BOOK REVIEWS

BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 14, Number 2, April 1986

0273-0979/86 $1.00 + $.25 per page

Recurrence in ergodic theory and combinatorial number theory, by H. Furstenberg, Princeton Univ. Press, 1981, vii + 199 pp., $25.00. ISBN 0-691-08269-3

The mathematical study of dynamical systems, comprising ergodic theory, topological dynamics, and differentiable dynamics (depending on whether the setting is a measure space, a topological space, or a differentiable manifold), arose from theoretical physics, especially Hamiltonian mechanics and statistical mechanics, and has drawn heavily from number theory, harmonic analysis, and differential geometry for both techniques and examples. Here are three important examples of applications of combinatorics or number theory to dynamical systems.

(1) Hadamard and Morse developed symbolic dynamics as a systematic way of making abstract dynamical systems susceptible to combinatorial analysis. If X is the phase space, or set of possible states, of a dynamical system, and $T: X \to X$ is a transformation which "makes time go by", so that $T^n x$ is the state at time n of a system which at time 0 is in state x, and if $\{X_0, X_1\}$ is a carefully chosen partition of X into two disjoint sets, then the doubly infinite $0,1$ sequence $\omega(x)$ for which $\omega(x)_k = i$ if and only if $T^k x \in X_i$ will capture much of the information in the entire trajectory (or orbit, or history) $\{T^n x: n \in \mathbf{Z}\}$ of x. In this way, one can use combinatorial properties of the result of the "coding" $x \to \omega(x)$ to study dynamically interesting properties of the system, such as the existence of periodic or recurrent orbits, the presence or absence of metric or topological transitivity or other mixing properties, and the entropy.

(2) The theorem of Kronecker that if α is irrational then $\{n\alpha: n \in \mathbf{Z}\}$ is dense mod 1 says, in dynamical terms, that an irrational rotation R_α of the circle is *minimal*: there are no proper closed invariant sets. Weyl's equidistribution theorem which states that $\{n\alpha: n \in \mathbf{Z}\}$ is equidistributed mod 1 (for each subinterval I of the circle, $\operatorname{card}\{k: 1 \leqslant k \leqslant n, (k\alpha \bmod 1) \in I\}/n$ converges to the length of I) implies that an irrational rotation of the circle is *uniquely ergodic*: there is only one R_α-invariant Borel probability measure on the circle (Lebesgue measure). Extensions and variations of these theorems have provided many nice examples in dynamical systems theory.

(3) In the "small denominators" problems of celestial mechanics, the existence of a solution of an equation connected with a dynamical system

depends on the speed with which a parameter of the system can be approximated by rational numbers. The implications of this finding for the stability or instability of the system have been the subject of profound investigations by Poincaré, Siegel, Kolmogorov, Arnold, and Moser, among others. It may not be possible to convince everyone that all the results obtained in this way have actual physical relevance, but the general theory of dynamical systems developed in this century with the help of number theory, harmonic analysis, and geometry has certainly illuminated several aspects of theoretical physics.

Dynamical systems theory has also begun to repay its debts to its supporting subjects—by producing new proofs, from a dynamical viewpoint, of known results, sometimes thus providing new unity to disparate results or new clarity to statements previously accompanied perhaps only by computational or obscure proofs; by producing new theorems in these subjects; and by leading to new questions and even new classes of questions in these older parts of mathematics. Thus the interplay of dynamical systems with number theory, geometry, and harmonic analysis continues to deepen and to enrich itself.

Here are three examples of this phenomenon (others can be found in [**9**]).

(1) Another equidistribution theorem of Weyl says that if p is a polynomial with real coefficients, at least one of which is irrational, then $\{p(n): n \in \mathbf{Z}\}$ is equidistributed mod 1. Dynamical proofs of this [**1**–**4**, **6**–**8**] turn the tables by showing first that a (skew product) dynamical system built by looking at the polynomial p is uniquely ergodic and thus easily reading off the result. (The present book proves the denseness of $\{p(n): n \in \mathbf{Z}\}$ by the same method. This is an easier argument which nicely illustrates the main ideas.)

(2) The class of almost periodic functions and its extension, introduced by Bochner, of the class of almost automorphic functions, can be studied naturally in a dynamic context. If f is a bounded function on a group, say on the integers, then we can consider f as a point in $K^{\mathbf{Z}}$, where K is a compact subset of $\mathbf{C}$ which contains the range of f. Translation in $\mathbf{Z}$ produces the shift operator $\sigma: K^{\mathbf{Z}} \to K^{\mathbf{Z}}$, for which $(\sigma f)(n) = f(n+1)$. Recurrence properties of the point f in the *Bebutov dynamical system* $(K^{\mathbf{Z}}, \sigma)$ are frequently related to properties of f which are of interest to harmonic analysts. Work of Veech, Knapp, and Ellis, among others, has developed this connection in the cases of almost automorphic, distal, and point distal functions.

(3) A famous theorem (the solution of a conjecture popularized by the Dutch mathematician Baudet in Göttingen in the 1920s) of van der Waerden, Artin, and Schreier says that if the natural numbers are divided into finitely many sets—$\mathbf{N} = S_1 \cup S_1 \cup \cdots \cup S_r$—then at least one of these sets contains arbitrarily long arithmetic progressions—there is a $j = 1, 2, \ldots, r$ such that given any $l = 1, 2, \ldots$ we can find $s, t \in \mathbf{N}$ such that $s, s+t, s+2t, \ldots, s+lt \in S_j$. An extension of this result was conjectured by Erdös and Turán, and Erdös of course also offered a monetary reward ($1000—he knew it was a hard problem) for its solution: Any subset of $\mathbf{N}$ which has positive upper density contains arbitrarily long arithmetic progressions. After Roth showed that one could always find arithmetic progressions of length 3 in any subset of $\mathbf{N}$ with positive upper density, Szemerédi gave the (difficult and intricate) proof of the full conjecture and claimed the reward. Furstenberg, inspired by a lecture in

which K. Jacobs described Szemerédi's new theorem, saw how to prove these results on the basis of topological dynamics and ergodic theory. Together with collaborators such as Katznelson, Weiss, and Ornstein, he has continued to work out the details, ramifications, improvements, and extensions. The present book is an exposition of these results at one stage in their (still continuing) development.

It is not surprising that the young subject of dynamical systems should call upon established areas of mathematics like number theory, harmonic analysis, and geometry for support when it needs techniques and examples, but how is it possible that an apparently abstract and imprecise subject like ergodic theory (dealing in qualitative and almost everywhere statements) could have anything important or informative to say about that crystalline kernel of the mass of mathematics, combinatorial number theory? I can give a vague and a precise indication of an answer. For the vague indication, please excuse some momentary philosophy. The idea of recurrence is the hinge connecting ergodic theory and combinatorial number theory: in the one, it determines stabilities and rhythms, while in the other it is the ultimate basis of repeated patterns. Indeed, there cannot be patterns without some sort of recurrence, and the task of all branches of mathematics is the search for patterns, just as for all of science it is the search for recurrent, hence to some degree predictable, phenomena. It is no accident that the parts of number theory and harmonic analysis which rely the most on recurrence phenomena are the ones which are the ripest for dynamical applications and investigation from the dynamical viewpoint.

For the precise indication, given $S \subset \mathbf{Z}$ with positive upper density (in fact, what Furstenberg calls positive *upper Banach density*,

$$\limsup_{\substack{\operatorname{card}(I)\to\infty \\ I \text{ an interval}}} \frac{\operatorname{card}(S \cap I)}{\operatorname{card}(I)} > 0,$$

is good enough), let $X \subset \{0,1\}^{\mathbf{Z}}$ be the orbit closure under the shift σ of the characteristic function χ_S of S (another Bebutov system), and let $A = \{x \in X: x(0) = 1\}$. Then (X, σ) is a dynamical system whose properties are related to the combinatorial properties of the point $\chi_S \in X$—especially when recurrence is involved. We want to find arithmetic progressions of arbitrary length, l, in S. What does it mean to say that $s, s+t, s+2t, \ldots, s+lt \in S$? Of course, just that $\chi_S(s) = \chi_S(s+t) = \chi_S(s+2t) = \cdots = \chi_S(s+lt) = 1$. This would follow immediately if we knew that $A \cap T^{-t}A \cap T^{-2t}A \cap \cdots \cap T^{-lt}A \neq \varnothing$: for if $\omega \in A \cap T^{-t}A \cap T^{-2t}A \cap \cdots \cap T^{-lt}A$, then $\omega(0) = \omega(t) = \omega(2t) = \cdots = \omega(lt) = 1$. Since $\omega \in X =$ closure of orbit of χ_S under σ, there is an s such that the initial $(lt+1)$-block of ω agrees with the initial $(lt+1)$-block of $\sigma^s\chi_S$. Then $\chi_S(s) = \chi_S(s+t) = \chi_S(s+2t) = \cdots = \chi_S(s+lt) = 1$. To prove that given l one can find t with $A \cap T^{-t}A \cap T^{-2t}A \cap \cdots \cap T^{-lt}A \neq \varnothing$, one makes the problem more general (therefore harder) and more abstract (therefore easier). The proof will be complete if one can show (1) there is a σ-invariant measure μ on $\{0,1\}^{\mathbf{Z}}$ for which the set A above has $\mu(A) > 0$, and (2) whenever $T_1, \ldots, T_r$ are commuting μ-preserving transformations on a measure space $(X, \mathscr{B}, \mu)$ and $A \in \mathscr{B}$ with $\mu(A) > 0$, then

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(T_1^{-n}A \cap T_2^{-n}A \cap \cdots \cap T_r^{-n}A) > 0.$$

The proof of (1) is fairly easy. For the van der Waerden–Artin–Schreier Theorem, measure theory is not involved, and the result follows from the above connection and an easily accessible topological multiple recurrence theorem. The proof of (2) is hard and takes up Chapters 4–7 of the present nine-chapter book. This theorem (2) is an amazing extension of the famous Poincaré Recurrence Theorem, according to which, if T: $X \to X$ is a measure-preserving transformation on a measure space $(X, \mathscr{B}, \mu)$ with $\mu(X) < \infty$, and if $A \in \mathscr{B}$ with $\mu(A) > 0$, then there is an $n > 1$ for which $\mu(T^{-n}A \cap A) > 0$. This Multiple Recurrence Theorem is not too hard to prove for a weakly mixing system, where it follows from average independence of the sets $T_i^{-n}A$, or for an equicontinuous system, where it follows from uniform recurrence. In order to prove his Multiple Recurrence Theorem for all measure-preserving $\mathbf{Z}^r$ actions, Furstenberg then needs to (1) develop relative ergodic theory: notions of ergodicity, weak mixing, almost periodic, mean ergodic theorem, etc., for a factor map or extension (measurable measure-preserving map $(X, \mathscr{B}, \mu, T) \to (Y, \mathscr{C}, \nu, S)$ which commutes with T and S); (2) show that the multiple recurrence property lifts under weakly mixing and compact extensions, indeed under controlled combinations of the two called primitive extensions, and that it is preserved by inverse limits; (3) prove a structure theorem asserting that every system is an inverse limit of primitive extensions, beginning with the one-point system. A key ingredient in (3) is the proof that given a nontrivial factor map $X \to Z$, one can interpolate $X \to Y \to Z$ with $Y \to Z$ primitive and nontrivial; then Zorn's Lemma helps to finish (3). All of this involves a virtuoso display of intertwined ergodic theory, topological dynamics, measure theory, and functional analysis.

The structure theory with the techniques and detailed arguments used to develop them in Chapters 4–7 may be useful to ergodic theorists, but the casual reader will be happy to know that Furstenberg, along with Katznelson and Ornstein, has found a shorter path to the Multiple Recurrence Theorem (although the main ideas are still the same)—see **[5]**. For this reviewer, the main value of the present book is in the way it sets the topological and measure-theoretic multiple recurrence theorems, with their corollaries the van der Waerden–Artin–Schreier and Szemerédi Theorems, into the larger context of topological dynamics–ergodic theory and combinatorial number theory–diophantine approximation, and the way in which it expounds the author's fruitful viewpoint from which the connections among these subjects are clear. Thus I would read **[5]** for the detailed argument, but this book for examples, motivations, related questions and overall understanding.

The Poincaré Recurrence Theorem is very easy to prove; and it is not too hard, though a bit tricky, to show that if $\mu(A) > 0$ then $W = \{n \geqslant 1$: $\mu(T^{-n}A \cap A) > 0\}$ is *syndetic*, or *relatively dense*: there is a K such that every subinterval of $\mathbf{N}$ of length K contains an element of W. (See p. 74 for a neat proof if you get stuck.) But how could one possibly prove that if $\mu(A) > 0$

then there is always an n such that $\mu(T^{-n^2}A \cap A) > 0$? The proof is much easier than several in this book (see p. 72), but the statement is indicative of the manifold novelties of the author's approach. Again, it is easy to prove (Dirichlet—pigeonhole principle) that given $\alpha \in \mathbf{R}$ and $\varepsilon > 0$, there are integers m and n with $|\alpha - m/n| < \varepsilon/n$. But how can you show that given $\varepsilon > 0$ you can always find m and n with $|\alpha - m/n^2| < \varepsilon/n^2$? This was originally proved by Hardy and Littlewood; p. 22 gives a proof based on skew-product dynamical systems, and p. 48 by a multidimensional version of the van der Waerden–Artin–Schreier Theorem. There is much more in this book, including variations of the results mentioned already involving *IP* sets, sets of recurrence times, difference sets, and sets containing arbitrarily long intervals. Some great names, like Hilbert, Schur, and Radó, have been associated with these problems in the past. There are insights on every page, helpful and illuminating examples—both easy and significant—abound, and provocative ideas leap out at the reader, inducing experimentation and fueling speculation. The author's extremely productive viewpoint leads to a multidimensional array of new types of questions in dynamical systems and number theory; the theory of mild mixing begun in the last chapter and the higher-dimensional Szemerédi Theorem (a combinatorial number theory result found by these dynamical techniques) are examples for the ambitious reader to emulate.

With all this inspiration, perhaps a list of precisely formulated open questions would have been superfluous or too quickly obsolete, but I still think it would have been useful. I could also quibble about the references, which are not always given, and, when present, are not always to the primary source. But no one should come to this book for his or her first contact with dynamical systems or combinatorial number theory; so, given the magnitude of the achievement in first discovering these results and then presenting them in such a clear and complete manner, the quibbles are negligible.

References

1. N. M. Akulsničev, *On a dynamical system related to the distribution of the fractional parts of second-degree polynomials*, Dokl. Akad. Nauk SSR **143** (1962), 503–505; English transl. Soviet Math. Dokl. **3** (1962), 405–407.

2. J. Cigler, *On a theorem of H. Weyl*, Compositio Math. **21** (1969), 151–154.

3. H. Furstenberg, *Stationary processes and prediction theory*, Ann. of Math. Studies No. 44, Princeton Univ. Press, Princeton, N.J., 1960.

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

5. H. Furstenberg, Y. Katznelson and D. Ornstein, *The ergodic theoretical proof of Szemerédi's theorem*, Bull. Amer. Math. Soc. **7** (1982), 527–552.

6. F. J. Hahn, *On affine transformations of compact abelian groups*, Amer. J. Math. **85** (1963), 428–446. *Erratum* **86** (1964).

7. ______, *Skew product transformations and the algebras generated by* $\exp(p(n))$, Illinois J. Math. **9** (1965), 178–190.

8. A. G. Postnikov, *Ergodic problems in the theory of congruences and of diophantine approximations*, Trudy Mat. Inst. Steklov **82** (1966); English transl. Proc. Steklov Inst. Math. **82** (1967).

9. William A. Veech, *Topological dynamics*, Bull. Amer. Math. Soc. **83** (1977), 775–830.

Karl Petersen