into two subintervals. We code the orbit of any point  $x \in [0, 1]$  by 0's and 1's by letting  $x_k = i$  if  $T^k x \in P_i, k = 0, 1, 2, \dots$ . Then reading the expansion of  $x$  amounts to applying to the coding the shift transformation and projection onto the first coordinate. This is equivalent to following the orbit of  $x$  under  $T$  and noting which element of the partition  $\mathcal{P}$  is entered at each time. Reappearances of blocks amount to recurrence to cylinder sets as  $x$  is moved by  $T$ , frequencies of blocks correspond to ergodic averages, and Borel's theorem on normal numbers is seen as a special case of the Ergodic Theorem.

Another example concerns Szemerédi's Theorem [34], which states that every subset  $A \subset \mathbb{N}$  of the natural numbers which has positive upper density contains arbitrarily long arithmetic progressions: given  $L \in \mathbb{N}$  there are  $s, m \in \mathbb{N}$  such that  $s, s + m, \dots, s + (L - 1)m \in A$ . Szemerédi's proof was ingenious, direct, and long. Furstenberg [15] saw how to obtain this result as a corollary of a strengthening of

Poincaré's Recurrence Theorem in ergodic theory, which he then proved. Again we have an apparently static situation: a set  $A \subset \mathbb{N}$  of positive density in which we seek arbitrarily long regularly spaced subsets. Furstenberg proposed to consider the characteristic function  $\mathbf{1}_A$  of  $A$  as a point in the space  $\{0, 1\}^{\mathbb{N}}$  of 0's and 1's and to form the orbit closure  $X$  of this point under the shift transformation  $\sigma$ . Because  $A$  has positive density, it is possible to find a shift-invariant measure  $\mu$  on  $X$  which gives positive measure to the cylinder set  $B = [1] = \{x \in X : x_1 = 1\}$ . Furstenberg's Multiple Recurrence Theorem says that given  $L \in \mathbb{N}$  there is  $m \in \mathbb{N}$  such that  $\mu(B \cap T^{-m}B \cap \dots \cap T^{-(L-1)m}B) > 0$ . If  $y$  is a point in this intersection, then  $y$  contains a block of  $L$  1's, each at distance  $m$  from the next. And since  $y$  is in the orbit closure of  $\mathbf{1}_A$ , this block also appears in the sequence  $\mathbf{1}_A \in \{0, 1\}^{\mathbb{N}}$ , yielding the result.

Aspects of the dynamical argument remain in new combinatorial and harmonic-analytic proofs of the Szemerédi Theorem by T. Gowers [16, 17] and T. Tao [35], as well as the extension to the (density zero) set of prime numbers by B. Green and T. Tao [18, 36].

**3.2. A range of actions.** Here is a sample of dynamical systems of various kinds:

- (1) A semigroup or group  $G$  acts on a set  $X$ . There is given a map  $G \times X \rightarrow X$ ,  $(g, x) \rightarrow gx$ , and it is assumed that

$$(3.1) \quad g_1(g_2x) = (g_1g_2)x \quad \text{for all } g_1, g_2 \in G, x \in X$$

$$(3.2) \quad ex = x \quad \text{for all } x \in X, \text{ if } G \text{ has an identity element } e.$$

- (2) A continuous linear operator  $T$  acts on a Banach or Hilbert space  $V$ .
- (3)  $\mathcal{B}$  is a Boolean  $\sigma$ -algebra (a set together with a zero element  $0$  and operations  $\vee, \wedge, '$  which satisfy the same rules as  $\emptyset, \cup, \cap, ^c$  (complementation) do for  $\sigma$ -algebras of sets);  $\mathcal{N}$  is a  $\sigma$ -ideal in  $\mathcal{B}$  ( $N \in \mathcal{N}, B \in \mathcal{B}, B \wedge N = B$  implies  $B \in \mathcal{N}$ ; and  $N_1, N_2, \dots \in \mathcal{N}$  implies  $\bigvee_{n=1}^{\infty} N_n \in \mathcal{N}$ ); and  $S : \mathcal{B} \rightarrow \mathcal{B}$  preserves the Boolean  $\sigma$ -algebra operations and  $S\mathcal{N} \subset \mathcal{N}$ .
- (4)  $\mathcal{B}$  is a Boolean  $\sigma$ -algebra,  $\mu$  is a countably additive positive (nonzero except on the zero element of  $\mathcal{B}$ ) function on  $\mathcal{B}$ , and  $S : \mathcal{B} \rightarrow \mathcal{B}$  is as above. Then  $(\mathcal{B}, \mu)$  is a *measure algebra* and  $S$  is a *measure algebra endomorphism*.
- (5)  $(X, \mathcal{B}, \mu)$  is a measure space ( $X$  is a set,  $\mathcal{B}$  is a  $\sigma$ -algebra of subsets of  $X$ , and  $\mu : \mathcal{B} \rightarrow [0, \infty]$  is countably additive: If  $B_1, B_2, \dots \in \mathcal{B}$  are pairwise disjoint, then  $\mu(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu(B_n)$ );  $T : X \rightarrow X$  is measurable ( $T^{-1}\mathcal{B} \subset \mathcal{B}$ ) and nonsingular ( $\mu(B) = 0$  implies  $\mu(T^{-1}B) = 0$ —or, more stringently,  $\mu$  and  $\mu T^{-1}$  are equivalent in the sense of absolute continuity).
- (6)  $(X, \mathcal{B}, \mu)$  is a measure space,  $T : X \rightarrow X$  is a one-to-one onto map such that  $T$  and  $T^{-1}$  are both measurable (so that  $T^{-1}\mathcal{B} = \mathcal{B} = T\mathcal{B}$ ), and  $\mu(T^{-1}B) = \mu(B)$  for all  $B \in \mathcal{B}$ . (In practice often  $T$  is not one-to-one, or onto, or even well-defined on all of  $X$ , but only after a set of measure

zero is deleted.) This is the case of most interest for us, and then we call  $(X, \mathcal{B}, \mu, T)$  a *measure-preserving system*. We also allow for the possibility that  $T$  is not invertible, or that some other group (such as  $\mathbb{R}$  or  $\mathbb{Z}^d$ ) or semigroup acts on  $X$ , but the case of  $\mathbb{Z}$  actions will be the main focus of this article.

- (7)  $X$  is a compact metric space and  $T : X \rightarrow X$  is a homeomorphism. Then  $(X, T)$  is a *topological dynamical system*.
- (8)  $M$  is a compact manifold ( $\mathcal{C}^k$  for some  $k \in [1, \infty]$ ) and  $T : M \rightarrow M$  is a diffeomorphism (one-to-one and onto, with  $T$  and  $T^{-1}$  both  $\mathcal{C}^k$ ). Then  $(M, T)$  is a *smooth dynamical system*. Such examples can arise from solutions of an autonomous differential equation given by a vector field on  $M$ . Recall that in  $\mathbb{R}^n$ , an ordinary differential equation initial-value problem  $x' = f(x), x(0) = x_0$  has a unique solution  $x(t)$  as long as  $f$  satisfies appropriate smoothness conditions. The existence and uniqueness theorem for differential equations then produces a flow according to  $T_t x_0 = x(t)$ , satisfying  $T_{s+t} x_0 = T_s(T_t x_0)$ . Restricting to a compact invariant set (if there is one) and taking  $T = T_1$  (the time 1 map) gives us a smooth system  $(M, f)$ .

Naturally there are relations and inclusions among these examples of actions. Often problems can be clarified by forgetting about some of the structure that is present or by adding desirable structure (such as topology) if it is not. There remain open problems about representation and realization; for example, taking into account necessary restrictions, which measure-preserving systems can be realized as smooth systems preserving a smooth measure? Sometimes interesting aspects of the dynamics of a smooth system can be due to the presence of a highly non-smooth subsystem, for example a compact lower-dimensional invariant set. Thus one should be ready to deal with many kinds of dynamical systems.

#### 4. WHERE DO MEASURE-PRESERVING SYSTEMS COME FROM?

**4.1. Systems in equilibrium.** Besides physical systems, abstract dynamical systems can also represent aspects of biological, economic, or other real world systems. Equilibrium does not mean stasis, but rather that the changes in the system are governed by laws which are not themselves changing. The presence of an invariant measure means that the probabilities of observable events do not change with time. (But of course what happens at time 2 can still depend on what happens at time 1, or, for that matter, at time 3.)

We consider first the example of the wide and important class of Hamiltonian systems. Many systems that model physical situations, for example a large number of ideal charged particles in a container, can be studied by means of Hamilton's equations. The state of the entire system at any time is supposed to be specified by a vector  $(q, p) \in \mathbb{R}^{2n}$ , the *phase space*, with  $q$  listing the coordinates of the positions of all of the particles, and  $p$  listing the coordinates of their momenta. We assume that there is a time-independent *Hamiltonian function*  $H(q, p)$  such that the time

development of the system satisfies *Hamilton's equations*:

$$(4.1) \quad \frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \quad i = 1, \dots, n.$$

Often the Hamiltonian function is the sum of kinetic and potential energy:

$$(4.2) \quad H(q, p) = K(p) + U(q).$$

The potential energy  $U(q)$  may depend on interactions among the particles or with an external field, while the kinetic energy  $K(p)$  depends on the velocities and masses of the particles.

As discussed above, solving these equations with initial state  $(q, p)$  for the system produces a flow  $(q, p) \rightarrow T_t(q, p)$  in phase space. According to *Liouville's Theorem*, this flow preserves Lebesgue measure on  $\mathbb{R}^{2n}$ . Calculating  $dH/dt$  by means of the Chain Rule and using Hamilton's equations shows that  $H$  is constant on orbits of the flow, and thus each set of constant energy  $X(H_0) = \{(q, p) : H(q, p) = H_0\}$  is an invariant set. Thus one should consider the flow restricted to the appropriate invariant set. It turns out that there are also natural invariant measures on the sets  $X(H_0)$ , namely the ones given by rescaling the volume element  $dS$  on  $X(H_0)$  by the factor  $1/|\nabla H|$ . For details, see [25].

Systems in equilibrium can also be hiding inside systems not in equilibrium, for example if there is an attractor supporting an SRB measure (for Sinai, Ruelle, and Bowen)(for definitions of the terms used here and more explanations, see the article in this collection by A. Wilkinson). Suppose that  $T : M \rightarrow M$  is a diffeomorphism on a compact manifold as above, and that  $m$  is a version of Lebesgue measure on  $M$ , say given by a smooth volume form. We consider  $m$  to be a “physical measure”, corresponding to laboratory measurements of observable quantities, whose values can be determined to lie in certain intervals in  $\mathbb{R}$ . Quite possibly  $m$  is not itself invariant under  $T$ , and an experimenter might observe strange or chaotic behavior whenever the state of the system gets close to some compact invariant set  $X$ . The dynamics of  $T$  restricted to  $X$  can in fact be quite complicated—maybe a full shift, which represents completely undeterministic behavior (for example if there is a horseshoe present), or a shift of finite type, or some other complicated topological dynamical system. Possibly  $m(X) = 0$ , so that  $X$  is effectively invisible to the observer except through its effects. It can happen that there is a  $T$ -invariant measure  $\mu$  supported on  $X$  such that

$$(4.3) \quad \frac{1}{n} \sum_{k=0}^{n-1} mT^{-k} \rightarrow \mu \quad \text{weak } ^*,$$

and then the long-term equilibrium dynamics of the system is described by  $(X, T, \mu)$ . For a recent survey on SRB measures, see [39].

**4.2. Stationary stochastic processes.** A *stationary process* is a family  $\{f_t : t \in T\}$  of random variables (measurable functions) on a probability space  $(\Omega, \mathcal{F}, P)$ . Usually  $T$  is  $\mathbb{Z}, \mathbb{N}$ , or  $\mathbb{R}$ . For the remainder of this section let us fix  $T = \mathbb{Z}$  (although the following definition could make sense for  $T$  any semigroup).

We say that the process  $\{f_n : n \in \mathbb{Z}\}$  is *stationary* if its finite-dimensional distributions are translation invariant, in the sense that for each  $r = 1, 2, \dots$ , each  $n_1, \dots, n_r \in \mathbb{Z}$ , each choice of Borel sets  $B_1, \dots, B_r \subset \mathbb{R}$ , and each  $s \in \mathbb{Z}$ , we have

$$(4.4) \quad P\{\omega : f_{n_1}(\omega) \in B_1, \dots, f_{n_r}(\omega) \in B_r\} = P\{\omega : f_{n_1+s}(\omega) \in B_1, \dots, f_{n_r+s}(\omega) \in B_r\}.$$

The  $f_n$  represent measurements made at times  $n$  of some random phenomenon, and the probability that a particular finite set of measurements yield values in certain ranges is supposed to be independent of time.

Each stationary process  $\{f_n : n \in \mathbb{Z}\}$  on  $(\Omega, \mathcal{F}, P)$  corresponds to a shift-invariant probability measure  $\mu$  on the set  $\mathbb{R}^{\mathbb{Z}}$  (with its Borel  $\sigma$ -algebra) and a single observable, namely the projection  $\pi_0$  onto the 0'th coordinate, as follows. Define

$$(4.5) \quad \phi : \Omega \rightarrow \mathbb{R}^{\mathbb{Z}} \quad \text{by } \phi(\omega) = (f_n(\omega))_{-\infty}^{\infty},$$

and for each Borel set  $E \subset \mathbb{R}^{\mathbb{Z}}$ , define  $\mu(E) = P(\phi^{-1}E)$ . Then examining the values of  $\mu$  on cylinder sets— for Borel  $B_1, \dots, B_r \subset \mathbb{R}$ ,

$$(4.6) \quad \mu\{x \in \mathbb{R}^{\mathbb{Z}} : x_{n_i} \in B_i, i = 1, \dots, r\} = P\{\omega \in \Omega : f_{n_i}(\omega) \in B_i, i = 1, \dots, r\}$$

—and using stationarity of  $(f_n)$  shows that  $\mu$  is invariant under  $\sigma$ . Moreover, the processes  $(f_n)$  on  $\Omega$  and  $\pi_0 \circ \sigma^n$  on  $\mathbb{R}^{\mathbb{Z}}$  have the same finite-dimensional distributions, so they are equivalent for the purposes of probability theory.

## 5. CONSTRUCTION OF MEASURES

We review briefly (following [33]) the construction of measures largely due to C. Carathéodory [8], with input from M. Fréchet [13], H. Hahn [19], A. N. Kolmogorov [26], and others, then discuss the application to construction of measures on shift spaces and of stochastic processes in general.

**5.1. The Carathéodory construction.** A *semialgebra* is a family  $\mathcal{S}$  of subsets of a set  $X$  which is closed under finite intersections and such that the complement of any member of  $\mathcal{S}$  is a finite disjoint union of members of  $\mathcal{S}$ . Key examples are

- (1) the family  $\mathcal{H}$  of half-open subintervals  $[a, b)$  of  $[0, 1)$ ;
- (2) in the space  $X = A^{\mathbb{Z}}$  of doubly infinite sequences on a finite alphabet  $A$ , the family  $\mathcal{C}$  of *cylinder sets* (determined by fixing finitely many entries)

$$(5.1) \quad \{x \in A^{\mathbb{Z}} : x_{n_1} = a_1, \dots, x_{n_r} = a_r\};$$

- (3) the family  $\mathcal{C}_1$  of *anchored cylinder sets*

$$(5.2) \quad \{x \in A^{\mathbb{N}} : x_1 = a_1, \dots, x_r = a_r\}$$

in the space  $X = A^{\mathbb{N}}$  of one-sided infinite sequences on a finite alphabet  $A$ .

An *algebra* is a family of subsets of a set  $X$  which is closed under finite unions, finite intersections, and complements. A  $\sigma$ -*algebra* is a family of subsets of a set  $X$  which is closed under countable unions, countable intersections, and complements.

If  $\mathcal{S}$  is a semialgebra of subsets of  $X$ , the *algebra*  $\mathcal{A}(\mathcal{S})$  generated by  $\mathcal{S}$  is the smallest algebra of subsets of  $X$  which contains  $\mathcal{S}$ .  $\mathcal{A}(\mathcal{S})$  is the intersection of all the subalgebras of the set  $2^X$  of all subsets of  $X$  and consists exactly of all finite disjoint unions of elements of  $\mathcal{S}$ . Given an algebra  $\mathcal{A}$ , the  $\sigma$ -*algebra*  $\mathcal{B}(\mathcal{A})$  generated by  $\mathcal{A}$  is the smallest  $\sigma$ -algebra of subsets of  $X$  which contains  $\mathcal{A}$ .

A *nonnegative set function* on  $\mathcal{S}$  is a function  $\mu : \mathcal{S} \rightarrow [0, \infty]$  such that  $\mu(\emptyset) = 0$  if  $\emptyset \in \mathcal{S}$ . We say that such a  $\mu$  is

- *finitely additive* if whenever  $S_1, \dots, S_n \in \mathcal{S}$  are pairwise disjoint and  $S = \cup_{i=1}^n S_i \in \mathcal{S}$ , we have  $\mu(S) = \sum_{i=1}^n \mu(S_i)$ ;
- *countably additive* if whenever  $S_1, S_2, \dots \in \mathcal{S}$  are pairwise disjoint and  $S = \cup_{i=1}^{\infty} S_i \in \mathcal{S}$ , we have  $\mu(S) = \sum_{i=1}^{\infty} \mu(S_i)$ ; and
- *countably subadditive* if whenever  $S_1, S_2, \dots \in \mathcal{S}$  and  $S = \cup_{i=1}^{\infty} S_i \in \mathcal{S}$ , we have  $\mu(S) \leq \sum_{i=1}^{\infty} \mu(S_i)$ .

A *measure* is a countably additive nonnegative set function defined on a  $\sigma$ -algebra.

**Proposition 5.1.** *Let  $\mathcal{S}$  be a semialgebra and  $\mu$  a nonnegative set function on  $\mathcal{S}$ . In order that  $\mu$  have an extension to a finitely additive set function on the algebra  $\mathcal{A}(\mathcal{S})$  generated by  $\mathcal{S}$ , it is necessary and sufficient that  $\mu$  be finitely additive on  $\mathcal{S}$ .*

*Proof.* The stated condition is obviously necessary. Conversely, given  $\mu$  which is finitely additive on  $\mathcal{S}$ , it is natural to define

$$(5.3) \quad \mu\left(\bigcup_{i=1}^n S_i\right) = \sum_{i=1}^n \mu(S_i)$$

whenever  $A = \cup_{i=1}^n S_i$  (with the  $S_i$  pairwise disjoint) is in the algebra  $\mathcal{A}(\mathcal{S})$  generated by  $\mathcal{S}$ . It is necessary to verify that  $\mu$  is then well defined on  $\mathcal{A}(\mathcal{S})$ , since each element of  $\mathcal{A}(\mathcal{S})$  may have more than one representation as a finite disjoint union of members of  $\mathcal{S}$ . But, given two such representations of a single set  $A$ , forming the common refinement and applying finite additivity on  $\mathcal{S}$  shows that  $\mu$  so defined assigns the same value to  $A$  both times. Then finite additivity on  $\mathcal{A}(\mathcal{S})$  of the extended  $\mu$  is clear.  $\square$

**Proposition 5.2.** *Let  $\mathcal{S}$  be a semialgebra and  $\mu$  a nonnegative set function on  $\mathcal{S}$ . In order that  $\mu$  have an extension to a countably additive set function on the algebra  $\mathcal{A}(\mathcal{S})$  generated by  $\mathcal{S}$ , it is necessary and sufficient that  $\mu$  be (i) finitely additive and (ii) countably subadditive on  $\mathcal{S}$ .*

*Proof.* Conditions (i) and (ii) are clearly necessary. If  $\mu$  is finitely additive on  $\mathcal{S}$ , then by Proposition 5.1  $\mu$  has an extension to a finitely additive nonnegative set function, which we will still denote by  $\mu$ , on  $\mathcal{A}(\mathcal{S})$ .

Let us see that this extension  $\mu$  is countably subadditive on  $\mathcal{A}(\mathcal{S})$ . Suppose that  $A_1, A_2, \dots \in \mathcal{A}(\mathcal{S})$  are pairwise disjoint and their union  $A \in \mathcal{A}(\mathcal{S})$ . Then  $A$  is a

finite disjoint union of sets in  $\mathcal{S}$ , as is each  $A_i$ :

$$(5.4) \quad \begin{aligned} A &= \bigcup_{i=1}^{\infty} A_i, & \text{each } A_i &= \bigcup_{k=1}^{n_i} S_{ik}, \\ A &= \bigcup_{j=1}^m R_j, & \text{each } A_i &\in \mathcal{A}(\mathcal{S}), \text{ each } S_{ik}, R_j \in \mathcal{S}. \end{aligned}$$

Since each  $R_j \in \mathcal{S}$ , by countable subadditivity of  $\mu$  on  $\mathcal{S}$ , and using  $R_j = R_j \cap A$ ,

$$(5.5) \quad \mu(R_j) = \mu\left(\bigcup_{i=1}^{\infty} \bigcup_{k=1}^{n_i} S_{ik} \cap R_j\right) \leq \sum_{i=1}^{\infty} \sum_{k=1}^{n_i} \mu(S_{ik} \cap R_j),$$

and hence, by finite additivity of  $\mu$  on  $\mathcal{A}(\mathcal{S})$ ,

$$(5.6) \quad \begin{aligned} \mu(A) &= \sum_{j=1}^m \mu(R_j) \leq \sum_{i=1}^{\infty} \sum_{k=1}^{n_i} \sum_{j=1}^m \mu(S_{ik} \cap R_j) \\ &= \sum_{i=1}^{\infty} \sum_{k=1}^{n_i} \mu(S_{ik}) = \sum_{i=1}^{\infty} \mu(A_i). \end{aligned}$$

Now finite additivity of  $\mu$  on an algebra  $\mathcal{A}$  implies that  $\mu$  is *monotonic* on the algebra: if  $A, B \in \mathcal{A}$  and  $A \subset B$ , then  $\mu(A) \leq \mu(B)$ . Thus if  $A_1, A_2, \dots \in \mathcal{A}(\mathcal{S})$  are pairwise disjoint and their union  $A \in \mathcal{A}(\mathcal{S})$ , then for each  $n$  we have  $\sum_{i=1}^n \mu(A_i) = \mu(\cup_{i=1}^n A_i) \leq \mu(A)$ , and hence  $\sum_{i=1}^{\infty} \mu(A_i) \leq \mu(A)$ .  $\square$

**Theorem 5.3.** *In order that a nonnegative set function  $\mu$  on an algebra  $\mathcal{A}$  of subsets of a set  $X$  have an extension to a (countably additive) measure on the  $\sigma$ -algebra  $\mathcal{B}(\mathcal{A})$  generated by  $\mathcal{A}$ , it is necessary and sufficient that  $\mu$  be countably additive on  $\mathcal{A}$ .*

Here is a sketch of how the extension can be constructed. Given a countably additive nonnegative set function  $\mu$  on an algebra  $\mathcal{A}$  of subsets of a set  $X$ , one defines the *outer measure*  $\mu^*$  that it determines on the family  $2^X$  of all subsets of  $X$  by

$$(5.7) \quad \mu^*(E) = \inf\left\{\sum_{i=1}^{\infty} \mu(A_i) : A_i \in \mathcal{A}, E \subset \cup_{i=1}^{\infty} A_i\right\}.$$

Then  $\mu^*$  is a nonnegative, countably subadditive, monotonic set function on  $2^X$ .

Define a set  $E$  to be  $\mu^*$ -*measurable* if for all  $T \subset X$ ,

$$(5.8) \quad \mu^*(T) = \mu^*(T \cap E) + \mu^*(T \cap E^c).$$

This ingenious definition can be partly motivated by noting that if  $\mu^*$  is to be finitely additive on the family  $\mathcal{M}$  of  $\mu^*$ -measurable sets, which should contain  $X$ , then at least this condition must hold when  $T = X$ . It is amazing that then this definition readily, with just a little set theory and a few  $\epsilon$ 's, yields the following theorem.

**Theorem 5.4.** *Let  $\mu$  be a countably additive nonnegative set function on an algebra  $\mathcal{A}$  of subsets of a set  $X$ , and let  $\mu^*$  be the outer measure that it determines on the family  $2^X$  of all subsets of  $X$  as above. Then the family  $\mathcal{M}$  of  $\mu^*$ -measurable subsets of  $X$  is a  $\sigma$ -algebra containing  $\mathcal{A}$  (and hence  $\mathcal{B}(\mathcal{A})$ ) and all subsets of  $X$  which have  $\mu^*$  measure 0. The restriction  $\mu^*|_{\mathcal{M}}$  is a (countably additive) measure which agrees on  $\mathcal{A}$  with  $\mu$ . If  $\mu$  is  $\sigma$ -finite on  $\mathcal{A}$  (so that there are  $X_1, X_2, \dots \in \mathcal{A}$  with  $\mu(X_i) < \infty$  for all  $i$  and  $X = \cup_{i=1}^{\infty} X_i$ ), then  $\mu$  on  $\mathcal{B}(\mathcal{A})$  is the only extension of  $\mu$  on  $\mathcal{A}$  to  $\mathcal{B}(\mathcal{A})$ .*

In this way, beginning with the semialgebra  $\mathcal{H}$  of half-open subintervals of  $[0, 1)$  and  $\mu[a, b) = b - a$ , one arrives at Lebesgue measure on the  $\sigma$ -algebra  $\mathcal{M}$  of Lebesgue measurable sets and on  $\mathcal{B}(\mathcal{H})$  of Borel sets.

**5.2. Measures on shift spaces.** The measures that determine stochastic processes are also frequently constructed by specifying data on a semialgebra of cylinder sets. Given a finite alphabet  $A$ , denote by  $\Omega(A) = A^{\mathbb{Z}}$  and  $\Omega^+(A) = A^{\mathbb{N}}$  the sets of two and one-sided sequences, respectively, with entries from  $A$ . These are compact metric spaces, with  $d(x, y) = 2^{-n}$  when  $n = \inf\{|k| : x_k \neq y_k\}$ . In both cases, the *shift transformation*  $\sigma$  defined by  $(\sigma x)_n = x_{n+1}$  for all  $n$  is a homeomorphism.

Suppose (cf. [3]) that for every  $k = 1, 2, \dots$  we are given a function  $g_k : A^k \rightarrow [0, 1]$ , and that these functions satisfy, for all  $k$ ,

- (1)  $g_k(B) \geq 0$  for all  $B \in A^k$ ;
- (2)  $\sum_{i \in A} g_{k+1}(Bi) = g_k(B)$  for all  $B \in A^k$ ;
- (3)  $\sum_{i \in A} g_1(i) = 1$ .

Then Theorems 5.3 and 5.4 imply that there is a unique shift-invariant measure  $\mu$  on the Borel subsets of  $\Omega^+(A)$  such that for all  $k = 1, 2, \dots$  and  $B \in A^k$

$$(5.9) \quad \mu\{x \in \Omega^+(A) : x_1 \dots x_k = B\} = g_k(B).$$

If in addition the  $g_k$  also satisfy

- (4)  $\sum_{i \in A} g_{k+1}(iB) = g_k(B)$  for all  $k = 1, 2, \dots$  and all  $B \in A^k$ ,

then there is a unique shift-invariant measure  $\mu$  on the Borel subsets of  $\Omega^+(A)$  (also  $\Omega(A)$ ) such that for all  $n$ , all  $k = 1, 2, \dots$  and  $B \in A^k$

$$(5.10) \quad \mu\{x \in \Omega^+(A) : x_n \dots x_{n+k-1} = B\} = g_k(B).$$

This follows from the Carathéodory theorem by beginning with the semialgebra  $\mathcal{C}_1$  of anchored cylinder sets or the semialgebra  $\mathcal{C}$  of cylinder sets determined by finitely many consecutive coordinates, respectively.

There are two particularly important examples of this construction. First, let our finite alphabet be  $A = \{0, \dots, d-1\}$ , and let  $p = (p_0, \dots, p_{d-1})$  be a probability

vector: all  $p_i \geq 0$  and  $\sum_{i=0}^{d-1} p_i = 1$ . For any block  $B = b_1 \dots b_k \in A^k$ , define

$$(5.11) \quad g_k(B) = p_{b_1} \dots p_{b_k}.$$

The resulting measure  $\mu_p$  is the product measure on  $\Omega(A) = A^{\mathbb{Z}}$  of infinitely many copies of the probability measure determined by  $p$  on the finite sample space  $A$ . The measure-preserving system  $(\Omega, \mathcal{B}, \mu, \sigma)$  (with  $\mathcal{B}$  the  $\sigma$ -algebra of Borel subsets of  $\Omega(A)$ , or its completion), is denoted by  $\mathcal{B}(p)$  and is called the *Bernoulli system* determined by  $p$ . This system models an infinite number of independent repetitions of an experiment with finitely many outcomes, the  $i$ 'th of which has probability  $p_i$  on each trial.

This construction can be generalized to model stochastic processes which have some memory. Again let  $A = \{0, \dots, d-1\}$ , and let  $p = (p_0, \dots, p_{d-1})$  be a probability vector. Let  $P$  be a  $d \times d$  *stochastic matrix* with rows and columns indexed by  $A$ . This means that all entries of  $P$  are nonnegative, and the sum of the entries in each row is 1. We regard  $P$  as giving the transition probabilities between pairs of elements of  $A$ . Now we define for any block  $B = b_1 \dots b_k \in A^k$

$$(5.12) \quad g_k(B) = p_{b_1} P_{b_1 b_2} P_{b_2 b_3} \dots P_{b_{k-1} b_k}.$$

Using the  $g_k$  to define a nonnegative set function  $\mu_{p,P}$  on the semialgebra  $\mathcal{C}_1$  of anchored cylinder subsets of  $\Omega^+(A)$ , one can verify that  $\mu_{p,P}$  is (vacuously) finitely additive and countably subadditive on  $\mathcal{C}_1$  and therefore extends to a measure on the Borel  $\sigma$ -algebra of  $\Omega^+(A)$ , and its completion. The resulting stochastic process is a (one-step, finite-state) *Markov process*. If  $p$  and  $P$  also satisfy

$$(5.13) \quad pP = p,$$

then condition (4) above is satisfied, and the Markov process is stationary. In this case we call the (one or two-sided) measure-preserving system the *Markov shift* determined by  $p$  and  $P$ . Points in the space are conveniently pictured as infinite paths in a directed graph with vertices  $A$  and edges corresponding to the nonzero entries of  $P$ . A process with a longer memory, say of length  $m$ , can be produced by repeating the foregoing construction after recoding with a *sliding block code* to the new alphabet  $A^m$ : for each  $\omega \in \Omega(A)$ , let  $(\phi(\omega))_n = \omega_n \omega_{n+1} \dots \omega_{n+m-1} \in A^m$ .

**5.3. The Kolmogorov Consistency Theorem.** There is a generalization of this method to the construction of stochastic processes indexed by any set  $T$ . (Most frequently  $T = \mathbb{Z}, \mathbb{N}, \mathbb{R}, \mathbb{Z}^d$ , or  $\mathbb{R}^d$ ). We give a brief description, following [4].

Let  $T$  be an arbitrary index set. We aim to produce a  $\mathbb{R}$ -valued stochastic process indexed by  $T$ , that is to say, a Borel probability measure  $P$  on  $\Omega = \mathbb{R}^T$ , which has prespecified finite-dimensional distributions. Suppose that for every ordered  $k$ -tuple  $t_1, \dots, t_k$  of *distinct* elements of  $T$  we are given a Borel probability measure  $\mu_{t_1 \dots t_k}$  on  $\mathbb{R}^k$ . Denoting  $f \in \mathbb{R}^T$  also by  $(f_t : t \in T)$ , we want it to be the case that, for each  $k$ , each choice of distinct  $t_1, \dots, t_k \in T$ , and each Borel set  $B \subset \mathbb{R}^k$ ,

$$(5.14) \quad P\{(f_t : t \in T) : (f_{t_1}, \dots, f_{t_k}) \in B\} = \mu_{t_1 \dots t_k}(B).$$

For consistency, we will need, for example, that

$$(5.15) \quad \mu_{t_1 t_2}(B_1 \times B_2) = \mu_{t_2 t_1}(B_2 \times B_1), \quad \text{since}$$

$$(5.16) \quad P\{(f_{t_1}, f_{t_2}) \in A_1 \times A_2\} = P\{(f_{t_2}, f_{t_1}) \in A_2 \times A_1\}.$$

Thus we assume:

(1) For any  $k = 1, 2, \dots$  and permutation  $\pi$  of  $1, \dots, k$ , if  $\phi_\pi : \mathbb{R}^k \rightarrow \mathbb{R}^k$  is defined by

$$(5.17) \quad \phi_\pi(x_{\pi_1}, \dots, x_{\pi_k}) = (x_1, \dots, x_k),$$

then for all  $k$  and all Borel  $B \subset \mathbb{R}^k$

$$(5.18) \quad \mu_{t_1 \dots t_k}(B) = \mu_{t_{\pi_1} \dots t_{\pi_k}}(\phi_\pi^{-1} B).$$

Further, since leaving the value of one of the  $f_{t_j}$  free does not change the probability in (5.14), we also should have

(2) For any  $k = 1, 2, \dots$ , distinct  $t_1, \dots, t_k, t_{k+1} \in T$ , and Borel set  $B \subset \mathbb{R}^k$ ,

$$(5.19) \quad \mu_{t_1 \dots t_k}(B) = \mu_{t_1 \dots t_k t_{k+1}}(B \times \mathbb{R}).$$

**Theorem 5.5** (Kolmogorov Consistency Theorem [26]). *Given a system of probability measures  $\mu_{t_1 \dots t_k}$  as above indexed by finite ordered subsets of a set  $T$ , in order that there exist a probability measure  $P$  on  $\mathbb{R}^T$  satisfying (5.14) it is necessary and sufficient that the system satisfy (1) and (2) above.*

When  $T = \mathbb{Z}, \mathbb{R}$ , or  $\mathbb{N}$ , as in the example with the  $g_k$  above, the problem of consistency with regard to permutations of indices does not arise, since we tacitly use the order in  $T$  in specifying the finite-dimensional distributions.

In case  $T$  is a semigroup, by adding conditions on the given data  $\mu_{t_1 \dots t_k}$  it is possible to extend this construction also to produce *stationary* processes indexed by  $T$ , in parallel with the above constructions for  $T = \mathbb{Z}$  or  $\mathbb{N}$ .

## 6. INVARIANT MEASURES ON TOPOLOGICAL DYNAMICAL SYSTEMS

**6.1. Existence of invariant measures.** Let  $X$  be a compact metric space and  $T : X \rightarrow X$  a homeomorphism (although usually it is enough just that  $T$  be a continuous map). Denote by  $\mathcal{C}(X)$  the Banach space of continuous real-valued functions on  $X$  with the supremum norm and by  $\mathcal{M}(X)$  the set of Borel probability measures on  $X$ . Given the weak\* topology, according to which

$$(6.1) \quad \mu_n \rightarrow \mu \quad \text{if and only if} \quad \int_X f_n d\mu \rightarrow \int_X f d\mu \quad \text{for all } f \in \mathcal{C}(X),$$

$\mathcal{M}(X)$  is a convex subset of the dual space  $\mathcal{C}(X)^*$  of all continuous linear functionals from  $\mathcal{C}(X)$  to  $\mathbb{R}$ . With the weak\* topology it is metrizable and (by Alaoglu's Theorem) compact.

Denote by  $\mathcal{M}_T(X)$  the set of  $T$ -invariant Borel probability measures on  $X$ . A Borel probability measure  $\mu$  on  $X$  is in  $\mathcal{M}(X)$  if and only if

$$(6.2) \quad \mu(T^{-1}B) = \mu(B) \quad \text{for all Borel sets } B \subset X,$$

equivalently,

$$(6.3) \quad \mu(fT) = \int_X f \circ T d\mu = \int_X f d\mu \quad \text{for all } f \in \mathcal{C}(X).$$

**Proposition 6.1.** *For every compact topological dynamical system  $(X, T)$  (with  $X$  not empty) there is always at least one  $T$ -invariant Borel probability measure on  $X$ .*

*Proof.* Let  $m$  be any Borel probability measure on  $X$ . For example, we could pick a point  $x_0 \in X$  and let  $m$  be the point mass  $\delta_{x_0}$  at  $x_0$  defined by

$$(6.4) \quad \delta_{x_0}(f) = f(x_0) \quad \text{for all } f \in \mathcal{C}(X).$$

Form the averages

$$(6.5) \quad A_n m = \frac{1}{n} \sum_{i=0}^{n-1} mT^{-i},$$

which are also in  $\mathcal{M}(X)$ . By compactness,  $\{A_n m\}$  has a weak\* cluster point  $\mu$ , so that there is a subsequence

$$(6.6) \quad A_{n_k} m \rightarrow \mu \quad \text{weak*}.$$

Then  $\mu \in \mathcal{M}(X)$ ; and  $\mu$  is  $T$ -invariant, because for each  $f \in \mathcal{C}(X)$

$$(6.7) \quad |\mu(fT) - \mu(f)| = \lim_{k \rightarrow \infty} \frac{1}{n_k} |\mu(fT^{n_k}) - \mu(f)| = 0,$$

both terms inside the absolute value signs being bounded. □

**6.2. Ergodicity and unique ergodicity.** Among the  $T$ -invariant measures on  $X$  are the *ergodic* ones, those for which  $(X, \mathcal{B}, \mu, T)$  (with  $\mathcal{B}$  the  $\sigma$ -algebra of Borel subsets of  $X$ ) forms an ergodic measure-preserving system. This means that there are no proper  $T$ -invariant measurable sets:

$$(6.8) \quad B \in \mathcal{B}, \mu(T^{-1}B \Delta B) = 0 \quad \text{implies } \mu(B) = 0 \text{ or } 1.$$

Equivalently (using the Ergodic Theorem),  $(X, \mathcal{B}, \mu, T)$  is ergodic if and only if for each  $f \in L^1(X, \mathcal{B}, \mu)$

$$(6.9) \quad \frac{1}{n} \sum_{k=1}^{n-1} f(T^k x) \rightarrow \int_X f d\mu \quad \text{almost everywhere .}$$

It can be shown that the ergodic measures on  $(X, T)$  are exactly the *extreme points* of the compact convex set  $\mathcal{M}_T(X)$ , namely those  $\mu \in \mathcal{M}_T(X)$  for which there do not exist  $\mu_1, \mu_2 \in \mathcal{M}_T(x)$  with  $\mu_1 \neq \mu_2$  and  $s \in (0, 1)$  such that

$$(6.10) \quad \mu = s\mu_1 + (1-s)\mu_2.$$

The Krein-Milman Theorem states that in a locally convex topological vector space such as  $\mathcal{C}(X)^*$  every compact convex set is the closed convex hull of its extreme points. Thus every nonempty such set has extreme points, and so there always

exist ergodic measures for  $(X, T)$ . A topological dynamical system  $(X, T)$  is called *uniquely ergodic* if there is only *one*  $T$ -invariant Borel probability measure on  $X$ , in which case, by the foregoing discussion, that measure must be ergodic.

There are many examples of topological dynamical systems which are uniquely ergodic and of others which are not. For now, we just remark that translation by a generator on a compact monothetic group is always uniquely ergodic, while group endomorphisms and automorphisms tend to be not uniquely ergodic. Bernoulli and (nonatomic) Markov shifts are not uniquely ergodic, because they have many periodic orbits, each of which supports an ergodic measure.

## 7. FINDING FINITE INVARIANT MEASURES EQUIVALENT TO A QUASI-INVARIANT MEASURE

Let  $(X, \mathcal{B}, m)$  be a  $\sigma$ -finite measure space, and suppose that  $T : X \rightarrow X$  is an invertible *nonsingular* transformation. Thus we assume that  $T$  is one-to-one and onto (maybe after a set of measure 0 has been deleted), that  $T$  and  $T^{-1}$  are both measurable, so that

$$(7.1) \quad T\mathcal{B} = \mathcal{B} = T^{-1}\mathcal{B},$$

and that  $T$  and  $T^{-1}$  preserve the  $\sigma$ -ideal of sets of measure 0:

$$(7.2) \quad m(B) = 0 \quad \text{if and only if} \quad m(T^{-1}B) = 0 \quad \text{if and only if} \quad m(TB) = 0.$$

In this situation we say that  $m$  is *quasi-invariant* for  $T$ .

A nonsingular system  $(X, \mathcal{B}, m, T)$  as above may model a nonequilibrium situation in which events that are impossible (measure 0) at any time are also impossible at any other time. When dealing with such a system, it can be useful to know whether there is a  $T$ -invariant measure  $\mu$  that is equivalent to  $m$  (in the sense of absolute continuity—they have the same sets of measure 0—in which case we write  $\mu \sim m$ ), for then one would have available machinery of the measure-preserving situation, such as the Ergodic Theorem and entropy in their simplest forms. Also, it is most useful if the measures are  $\sigma$ -finite, so that tools such as the Radon-Nikodym and Tonelli-Fubini theorems will be available.

We may assume that  $m(X) = 1$ . For if  $X = \cup_{i=1}^{\infty} X_i$  with each  $X_i \in \mathcal{B}$  and  $m(X_i) < \infty$ , disjointifying (replace  $X_i$  by  $X_i \setminus X_{i-1}$  for  $i \geq 2$ ) and deleting any  $X_i$  that have measure 0, we may replace  $m$  by

$$(7.3) \quad \sum_{i=1}^{\infty} \frac{m|_{X_i}}{2^i m(X_i)}.$$

**Definition 7.1.** Let  $(X, \mathcal{B}, m)$  be a probability space and  $T : X \rightarrow X$  a nonsingular transformation. We say that  $A, B \in \mathcal{B}$  are  *$T$ -equivalent*, and write  $A \sim_T B$ , if there are two sequences of pairwise disjoint sets,  $A_1, A_2, \dots$  and  $B_1, B_2, \dots$  and integers  $n_1, n_2, \dots$  such that

$$(7.4) \quad A = \bigcup_{i=1}^{\infty} A_i, \quad B = \bigcup_{i=1}^{\infty} B_i, \quad \text{and} \quad T^{n_i} A_i = B_i \quad \text{for all } i.$$

**Definition 7.2.** Let  $(X, \mathcal{B}, m, T)$  be as above. A measurable set  $A \subset X$  is called *T-nonshrinkable* if  $A$  is not  $T$ -equivalent to any proper subset: whenever  $B \subset A$  and  $B \sim_T A$  we have  $m(A \setminus B) = 0$ .

**Theorem 7.3** (Hopf [23]). *Let  $(X, \mathcal{B}, m)$  be a probability space and  $T : X \rightarrow X$  a nonsingular transformation. There exists a finite invariant measure  $\mu \sim m$  if and only if  $X$  is  $T$ -nonshrinkable.*

*Proof.* We present just the easy half. If  $\mu \sim m$  is  $T$ -invariant and  $X \sim_T B$ , with corresponding decompositions  $X = \cup_{i=1}^{\infty} X_i, B = \cup_{i=1}^{\infty} B_i$ , then

$$(7.5) \quad \begin{aligned} \mu(B) &= \sum_{i=1}^{\infty} \mu(B_i) = \sum_{i=1}^{\infty} \mu(T^{n_i} X_i) \\ &= \sum_{i=1}^{\infty} \mu(X_i) = \mu(X), \end{aligned}$$

so that  $\mu(X \setminus B) = 0$  and hence  $m(X \setminus B) = 0$ .

For the converse, one tries to show that if  $X$  is  $T$ -nonshrinkable, then for each  $A \in \mathcal{B}$  the following limit exists:

$$(7.6) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} m(T^k A).$$

□

The condition of  $T$ -nonshrinkability not being easy to check, subsequent authors gave various necessary and sufficient conditions for the existence of a finite equivalent invariant measure:

**7.4** Dowker [11]. *Whenever  $A \in \mathcal{B}$  and  $m(A) > 0$ ,  $\liminf_{n \rightarrow \infty} m(T^n A) > 0$ .*

**7.5** Calderón [6]. *Whenever  $A \in \mathcal{B}$  and  $m(A) > 0$ ,  $\liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{\infty} m(T^k A) > 0$ .*

**7.6** Dowker [12]. *Whenever  $A \in \mathcal{B}$  and  $m(A) > 0$ ,  $\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{\infty} m(T^k A) > 0$ .*

Hajian and Kakutani [20] showed that the condition

$$(7.7) \quad m(A) > 0 \quad \text{implies} \quad \limsup_{n \rightarrow \infty} m(T^n A) > 0$$

is *not* sufficient for existence of a finite equivalent invariant measure. They also gave another necessary and sufficient condition.

**Definition 7.7.** A measurable set  $W \subset X$  is called *wandering* if the sets  $T^i W, i \in \mathbb{Z}$ , are pairwise disjoint.  $W$  is called *weakly wandering* if there are infinitely many integers  $n_i$  such that  $T^{n_i} W$  and  $T^{n_j} W$  are disjoint whenever  $n_i \neq n_j$ .

**Theorem 7.8** (Hajian-Kakutani [20]). *Let  $(X, \mathcal{B}, m)$  be a probability space and  $T : X \rightarrow X$  a nonsingular transformation. There exists a finite invariant measure  $\mu \sim m$  if and only if there are no weakly wandering sets of positive measure.*

8. FINDING  $\sigma$ -FINITE INVARIANT MEASURES EQUIVALENT TO A  
QUASI-INVARIANT MEASURE

**8.1. First necessary and sufficient conditions.** While being able to replace a quasi-invariant measure by an equivalent finite invariant measure would be great, it may be impossible, and then finding a  $\sigma$ -finite equivalent measure would still be pretty good. Hopf's nonshrinkability condition was extended to the  $\sigma$ -finite case by Halmos:

**Theorem 8.1** (Halmos [21]). *Let  $(X, \mathcal{B}, m)$  be a probability space and  $T : X \rightarrow X$  a nonsingular transformation. There exists a  $\sigma$ -finite invariant measure  $\mu \sim m$  if and only if  $X$  is a countable union of  $T$ -nonshrinkable sets.*

Another necessary and sufficient condition is given easily in terms of solvability of a *cohomological functional equation* involving the Radon-Nikodym derivative  $w$  of  $mT$  with respect to  $m$ , defined by

$$(8.1) \quad m(TB) = \int_B w \, dm \quad \text{for all } B \in \mathcal{B}.$$

**Proposition 8.2** [21]. *Let  $(X, \mathcal{B}, m)$  be a probability space and  $T : X \rightarrow X$  a nonsingular transformation. There exists a  $\sigma$ -finite invariant measure  $\mu \sim m$  if and only if there is a measurable function  $f : X \rightarrow (0, \infty)$  such that*

$$(8.2) \quad f(Tx) = w(x)f(x) \quad \text{a.e.}$$

*Proof.* If  $\mu \sim m$  is  $\sigma$ -finite and  $T$ -invariant, let  $f = dm/d\mu$  be the Radon-Nikodym derivative of  $m$  with respect to  $\mu$ , so that

$$(8.3) \quad m(B) = \int_B f \, d\mu \quad \text{for all } B \in \mathcal{B}.$$

Then for all  $B \in \mathcal{B}$ , since  $\mu T = \mu$ ,

$$(8.4) \quad \begin{aligned} m(TB) &= \int_{TB} f \, d\mu = \int_B fT \, d\mu, \quad \text{while also} \\ m(TB) &= \int_B w \, dm = \int_B wf \, dm, \end{aligned}$$

so that  $fT = wf$  a.e..

Conversely, given such an  $f$ , let

$$(8.5) \quad \mu(B) = \int_B \frac{1}{f} \, dm \quad \text{for all } B \in \mathcal{B}.$$

Then for all  $B \in \mathcal{B}$

$$(8.6) \quad \begin{aligned} \mu(TB) &= \int_{TB} \frac{1}{f} \, dm = \int_B \frac{1}{fT} \, dmT \\ &= \int_B \frac{1}{fT} w \, dm = \int_B \frac{1}{f} \, dm = \mu(B). \end{aligned}$$

□

## 8.2. Conservativity and recurrence.

**Definition 8.3.** A nonsingular system  $(X, \mathcal{B}, m, T)$  (with  $m(X) = 1$ ) is called *conservative* if there are no wandering sets of positive measure. It is called *completely dissipative* if there is a wandering set  $W$  such that

$$(8.7) \quad m\left(\bigcup_{i=-\infty}^{\infty} T^i W\right) = m(X).$$

Note that if  $(X, \mathcal{B}, m, T)$  is completely dissipative, it is easy to construct a  $\sigma$ -finite equivalent invariant measure. With  $W$  as above, define  $\mu = m$  on  $W$  and push  $\mu$  along the orbit of  $W$ , letting  $\mu = mT^{-n}$  on each  $T^n W$ . We want to claim that this allows us to restrict attention to the conservative case, which follows once we know that the system splits into a conservative and a completely dissipative part.

**Theorem 8.4** (Hopf Decomposition [24]). *Given a nonsingular map  $T$  on a probability space  $(X, \mathcal{B}, m)$ , there are disjoint measurable sets  $C$  and  $D$  such that*

- (1)  $X = C \cup D$ ;
- (2)  $C$  and  $D$  are invariant:  $TC = C = T^{-1}C$ ,  $TD = D = T^{-1}D$ ;
- (3)  $T|_C$  is conservative;
- (4) If  $D \neq \emptyset$ , then  $T|_D$  is completely dissipative.

*Proof.* Assume that the family  $\mathcal{W}$  of wandering sets with positive measure is nonempty, since otherwise we can take  $C = X$  and  $D = \emptyset$ . Partially order  $\mathcal{W}$  by

$$(8.8) \quad W_1 \leq W_2 \quad \text{if} \quad m(W_1 \setminus W_2) = 0.$$

We want to apply Zorn's Lemma to find a maximal element in  $\mathcal{W}$ . Let  $\{W_\lambda : \lambda \in \Lambda\}$  be a chain (linearly ordered subset) in  $\mathcal{W}$ . Just forming  $\cup_{\lambda \in \Lambda} W_\lambda$  may result in a nonmeasurable set, so we have to use the measure to form a measure-theoretic essential supremum of the chain. So let

$$(8.9) \quad s = \sup\{m(W_\lambda) : \lambda \in \Lambda\},$$

so that  $s \in (0, 1]$ . If there is a  $\lambda$  such that  $m(W_\lambda) = s$ , let  $W$  be that  $W_\lambda$ . Otherwise, for each  $k$  choose  $\lambda_k \in \Lambda$  so that

$$(8.10) \quad s_k = m(W_{\lambda_k}) \uparrow s,$$

and let

$$(8.11) \quad W = \bigcup_{k=1}^{\infty} W_{\lambda_k}.$$

We claim that in either case  $W$  is an upper bound for the chain  $\{W_\lambda : \lambda \in \Lambda\}$ . In both cases we have  $m(W) = s$ .

Note that if  $\lambda, \tau \in \Lambda$  are such that  $m(W_\lambda) \leq m(W_\tau)$ , then  $W_\lambda \leq W_\tau$ . For if  $W_\tau \leq W_\lambda$ , then  $m(W_\tau \setminus W_\lambda) = 0$ , and thus

$$(8.12) \quad \begin{aligned} m(W_\tau) &= m(W_\tau \cap W_\lambda) + m(W_\tau \setminus W_\lambda) = m(W_\tau \cap W_\lambda) \\ &\leq m(W_\tau \cap W_\lambda) + m(W_\lambda \setminus W_\tau) \\ &= m(W_\lambda) \leq m(W_\tau), \end{aligned}$$

so that  $m(W_\lambda \setminus W_\tau) = 0$ ,  $W_\lambda \leq W_\tau$ , and hence  $W_\lambda = W_\tau$ .

Thus in the first case  $W \in \mathcal{W}$  is an upper bound for the chain. In the second case, by discarding the measure 0 set

$$(8.13) \quad Z = \bigcup_{k=1}^{\infty} (W_{\lambda_k} \setminus W_{\lambda_{k+1}}),$$

we may assume that  $W$  is the *increasing* union of the  $W_{\lambda_k}$ . Then  $W \geq W_{\lambda_k}$  for all  $k$ , and  $W$  is wandering: if some  $T^n W \cap W \neq \emptyset$ , then there must be a  $k$  such that  $T^n W_{\lambda_k} \cap W_{\lambda_k} \neq \emptyset$ .

Moreover,  $W_\lambda \leq W$  for all  $\lambda \in \Lambda$ . For let  $\lambda \in \Lambda$  be given. Choose  $k$  with  $s_k = m(W_{\lambda_k}) > m(W_\lambda)$ . By the above, we have  $W_{\lambda_k} \geq W_\lambda$ . Since  $W$  is the increasing union of the  $W_{\lambda_k}$ , we have  $W \geq W_{\lambda_k}$  for all  $k$ . Therefore  $W \geq W_\lambda$ , and  $W$  is an upper bound in  $\mathcal{W}$  for the given chain.

By Zorn's Lemma, there is a maximal element  $W^*$  in  $\mathcal{W}$ . Then  $D = \bigcup_{i=-\infty}^{\infty} T^i W^*$  is  $T$ -invariant,  $T|_D$  is completely dissipative, and  $C = X \setminus D$  cannot contain any wandering set of positive measure, by maximality of  $W^*$ , so  $T|_C$  is conservative.  $\square$

Because of this decomposition, when looking for a  $\sigma$ -finite equivalent invariant measure we may assume that the nonsingular system  $(X, \mathcal{B}, m, T)$  is conservative, for if not we can always construct one on the dissipative part.

*Remark 8.5.* If  $(X, \mathcal{B}, m)$  is nonatomic and  $T : X \rightarrow X$  is nonsingular, invertible, and *ergodic*, in the sense that if  $A \in \mathcal{B}$  satisfies  $T^{-1}A = A = TA$  then either  $m(A) = 0$  or  $m(A^c) = 0$ , then  $T$  is conservative. For if  $W$  is a wandering set of positive measure, taking any  $A \subset W$  with  $0 < m(A) < m(W)$  and forming  $\bigcup_{i=-\infty}^{\infty} T^i A$  will produce an invariant set of positive measure whose complement also has positive measure.

We want to reduce the problem of existence of a  $\sigma$ -finite equivalent invariant measure to that of a finite one by using first-return maps to sets of finite measure. For this purpose it will be necessary to know that every conservative nonsingular system is *recurrent*: almost every point of each set of positive measure returns at some future time to that set. This is easy to see, because for each  $B \in \mathcal{B}$ , the set

$$(8.14) \quad B^* = \bigcup_{i=1}^{\infty} T^{-i} B$$

is wandering. In fact much more is true.

**Theorem 8.6** [21]. *For any nonsingular system  $(X, \mathcal{B}, m, T)$  the following properties are equivalent:*

(1) *The system is incompressible: for each  $B \in \mathcal{B}$  such that  $T^{-1}B \subset B$ , we have  $m(B \setminus T^{-1}B) = 0$ .*

(2) *The system is recurrent; for each  $B \in \mathcal{B}$ , with  $B^*$  defined as above,  $m(B \setminus B^*) = 0$ .*

(3) *The system is conservative: there are no wandering sets of positive measure.*

(4) *The system is infinitely recurrent: for each  $B \in \mathcal{B}$ , almost every point of  $B$  returns to  $B$  infinitely many times, equivalently,*

$$(8.15) \quad m\left(B \setminus \bigcap_{n=0}^{\infty} \bigcup_{i=n}^{\infty} T^{-i}B\right) = m\left(B \setminus \bigcap_{n=0}^{\infty} T^{-n}B^*\right) = 0.$$

There is a very slick proof by F. B. Wright [38] of this result in the even more general situation of a Boolean  $\sigma$ -algebra homomorphism (reproduced in [28]).

**8.3. Using first-return maps, and counterexamples to existence.** Now given a nonsingular conservative system  $(X, \mathcal{B}, m, T)$  and a set  $B \in \mathcal{B}$ , for each  $x \in B$  there is a smallest  $n_B(x) \geq 1$  such that

$$(8.16) \quad T^{n_B}(x) \in B.$$

We define the *first-return map*  $T_B : B \rightarrow B$  by

$$(8.17) \quad T_B(x) = T^{n_B(x)}(x) \quad \text{for all } x \in B.$$

Using derivative maps, it is easy to reduce the problem of existence of a  $\sigma$ -finite equivalent invariant measure to that of existence of finite equivalent invariant measures, in a way.

**Theorem 8.7** (see [14]). *Let  $T$  be a conservative nonsingular transformation on a probability space  $(X, \mathcal{B}, m)$ . Then there is a  $\sigma$ -finite  $T$ -invariant measure  $\mu \sim m$  if and only if there is an increasing sequence of sets  $B_n \in \mathcal{B}$  with  $\cup_{n=1}^{\infty} B_n = X$  such that for each  $n$  the first-return map  $T_{B_n}$  has a finite invariant measure equivalent to  $m$  restricted to  $B_n$ .*

*Proof.* Given a  $\sigma$ -finite equivalent invariant measure  $\mu$ , let the  $B_n$  be sets of finite  $\mu$ -measure that increase to  $X$ . Conversely, given such a sequence  $B_n$  with finite invariant measures  $\mu_n$  for the first-return maps  $T_{B_n}$ , extend  $\mu_1$  in the obvious way to an (at least  $\sigma$ -finite) invariant measure on the full orbit  $A_1 = \cup_{i=-\infty}^{\infty} T^i B_1$ . Then replace  $B_2$  by  $B_2 \setminus A_1$ , and continue.  $\square$

There are many more checkable conditions for existence of a  $\sigma$ -finite equivalent invariant measure in the literature. There are also examples of invertible ergodic nonsingular systems for which *there does not exist* any  $\sigma$ -finite equivalent invariant

measure due to Ornstein [27] and subsequently Chacon [9], Brunel [5], L. Arnold [2], and others.

**8.4. Invariant measures for maps of the interval or circle.** Finally we mention sample theorems from a huge array of such results about existence of finite invariant measures for maps of an interval or of the circle.

**Theorem 8.8** (“Folklore Theorem” [1]). *Let  $X = (0, 1)$  and denote by  $m$  Lebesgue measure on  $X$ . Let  $T : X \rightarrow X$  be a map for which there is a finite or countable partition  $\alpha = \{A_i\}$  of  $X$  into half-open intervals  $[a_i, b_i)$  satisfying the following conditions. Denote by  $A_i^0$  the interior of each interval  $A_i$ . Suppose that*

(1) *for each  $i$ ,  $T : A_i^0 \rightarrow X$  is one-to-one and onto;*

(2)  *$T$  is  $C^2$  on each  $A_i^0$ ;*

(3) *there is an  $n$  such that*

$$(8.18) \quad \inf_i \inf_{x \in A_i^0} |(T^n)'(x)| > 1;$$

(4) *for each  $i$ ,*

$$(8.19) \quad \sup_{x, y, z \in A_i^0} \left| \frac{T''(x)}{T'(y)T'(z)} \right| < \infty.$$

*Then for each measurable set  $B$ ,  $\lim_{n \rightarrow \infty} m(T^{-n}B) = \mu(B)$  exists and defines the unique  $T$ -invariant ergodic probability measure on  $X$  that is equivalent to  $m$ .*

*Moreover, the partition  $\alpha$  is weakly Bernoulli for  $T$ , so that the natural extension of  $T$  is isomorphic to a Bernoulli system.*

A key example to which the theorem applies is that of the *Gauss map*  $Tx = 1/x \bmod 1$  with the partition for which  $A_i = [1/(i+1), 1/i)$  for each  $i = 1, 2, \dots$ . Coding orbits to  $\mathbb{N}^{\mathbb{N}}$  by letting  $a(x) = a_i$  if  $x \in A_i$  carries  $T$  to the shift on the continued fraction expansion  $[a_1, a_2, \dots]$  of  $x$ . It was essentially known already to Gauss that  $T$  preserves the measure whose density with respect to Lebesgue measure is  $1/((1+x) \log 2)$ .

**Theorem 8.9** (see [22]). *Let  $X = S^1$ , the unit circle, and let  $T : X \rightarrow X$  be a (noninvertible)  $C^2$  map which is expanding, in the sense that  $|T'(x)| > 1$  everywhere. Then there is a unique finite invariant measure  $\mu$  equivalent to Lebesgue measure  $m$ , and in fact  $\mu$  is ergodic and the Radon-Nikodym derivative  $d\mu/dm$  has a continuous version.*

## 9. SOME MATHEMATICAL BACKGROUND

### 9.1. Lebesgue spaces.

**Definition 9.1.** Two measure spaces  $(X, \mathcal{B}, \mu)$  and  $(Y, \mathcal{C}, \nu)$  are *isomorphic* (sometimes also called *isomorphic mod 0*) if there are subsets  $X_0 \subset X$  and  $Y_0 \subset Y$  such that  $\mu(X_0) = 0 = \nu(Y_0)$  and a one-to-one onto map  $\phi : X \setminus X_0 \rightarrow Y \setminus Y_0$  such that  $\phi$  and  $\phi^{-1}$  are measurable and  $\mu(\phi^{-1}C) = \nu(C)$  for all measurable  $C \subset Y \setminus Y_0$ .

**Definition 9.2.** A *Lebesgue space* is a finite measure space that is isomorphic to a measure space consisting of a (possibly empty) finite subinterval of  $\mathbb{R}$  with the  $\sigma$ -algebra of Lebesgue measurable sets and Lebesgue measure, possibly together with countably many atoms (point masses).

The *measure algebra* of a measure space  $(X, \mathcal{B}, \mu)$  consists of the pair  $(\hat{\mathcal{B}}, \hat{\mu})$ , with  $\hat{\mathcal{B}}$  the Boolean  $\sigma$ -algebra (see Section 3.2, 3) of  $\mathcal{B}$  modulo the  $\sigma$ -ideal of sets of measure 0, together with the operations induced by set operations in  $\mathcal{B}$ , and  $\hat{\mu}$  is induced on  $\hat{\mathcal{B}}$  by  $\mu$  on  $\mathcal{B}$ . Every measure algebra  $(\hat{\mathcal{B}}, \hat{\mu})$  is a metric space with the metric  $d(A, B) = \hat{\mu}(A \Delta B)$  for all  $A, B \in \hat{\mathcal{B}}$ . It is *nonatomic* if whenever  $A, B \in \hat{\mathcal{B}}$  and  $A < B$  (which means  $A \wedge B = A$ ), either  $A = 0$  or  $A = B$ . A *homomorphism of measure algebras*  $\psi : (\hat{\mathcal{C}}, \hat{\nu}) \rightarrow (\hat{\mathcal{B}}, \hat{\mu})$  is a Boolean  $\sigma$ -algebra homomorphism such that  $\hat{\mu}(\hat{C}) = \hat{\nu}(C)$  for all  $\hat{C} \in \hat{\mathcal{C}}$ . The inverse of any factor map  $\phi : X \rightarrow Y$  from a measure space  $(X, \mathcal{B}, \mu)$  to a measure space  $(Y, \mathcal{C}, \nu)$  induces a homomorphism of measure algebras  $(\hat{\mathcal{C}}, \hat{\nu}) \rightarrow (\hat{\mathcal{B}}, \hat{\mu})$ . We say that a measure algebra is *normalized* if the measure of the maximal element is 1:  $\hat{\mu}(0') = 1$ .

We work within the class of Lebesgue spaces because (1) they are the ones commonly encountered in the wide range of naturally arising examples; (2) they allow us to assume if we wish that we are dealing with a familiar space such as  $[0, 1]$  or  $\{0, 1\}^{\mathbb{N}}$ ; and (3) they have the following useful properties.

- (Carathéodory [7]) Every normalized and nonatomic measure algebra whose associated metric space is separable (has a countable dense set) is measure-algebra isomorphic with the measure algebra of the unit interval with Lebesgue measure.
- (von Neumann [37]) Every complete separable metric space with a Borel probability measure on the completion of the Borel sets is a Lebesgue space.
- (von Neumann [37]) Every homomorphism  $\psi : (\hat{\mathcal{C}}, \hat{\nu}) \rightarrow (\hat{\mathcal{B}}, \hat{\mu})$  of the measure algebras of two Lebesgue spaces  $(Y, \mathcal{C}, \nu)$  and  $(X, \mathcal{B}, \mu)$  comes from a factor map: there are a set  $X_0 \subset X$  with  $\mu(X_0) = 0$  and a measurable map  $\phi : X \setminus X_0 \rightarrow Y$  such that  $\psi$  coincides with the map induced by  $\phi^{-1}$  from  $\hat{\mathcal{C}}$  to  $\hat{\mathcal{B}}$ .

9.1.1. *Rokhlin theory.* V. A. Rokhlin [31] provided an axiomatic, intrinsic characterization of Lebesgue spaces. The key ideas are the concept of a basis and the correspondence of factors with complete sub- $\sigma$ -algebras and (not necessarily finite or countable) measurable partitions of a special kind.

**Definition 9.3.** A *basis* for a complete measure space  $(X, \mathcal{B}, \mu)$  is a countable family  $\mathcal{C} = \{C_1, C_2, \dots\}$  of measurable sets which

*generates  $\mathcal{B}$* : For each  $B \in \mathcal{B}$  there is  $C \in \mathcal{B}(\mathcal{C})$  (the smallest  $\sigma$ -algebra of subsets of  $X$  that contains  $\mathcal{C}$ ) such that  $B \subset C$  and  $\mu(C \setminus B) = 0$ ; and

*separates the points of  $X$* : For each  $x, y \in X$  with  $x \neq y$ , there is  $C_i \in \mathcal{C}$  such that either  $x \in C_i, y \notin C_i$  or else  $y \in C_i, x \notin C_i$ .

Coarse sub- $\sigma$ -algebras of  $\mathcal{B}$  may not separate points of  $X$  and thus may lead to equivalence relations, partitions, and factor maps. Partitions of the following kind deserve careful attention.

**Definition 9.4.** Let  $(X, \mathcal{B}, \mu)$  be a complete measure space and  $\xi$  a partition of  $X$ , meaning that up to a set of measure 0,  $X$  is the union of the elements of  $\xi$ , which are pairwise disjoint up to sets of measure 0. We call  $\xi$  an *R-partition* if there is a countable family  $D = \{D_1, D_2, \dots\}$  of  $\xi$ -saturated sets (that is, each  $D_i$  is a union of elements of  $\xi$ ) such that

$$(9.1) \quad \begin{aligned} &\text{for all distinct } E, F \in \xi, \text{ there is } D_i \text{ such that either} \\ &E \subseteq D_i, F \not\subseteq D_i \quad (\text{so } F \subset D_i^c) \\ &\text{or } F \subseteq D_i, E \not\subseteq D_i \quad (\text{so } E \subset D_i^c). \end{aligned}$$

Any such family  $D$  is called a *basis* for  $\xi$ .

Note that each element of an *R-partition* is necessarily measurable: if  $C \in \xi$  with basis  $\{D_i\}$ , then

$$(9.2) \quad C = \bigcap \{D_i : C \subseteq D_i\}.$$

Every countable or finite measurable partition of a complete measure space is an *R-partition*. The orbit partition of a measure-preserving transformation is often *not* an *R-partition*. (For example, if the transformation is ergodic, the corresponding factor space will be trivial, consisting of just one cell, rather than corresponding to the partition into orbits as required.)

For any set  $B \subset X$ , let  $B^0 = B$  and  $B^1 = B^c = X \setminus B$ .

**Definition 9.5.** A basis  $\mathcal{C} = \{C_1, C_2, \dots\}$  for a complete measure space  $(X, \mathcal{B}, \mu)$  is called *complete*, and the space is called *complete with respect to the basis*, if for every 0,1-sequence  $e \in \{0, 1\}^{\mathbb{N}}$ ,

$$(9.3) \quad \bigcap_{i=1}^{\infty} C_i^{e_i} \neq \emptyset.$$

$\mathcal{C}$  is called *complete mod 0* (and  $(X, \mathcal{B}, \mu)$  is called *complete mod 0 with respect to  $\mathcal{C}$* , if there is a complete measure space  $(X', \mathcal{B}', \mu')$  with a complete basis  $\mathcal{C}'$  such that  $X$  is a full-measure subset of  $X'$ , and  $C_i = C'_i \cap X$  for all  $i = 1, 2, \dots$ .

From the definition of basis, each intersection in (9.3) contains at most one point. The space  $\{0, 1\}^{\mathbb{N}}$  with Bernoulli  $1/2, 1/2$  measure on the completion of the Borel sets has the complete basis  $C_i = \{\omega : \omega_i = 0\}$ .

**Proposition 9.6.** *If a measure space is complete mod 0 with respect to one basis, then it is complete mod 0 with respect to every basis.*

**Theorem 9.7.** ([31]) *A measure space is a Lebesgue space (that is, isomorphic mod 0 with the usual Lebesgue measure space of a possibly empty subinterval of  $\mathbb{R}$  possibly together with countably many atoms) if and only if it has a complete basis.*

In a Lebesgue space  $(X, \mathcal{B}, \mu)$  there is a one-to-one onto correspondence between complete sub- $\sigma$ -algebras of  $\mathcal{B}$  (that is, those for which the restriction of the measure yields a complete measure space) and  $R$ -partitions of  $X$ :

Given an  $R$ -partition  $\xi$ , let  $\mathcal{B}(\xi)$  denote the  $\sigma$ -algebra generated by  $\xi$ , which consists of all sets in  $\mathcal{B}$  that are  $\xi$ -saturated—unions of members of  $\xi$ —and let  $\overline{\mathcal{B}}(\xi)$  denote the completion of  $\mathcal{B}(\xi)$  with respect to  $\mu$ .

Conversely, given a complete sub- $\sigma$ -algebra  $\mathcal{C} \subset \mathcal{B}$ , define an equivalence relation on  $X$  by  $x \sim y$  if for all  $A \in \mathcal{C}$ , either  $x, y \in A$  or else  $x, y \in A^c$ . The measure algebra  $(\hat{\mathcal{C}}, \hat{\mu})$  has a countable dense set  $\hat{\mathcal{C}}_0$  (take a countable dense set  $\{\hat{B}_i\}$  for  $(\hat{\mathcal{B}}, \hat{\mu})$  and, for each  $i, j$  for which it is possible, choose  $\hat{C}_{ij}$  within distance  $1/2^j$  of  $\hat{B}_i$ ). Then representatives  $C_i \in \mathcal{C}$  of the  $\hat{C}_{ij}$  will be a basis for the partition  $\xi$  corresponding to the equivalence relation  $\sim$ .

Given any family  $\{\mathcal{B}_\lambda\}$ , of complete sub- $\sigma$ -algebras of  $\mathcal{B}$ , their *join* is the intersection of all the sub- $\sigma$ -algebras that contain their union:

$$(9.4) \quad \bigvee_{\lambda} \mathcal{B}_\lambda = \mathcal{B}(\bigcup_{\lambda} \mathcal{B}_\lambda),$$

and their *infimum* is just their intersection:

$$(9.5) \quad \bigwedge_{\lambda} \mathcal{B}_\lambda = \mathcal{B}(\bigcap_{\lambda} \mathcal{B}_\lambda).$$

These  $\sigma$ -algebra operations correspond to the supremum and infimum of the corresponding families of  $R$ -partitions. We say that a partition  $\xi_1$  is *finer* than a partition  $\xi_2$ , and write  $\xi_1 \geq \xi_2$ , if every element of  $\xi_2$  is a union of elements of  $\xi_1$ . Given any family  $\{\xi_\lambda\}$  of  $R$ -partitions, there is a coarsest  $R$ -partition  $\bigvee_{\lambda} \xi_\lambda$  which refines all of them, and a finest  $R$ -partition  $\bigwedge_{\lambda} \xi_\lambda$  which is coarser than all of them. We have

$$(9.6) \quad \bigvee_{\lambda} \mathcal{B}(\xi_\lambda) = \mathcal{B}(\bigvee_{\lambda} \xi_\lambda), \quad \bigwedge_{\lambda} \mathcal{B}(\xi_\lambda) = \mathcal{B}(\bigwedge_{\lambda} \xi_\lambda).$$

Now we discuss the relationship among factor maps  $\phi : X \rightarrow Y$  from a Lebesgue space  $(X, \mathcal{B}, \mu)$  to a complete measure space  $(Y, \mathcal{C}, \nu)$ , complete sub- $\sigma$ -algebras of  $\mathcal{B}$ , and  $R$ -partitions of  $X$ . Given such a factor map  $\phi$ ,  $\mathcal{B}_Y = \phi^{-1}\mathcal{C}$  is a complete sub- $\sigma$ -algebra of  $\mathcal{B}$ , and the equivalence relation  $x_1 \sim x_2$  if  $\phi(x_1) = \phi(x_2)$  determines an  $R$ -partition  $\xi_Y$ . (A basis for  $\xi$  can be formed from a countable dense set in  $\hat{\mathcal{B}}_Y$  as above.)

Conversely, given a complete sub- $\sigma$ -algebra  $\mathcal{C} \subset \mathcal{B}$ , the identity map  $(X, \mathcal{B}, \mu) \rightarrow (X, \mathcal{C}, \mu)$  is a factor map. Alternatively, given an  $R$ -partition of  $X$ , we can form a measure space  $(X/\xi, \mathcal{B}(\xi), \mu_\xi)$  and a factor map  $\phi_\xi : X \rightarrow X/\xi$  as follows. The space  $X/\xi$  is just  $\xi$  itself; that is, the points of  $X/\xi$  are the members (cells, or atoms) of

the partition  $\xi$ .  $\mathcal{B}(\xi)$  consists of the  $\xi$ -saturated sets in  $\mathcal{B}$  considered as subsets of  $\xi$ , and  $\mu_\xi$  is the restriction of  $\mu$  to  $\mathcal{B}(\xi)$ . Completeness of  $(X, \mathcal{B}, \mu)$  forces completeness of  $(X/\xi, \mathcal{B}(\xi), \mu_\xi)$ . The map  $\phi_\xi : X \rightarrow X/\xi$  is defined by letting  $\phi(x) = \xi(x) =$  the element of  $\xi$  to which  $x$  belongs. Thus for a Lebesgue space  $(X, \mathcal{B}, \mu)$ , there is a perfect correspondence among images under factor maps, complete sub- $\sigma$ -algebras of  $\mathcal{B}$ , and  $R$ -partitions of  $X$ .

**Theorem 9.8.** *If  $(X, \mathcal{B}, \mu)$  is a Lebesgue space and  $(Y, \mathcal{C}, \nu)$  is a complete measure space that is the image of  $(X, \mathcal{B}, \mu)$  under a factor map, then  $(Y, \mathcal{C}, \nu)$  is also a Lebesgue space.*

**Theorem 9.9.** *Let  $(X, \mathcal{B}, \mu)$  be a Lebesgue space,  $(Y, \mathcal{C}, \nu)$  a separable measure space (that is, one with a countable basis as above, equivalently one with a countable dense set in its measure algebra), and  $\phi : X \rightarrow Y$  a measurable map ( $\phi^{-1}\mathcal{C} \subset \mathcal{B}$ ). Then  $\phi$  is also forward measurable: if  $A \subset X$  is measurable, then  $\phi(A) \subset Y$  is measurable.*

**Theorem 9.10.** *Let  $(X, \mathcal{B}, \mu)$  be a Lebesgue space.*

(1) *Every measurable subset of  $X$ , with the restriction of  $\mathcal{B}$  and  $\mu$ , is a Lebesgue space. Conversely, if a subset  $A$  of  $X$  with the restrictions of  $\mathcal{B}$  and  $\mu$  is a Lebesgue space, then  $A$  is measurable ( $A \in \mathcal{B}$ ).*

(2) *The product of countably many Lebesgue spaces is a Lebesgue space.*

(3) *Every measure algebra isomorphism of  $(\hat{\mathcal{B}}, \hat{\mu})$  (defined as above) is induced by a point isomorphism mod 0.*

9.1.2. *Disintegration of measures.* Every  $R$ -partition  $\xi$  of a Lebesgue space  $(X, \mathcal{B}, \mu)$  has associated with it a *canonical system of conditional measures*: Using the notation of the preceding section, for  $\mu_\xi$ -almost every  $C \in X/\xi$ , there are a  $\sigma$ -algebra  $\mathcal{B}_C$  of subsets of  $C$  and a measure  $m_C$  on  $\mathcal{B}_C$  such that:

- (1)  $(C, \mathcal{B}_C, m_C)$  is a Lebesgue space;
- (2) for every  $A \in \mathcal{B}$ ,  $A \cap C \in \mathcal{B}_C$  for  $\mu_\xi$ -almost every  $C \in \xi$ ;
- (3) for every  $A \in \mathcal{B}$ , the map  $C \rightarrow m_C(A \cap C)$  is  $\mathcal{B}(\xi)$ -measurable on  $X/\xi$ ;
- (4) for every  $A \in \mathcal{B}$ ,

$$(9.7) \quad \mu(A) = \int_{X/\xi} m_C(A \cap C) d\mu_\xi(C).$$

It follows that for  $f \in L^1(X)$ , (a version of) its conditional expectation (see the next section) with respect to the factor algebra corresponding to  $\xi$  is given by

$$(9.8) \quad \mathbb{E}(f|\mathcal{B}(\xi)) = \int_C f dm_C \quad \text{on } \mu_\xi\text{-a.e. } C \in \xi,$$

since the right-hand side is  $\mathcal{B}(\xi)$ -measurable and for each  $A \in \mathcal{B}(\xi)$ , its integral over any  $B \in \mathcal{B}(\xi)$  is, as required,  $\mu(A \cap B)$  (use the formula on  $B/(\xi|B)$ ).

It can be shown that a canonical system of conditional measures for an  $R$ -partition of a Lebesgue space is essentially unique, in the sense that any two measures  $m_C$  and  $m'_C$  will be equal for  $\mu_\xi$ -almost all  $C \in \xi$ . Also, any partition of a Lebesgue space that has a canonical system of conditional measures must be an  $R$ -partition.

These conditional systems of measures can be used to prove the ergodic decomposition theorem and to show that every factor situation is essentially projection of a skew product onto the base (see [32]).

**Theorem 9.11.** *Let  $(X, \mathcal{B}, \mu)$  be a Lebesgue space. If  $\xi$  is an  $R$ -partition of  $X$ ,  $\{(C, \mathcal{B}_C, m_C)\}$  is a canonical system of conditional measures for  $\xi$ , and  $A \in \mathcal{B}$ , define  $\mu(A|C) = m_C(A \cap C)$ . Then:*

- (1) for every  $A \in \mathcal{B}$ ,  $\mu(A|\xi(x))$  is a measurable function of  $x \in X$ ;
- (2) if  $(\xi_n)$  is an increasing sequence of  $R$ -partitions of  $X$ , then for each  $A \in \mathcal{B}$

$$(9.9) \quad \mu(A|\xi_n(x)) \rightarrow \mu(A|\bigvee_n \xi_n(X)) \quad \text{a.e. } d\mu;$$

- (3) if  $(\xi_n)$  is a decreasing sequence of  $R$ -partitions of  $X$ , then for each  $A \in \mathcal{B}$

$$(9.10) \quad \mu(A|\xi_n(x)) \rightarrow \mu(A|\bigwedge_n \xi_n(X)) \quad \text{a.e. } d\mu.$$

This is a consequence of the Martingale and Reverse Martingale Convergence Theorems. The statements hold just as well for  $f \in L^1(X)$  as for  $f = \mathbf{1}_A$  for some  $A \in \mathcal{B}$ .

**9.2. Conditional expectation.** Let  $(X, \mathcal{B}, \mu)$  be a  $\sigma$ -finite measure space,  $f \in L^1(X)$ , and  $\mathcal{F} \subset \mathcal{B}$  a sub- $\sigma$ -algebra of  $\mathcal{B}$ . Then

$$(9.11) \quad \nu(F) = \int_F f d\mu$$

defines a finite signed measure on  $\mathcal{F}$  which is absolutely continuous with respect to  $\mu$  restricted to  $\mathcal{F}$ . So by the Radon-Nikodym Theorem there is a function  $g \in L^1(X, \mathcal{F}, \mu)$  such that

$$(9.12) \quad \nu(F) = \int_F g d\mu \quad \text{for all } F \in \mathcal{F}.$$

Any such function  $g$ , which is unique as an element of  $L^1(X, \mathcal{F}, \mu)$  (and determined only up to sets of  $\mu$ -measure 0) is called a version of the *conditional expectation* of  $f$  with respect to  $\mathcal{F}$ , and denoted by

$$(9.13) \quad g = \mathbb{E}(f|\mathcal{F}).$$

As an element of  $L^1(X, \mathcal{B}, \mu)$ ,  $\mathbb{E}(f|\mathcal{F})$  is characterized by the following two properties:

$$(9.14) \quad \mathbb{E}(f|\mathcal{F}) \quad \text{is } \mathcal{F}\text{-measurable;}$$

$$(9.15) \quad \int_F \mathbb{E}(f|\mathcal{F}) d\mu = \int_F f d\mu \quad \text{for all } F \in \mathcal{F}.$$

We think of  $\mathbb{E}(f|\mathcal{F})(x)$  as our expected value for  $f$  if we are *given the information in  $\mathcal{F}$* , in the sense that for each  $F \in \mathcal{F}$  we know whether or not  $x \in F$ . When  $\mathcal{F}$  is the  $\sigma$ -algebra generated by a finite measurable partition  $\alpha$  of  $X$  and  $f$  is the characteristic function of a set  $A \in \mathcal{B}$ , the conditional expectation gives the conditional probabilities of  $A$  with respect to all the sets in  $\alpha$ :

$$(9.16) \quad \mathbb{E}(\mathbf{1}_A|\mathcal{F})(x) = \mu(A|\alpha(x)) = \mu(A \cap F)/\mu(F) \quad \text{if } x \in F \in \alpha.$$

We write  $\mathbb{E}(f) = \mathbb{E}(f|\{\emptyset, X\}) = \int_X f d\mu$  for the expectation of any integrable function  $f$ . A measurable function  $f$  on  $X$  is *independent* of a sub- $\sigma$ -algebra  $\mathcal{F} \subset \mathcal{B}$  if for each  $(a, b) \subset \mathbb{R}$  and  $F \in \mathcal{F}$  we have

$$(9.17) \quad \mu(f^{-1}(a, b) \cap F) = \mu(f^{-1}(a, b)) \mu(F).$$

A function  $\tau : \mathbb{R} \rightarrow \mathbb{R}$  is *convex* if whenever  $t_1, \dots, t_n \geq 0$  and  $\sum_{i=1}^n t_i = 1$ ,

$$(9.18) \quad \phi\left(\sum_{i=1}^n t_i x_i\right) \leq \sum_{i=1}^n t_i \phi(x_i) \quad \text{for all } x_1, \dots, x_n \in \mathbb{R}.$$

**Theorem 9.12.** *Let  $(X, \mathcal{B}, \mu)$  be a probability space and  $\mathcal{F} \subset \mathcal{B}$  a sub- $\sigma$ -algebra.*

- (1)  $\mathbb{E}(\cdot|\mathcal{F})$  is a positive contraction on  $L^p(X)$  for each  $p \geq 1$ .
- (2) If  $f \in L^1(X)$  is  $\mathcal{F}$ -measurable, then  $\mathbb{E}(f|\mathcal{F}) = f$  a.e.. If  $f \in L^\infty(X)$  is  $\mathcal{F}$ -measurable, then  $\mathbb{E}(fg|\mathcal{F}) = f\mathbb{E}(g|\mathcal{F})$  for all  $g \in L^1(X)$ .
- (3) If  $\mathcal{F}_1 \subset \mathcal{F}_2$  are sub- $\sigma$ -algebras of  $\mathcal{B}$ , then  $\mathbb{E}(\mathbb{E}(f|\mathcal{F}_2)|\mathcal{F}_1) = \mathbb{E}(f|\mathcal{F}_1)$  a.e. for each  $f \in L^1(X)$ .
- (4) If  $f \in L^1(X)$  is independent of the sub- $\sigma$ -algebra  $\mathcal{F} \subset \mathcal{B}$ , then  $\mathbb{E}(f|\mathcal{F}) = \mathbb{E}(f)$  a.e..
- (5) If  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  is convex,  $f$  and  $\phi \circ f \in L^1(X)$ , and  $\mathcal{F} \subset \mathcal{B}$  is a sub- $\sigma$ -algebra, then  $\phi(\mathbb{E}(f|\mathcal{F})) \leq \mathbb{E}(\phi \circ f|\mathcal{F})$  a.e..

**9.3. The Spectral Theorem.** A *separable Hilbert space* is one with a countable dense set, equivalently a countable orthonormal basis. A *normal operator* is a continuous linear operator  $S$  on a Hilbert space  $\mathcal{H}$  such that  $SS^* = S^*S$ ,  $S^*$  being the adjoint operator defined by  $(Sf, g) = (f, S^*g)$  for all  $f, g \in \mathcal{H}$ . A continuous linear operator  $S$  is *unitary* if it is invertible and  $S^* = S^{-1}$ . Two operators  $S_1$  and  $S_2$  on Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , respectively, are called *unitarily equivalent* if there is a unitary operator  $U : \mathcal{H}_1 \rightarrow \mathcal{H}_2$  which carries  $S_1$  to  $S_2$ :  $S_2U = US_1$ . The following brief account follows [10, 30].

**Theorem 9.13.** *Let  $S : \mathcal{H} \rightarrow \mathcal{H}$  be a normal operator on a separable Hilbert space  $\mathcal{H}$ . Then there are mutually singular Borel probability measures  $\mu_\infty, \mu_1, \mu_2, \dots$  such that  $S$  is unitarily equivalent to the operator  $M$  on the direct sum Hilbert space*

$$(9.19) \quad L^2(\mathbb{C}, \mu_\infty) \oplus L^2(\mathbb{C}, \mu_1) \oplus \left( \bigoplus_{k=1}^2 L^2(\mathbb{C}, \mu_2) \right) \oplus \dots \oplus \left( \bigoplus_{k=1}^m L^2(\mathbb{C}, \mu_m) \right) \oplus \dots$$

defined by

$$(9.20) \quad M((f_{\infty,1}(z_{\infty,1}), f_{\infty,2}(z_{\infty,2}), \dots), (f_{1,1}(z_{1,1})), (f_{2,1}(z_{2,1}), f_{2,2}(z_{2,2})), \dots) = \\ (z_{\infty,1}f_{\infty,1}(z_{\infty,1}), z_{\infty,2}f_{\infty,2}(z_{\infty,2}), \dots), (z_{1,1}f_{1,1}(z_{1,1})), \\ (z_{2,1}f_{2,1}(z_{2,1}), z_{2,2}f_{2,2}(z_{2,2})), \dots).$$

The measures  $\mu_i$  are supported on the *spectrum*  $\sigma(S)$  of  $S$ , the (compact) set of all  $\lambda \in \mathbb{C}$  such that  $S - \lambda I$  does not have a continuous inverse. Some of the  $\mu_i$  may be 0. They are uniquely determined up to absolute continuity equivalence. The smallest absolute continuity class with respect to which all the  $\mu_i$  are absolutely continuous is called the *maximum spectral type* of  $S$ . A measure representing this type is  $\sum_i \mu_i/2^i$ . We have in mind the example for which  $\mathcal{H} = L^2(X, \mathcal{B}, \mu)$  and  $Sf = f \circ T$  (the ‘‘Koopman operator’’) for a measure-preserving system  $(X, \mathcal{B}, \mu, T)$  on a Lebesgue space  $(X, \mathcal{B}, \mu)$ , which is *unitary*: it is linear, continuous, invertible, preserves scalar products, and has spectrum equal to the unit circle.

The proof of Theorem 9.13 can be accomplished by first decomposing  $\mathcal{H}$  (in a careful way) into the direct sum of pairwise orthogonal *cyclic subspaces*  $\mathcal{H}_n$ : each  $\mathcal{H}_n$  is the closed linear span of  $\{S^i(S^*)^j f_n : i, j \geq 0\}$  for some  $f_n \in \mathcal{H}$ . This means that for each  $n$  the set  $\{p(S, S^*)f_n : p \text{ is a polynomial in two variables}\}$  is dense in  $\mathcal{H}_n$ . Similarly, by the Stone-Weierstrass Theorem the set  $\mathcal{P}_n$  of all polynomials  $p(z, \bar{z})$  is dense in the set  $\mathcal{C}(\sigma(S|\mathcal{H}_n))$  of continuous complex-valued functions on  $\sigma(S|\mathcal{H}_n)$ . We define a bounded linear functional  $\phi_n$  on  $\mathcal{P}_n$  by

$$(9.21) \quad \phi(p) = (p(S, S^*)f_n, f_n)$$

and extend it by continuity to a bounded linear functional on  $\mathcal{C}(\sigma(S|\mathcal{H}_n))$ . It can be proved that this functional is positive, and therefore, by the Riesz Representation Theorem, it corresponds to a positive Borel measure on  $\sigma(S|\mathcal{H}_n)$ .

The various  $L^2$  spaces and multiplication operators involved in the above theorem can be amalgamated into a coherent whole, resulting in the following convenient form of the Spectral Theorem for normal operators

**Theorem 9.14.** *Let  $S : \mathcal{H} \rightarrow \mathcal{H}$  be a normal operator on a separable Hilbert space  $\mathcal{H}$ . There are a finite measure space  $(X, \mathcal{B}, \mu)$  and a bounded measurable function  $h : X \rightarrow \mathbb{C}$  such that  $S$  is unitarily equivalent to the operator of multiplication by  $h$  on  $L^2(X, \mathcal{B}, \mu)$ .*

The form of the Spectral Theorem given in Theorem 9.13 is useful for discussing absolute continuity and multiplicity properties of the spectrum of a normal operator. Another form, involving spectral measures, has useful consequences such as the functional calculus.

**Theorem 9.15.** *Let  $S : \mathcal{H} \rightarrow \mathcal{H}$  be a normal operator on a separable Hilbert space  $\mathcal{H}$ . There is a unique projection-valued measure  $E$  defined on the Borel subsets of the spectrum  $\sigma(S)$  of  $S$  such that  $E(\sigma(S)) = I$  (= the identity on  $\mathcal{H}$ );*

$$(9.22) \quad (E(\bigcup_{i=1}^{\infty} A_i))f = \sum_{i=1}^{\infty} (E(A_i))f$$

whenever  $A_1, A_2, \dots$  are pairwise disjoint Borel subsets of  $\sigma(S)$  and  $f \in \mathcal{H}$ , with the series converging in norm; and

$$(9.23) \quad S = \int_{\sigma(S)} \lambda dE(\lambda).$$

Spectral integrals such as the one in (9.23) can be defined by reducing to complex measures  $\mu_{f,g}(A) = (E(A)f, g)$ , for  $f, g \in \mathcal{H}$  and  $A \subset \sigma(S)$  a Borel set. Given a bounded Borel measurable function  $\phi$  on  $\sigma(S)$ , the operator

$$(9.24) \quad V = \phi(S) = \int_{\sigma(S)} \phi(\lambda) dE(\lambda)$$

is determined by specifying that

$$(9.25) \quad (Vf, g) = \int_{\sigma(S)} \phi(\lambda) d\mu_{f,g} \quad \text{for all } f, g \in \mathcal{H}.$$

Then

$$(9.26) \quad S^k = \int_{\sigma(S)} \lambda^k dE(\lambda) \quad \text{for all } k = 0, 1, \dots$$

These spectral integrals sometimes behave a bit strangely:

$$(9.27) \quad \begin{aligned} \text{If } V_1 = \int_{\sigma(S)} \phi_1(\lambda) dE(\lambda) \quad \text{and } V_2 = \int_{\sigma(S)} \phi_2(\lambda) dE(\lambda), \text{ then} \\ V_1 V_2 = \int_{\sigma(S)} \phi_1(\lambda) \phi_2(\lambda) dE(\lambda). \end{aligned}$$

Finally, if  $f \in \mathcal{H}$  and  $\nu$  is a finite positive Borel measure that is absolutely continuous with respect to  $\mu_{f,f}$ , then there is  $g$  in the closed linear span of  $\{S^i(S^*)^j f : i, j \geq 0\}$  such that  $\nu = \mu_{g,g}$ .

Theorem 9.15 can be proved by applying Theorem 9.14, which allows us to assume that  $\mathcal{H} = L^2(X, \mathcal{B}, \mu)$  and  $S$  is multiplication by  $h \in L^\infty(X, \mathcal{B}, \mu)$ . For any Borel set  $A \subset \sigma(S)$ , let  $E(A)$  be the projection operator given by multiplication by the 0, 1-valued function  $\mathbf{1}_A \circ h$ .

## 10. FUTURE DIRECTIONS

The mathematical study of dynamical systems arose in the late nineteenth and early twentieth century, along with measure theory and probability theory, so it is a young field with many interesting open problems. New questions arise continually from applications and from interactions with other parts of mathematics. Basic aspects of the problems of classification, topological or smooth realization, and systematic construction of measure-preserving systems remain open. There is much work to be done to understand the relations among systems and different types of systems (factors and relative properties, joinings and disjointness, various notions

of equivalence with associated invariants). There is a continual need to determine properties of classes of systems and of particular systems arising from applications or other parts of mathematics such as probability, number theory, geometry, algebra, and harmonic analysis. Some of these questions are mentioned in more detail in the other articles in this collection.

## 11. BIBLIOGRAPHY

### PRIMARY LITERATURE

- [1] Roy L. Adler, *F-expansions revisited* (1973), 1–5. Lecture Notes in Math., Vol. 318. MR 0389825 (52 #10655)
- [2] Leslie K. Arnold, *On  $\sigma$ -finite invariant measures*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **9** (1968), 85–97. MR 0227362 (37 #2947)
- [3] Patrick Billingsley, *Ergodic Theory and Information*, Robert E. Krieger Publishing Co., Huntington, N.Y., 1978, ISBN 0-88275-666-4, Reprint of the 1965 original. MR 524567 (80b:28017)
- [4] ———, *Probability and Measure*, 3rd ed., Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, 1995, ISBN 0-471-00710-2, A Wiley-Interscience Publication. MR 1324786 (95k:60001)
- [5] Antoine Brunel, *Sur les mesures invariantes*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **5** (1966), 300–303. MR 0207957 (34 #7770) (French)
- [6] Alberto P. Calderón, *Sur les mesures invariantes*, C. R. Acad. Sci. Paris **240** (1955), 1960–1962. MR 0069257 (16,1008b) (French)
- [7] Constantin Carathéodory, *Die Homomorphieen von Somen und die Multiplikation von Inhaltfunktionen*, Annali della R. Scuola Normale Superiore di Pisa (Ser. 2) **8** (1939), 105–130.
- [8] ———, *Vorlesungen über Reelle Funktionen*, Third (corrected) edition, Chelsea Publishing Co., New York, 1968. MR 0225940 (37 #1530) (German)
- [9] R. V. Chacon, *A class of linear transformations*, Proc. Amer. Math. Soc. **15** (1964), 560–564. MR 0165071 (29 #2362)
- [10] John B. Conway, *A Course in Functional Analysis*, 2nd ed., Graduate Texts in Mathematics, vol. 96, Springer-Verlag, New York, 1990, ISBN 0-387-97245-5. MR 1070713 (91e:46001)
- [11] Yael Naim Dowker, *On measurable transformations in finite measure spaces*, Ann. of Math. (2) **62** (1955), 504–516. MR 0072929 (17,353a)
- [12] ———, *Sur les applications mesurables*, C. R. Acad. Sci. Paris **242** (1956), 329–331. MR 0074486 (17,593i) (French)
- [13] Maurice Fréchet, *Des familles et fonctions additives d'ensembles abstraits*, Fund. Math. **5** (1919/24), 206–251. (French)
- [14] Nathaniel A. Friedman, *Introduction to Ergodic Theory*, Van Nostrand Reinhold Co., New York, 1970, Van Nostrand Reinhold Mathematical Studies, No. 29. MR 0435350 (55 #8310)
- [15] Harry Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204–256. MR 0498471 (58 #16583)
- [16] W. T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), 465–588. MR 1844079 (2002k:11014)
- [17] *Erratum: "A new proof of Szemerédi's theorem" [Geom. Funct. Anal. **11** (2001), no. 3, 465–588; 1844079] by W. T. Gowers*, Geom. Funct. Anal. **11** (2001), 869. MR 1866805
- [18] Ben Green and Terence Tao, *The primes contain arbitrarily long arithmetic progressions*, arXiv:math.NT/0404188.
- [19] Hans Hahn, *Über die Multiplikation total-additiver Mengenfunktionen*, Annali Scuola Norm. Sup. Pisa **2** (1933), 429–452. (German)
- [20] Arshag B. Hajian and Shizuo Kakutani, *Weakly wandering sets and invariant measures*, Trans. Amer. Math. Soc. **110** (1964), 136–151. MR 0154961 (27 #4904)
- [21] Paul R. Halmos, *Invariant measures*, Ann. of Math. (2) **48** (1947), 735–754. MR 0021963 (9,137d)

- [22] Anatole Katok and Boris Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Encyclopedia of Mathematics and its Applications, vol. 54, Cambridge University Press, Cambridge, 1995, ISBN 0-521-34187-6, With a supplementary chapter by Katok and Leonardo Mendoza. MR 1326374 (96c:58055)
- [23] Eberhard Hopf, *Theory of measure and invariant integrals*, Trans. Amer. Math. Soc. **34** (1932), 373–393. MR 1501643
- [24] ———, *Ergodentheorie*, 1st ed., Ergebnisse der Mathematik und ihrer Grenzgebiete; 5. Bd., 2, J. Springer, Berlin, 1937.
- [25] A. I. Khinchin, *Mathematical Foundations of Statistical Mechanics*, Dover Publications Inc., New York, N. Y., 1949, Translated by G. Gamow. MR 0029808 (10,666c)
- [26] A. N. Kolmogorov, *Foundations of the Theory of Probability*, Chelsea Publishing Co., New York, 1956, Translation edited by Nathan Morrison, with an added bibliography by A. T. Bharucha-Reid. MR 0079843 (18,155e)
- [27] Donald S. Ornstein, *On invariant measures*, Bull. Amer. Math. Soc. **66** (1960), 297–300. MR 0146350 (26 #3872)
- [28] Karl Petersen, *Ergodic Theory*, Cambridge Studies in Advanced Mathematics, vol. 2, Cambridge University Press, Cambridge, 1989, ISBN 0-521-38997-6, Corrected reprint of the 1983 original. MR 1073173 (92c:28010)
- [29] H. Poincaré, *Les Méthodes Nouvelles de la Mécanique Céleste. Tomes I, II, III*, Les Grands Classiques Gauthier-Villars. [Gauthier-Villars Great Classics], Librairie Scientifique et Technique Albert Blanchard, Paris, 1987, ISBN 2-85367-093-7, Reprint of the 1892 original; With a foreword by J. Kovalevsky; Bibliothèque Scientifique Albert Blanchard. [Albert Blanchard Scientific Library]. MR 926906 (89e:01054a) (French)
- [30] Heydar Radjavi and Peter Rosenthal, *Invariant subspaces*, 2nd ed., Dover Publications Inc., Mineola, NY, 2003, ISBN 0-486-42822-2. MR 2003221 (2004e:47010)
- [31] V. A. Rohlin, *On the fundamental ideas of measure theory*, Amer. Math. Soc. Translation **1952** (1952), 55. MR 0047744 (13,924e)
- [32] ———, *New progress in the theory of transformations with invariant measure.*, Russian Math. Surveys **15** (1960), 1–22. MR 0132155 (24 #A2002)
- [33] H. L. Royden, *Real Analysis*, 3rd ed., Macmillan Publishing Company, New York, 1988, ISBN 0-02-404151-3. MR 1013117 (90g:00004)
- [34] E. Szemerédi, *On sets of integers containing no  $k$  elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245, Collection of articles in memory of Juriĭ Vladimirovič Linnik. MR 0369312 (51 #5547)
- [35] Terence Tao, *Szemerédi’s regularity lemma revisited*, Contrib. Discrete Math. **1** (2006), 8–28 (electronic). MR 2212136
- [36] ———, *Arithmetic progressions and the primes*, Collect. Math. (2006), 37–88. MR 2264205
- [37] J. von Neumann, *Einige Sätze über messbare Abbildungen*, Ann. of Math. (2) **33** (1932), 574–586. MR 1503077 (German)
- [38] Fred B. Wright, *The recurrence theorem*, Amer. Math. Monthly **68** (1961), 247–248. MR 0133427 (24 #A3261)
- [39] Lai-Sang Young, *What are SRB measures, and which dynamical systems have them?*, J. Statist. Phys. **108** (2002), 733–754, Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays. MR 1933431 (2003g:37042)

#### BOOKS AND REVIEWS

- [1] Patrick Billingsley, *Ergodic Theory and Information*, Robert E. Krieger Publishing Co., Huntington, N. Y., 1978, ISBN 0-88275-666-4, Reprint of the 1965 original. MR 524567 (80b:28017)
- [2] I. P. Cornfeld, S. V. Fomin, and Ya. G. Sinai, *Ergodic Theory*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 245, Springer-Verlag, New York, 1982, ISBN 0-387-90580-4, Translated from the Russian by A. B. Sosinskiĭ. MR 832433 (87f:28019)
- [3] Manfred Denker, Christian Grillenberger, and Karl Sigmund, *Ergodic Theory on Compact Spaces*, Springer-Verlag, Berlin, 1976, Lecture Notes in Mathematics, Vol. 527. MR 0457675 (56 #15879)

- [4] Nathaniel A. Friedman, *Introduction to Ergodic Theory*, Van Nostrand Reinhold Co., New York, 1970, Van Nostrand Reinhold Mathematical Studies, No. 29. MR 0435350 (55 #8310)
- [5] Eli Glasner, *Ergodic Theory via Joinings*, Mathematical Surveys and Monographs, vol. 101, American Mathematical Society, Providence, RI, 2003, ISBN 0-8218-3372-3. MR 1958753 (2004c:37011)
- [6] Paul R. Halmos, *Lectures on Ergodic Theory*, Chelsea Publishing Co., New York, 1960. MR 0111817 (22 #2677)
- [7] Anatole Katok and Boris Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Encyclopedia of Mathematics and its Applications, vol. 54, Cambridge University Press, Cambridge, 1995, ISBN 0-521-34187-6, With a supplementary chapter by Katok and Leonardo Mendoza. MR 1326374 (96c:58055)
- [8] Eberhard Hopf, *Ergodentheorie*, 1st ed., Ergebnisse der Mathematik und ihrer Grenzgebiete ; 5. Bd., 2, J. Springer, Berlin, 1937.
- [9] Konrad Jacobs, *Einige neuere Ergebnisse der Ergodentheorie*, Jber. Deutsch. Math.-Verein. **67** (1965), 143–182. MR 0186789 (32 #4244) (German)
- [10] Karl Petersen, *Ergodic Theory*, Cambridge Studies in Advanced Mathematics, vol. 2, Cambridge University Press, Cambridge, 1989, ISBN 0-521-38997-6, Corrected reprint of the 1983 original. MR 1073173 (92c:28010)
- [11] H. L. Royden, *Real Analysis*, 3rd ed., Macmillan Publishing Company, New York, 1988, ISBN 0-02-404151-3. MR 1013117 (90g:00004)
- [12] Daniel J. Rudolph, *Fundamentals of Measurable Dynamics*, Oxford Science Publications, The Clarendon Press Oxford University Press, New York, 1990, ISBN 0-19-853572-4, Ergodic theory on Lebesgue spaces. MR 1086631 (92e:28006)
- [13] Peter Walters, *An Introduction to Ergodic Theory*, Graduate Texts in Mathematics, vol. 79, Springer-Verlag, New York, 1982, ISBN 0-387-90599-5. MR 648108 (84e:28017)

DEPARTMENT OF MATHEMATICS, CB 3250, PHILLIPS HALL, UNIVERSITY OF NORTH CAROLINA,  
CHAPEL HILL, NC 27599 USA

*E-mail address:* `petersen@math.unc.edu`