Ergodic Theory and Applications to Additive Number Theory

A. M. Blumenthal

January 27, 2010

1 Dynamical Systems and Symbolic Flow

Definition A *measure preserving system* (an m.p.s) is the quadruple (X, β, μ, T), where X is a topological space, β is a σ algebra of subsets of X, μ is a measure on β and $T : X \to X$ is a continuous map with the property that if $A \in \beta$, then $\mu(TA) = \mu(A)$.

Definition A *measure* on X is a map $\mu : \beta \to \mathbb{R}_+$ for which $\mu(\varnothing) = 0$, and if $A, B \in \beta$ are disjoint, then $\mu(A \cup B) = \mu(A) + \mu(B)$. If $\mu(X) = 1$, then μ is said to be a *probability measure.*

Definition A σ *algebra* β of subsets of a space X is a nonempty subcollection $\beta \subset 2^X$ which is closed under complementation, arbitrary unions and intersections.

The following lemma follows the standard paradigm of recurrence results in ergodic theory: given a topological space X which satisfies a suitable 'smallness' condition (e.g. compactness, $\mu(X) < \infty$), and a transformation $T : X \to X$, there exist points of X which satisfy a certain 'almost-periodicity' condition under the action of T.

Lemma 1. *Poincare Recurrence Lemma*

If (X, β, μ, T) *is an m.p.s. and* $\mu(X) < \infty$, *then* $\forall B \in \beta$ *such that* $\mu(B) > 0$, $\exists N \in \mathbb{N}$ *such that* $\mu(B \cap T^N B) > 0$.

Proof. Suppose not. Then, note that $B \cap T^n B = \varphi$ for all $n \in \mathbb{Z}$, and $\cup_{n\in\mathbb{Z}} T^n B \subset X$. This means that $\mu(\cup_{n\in\mathbb{Z}} T^n B) = \sum_{n\in\mathbb{Z}} \mu(B) < \mu(X)$, but this is a contradiction since $\mu(X)$ is finite and is bounded from below by an unbounded quantity. $\square$

For proofs of Van der Waerden's and Szemeredi's Theorems, we will need to consider the notion of *symbolic flow.*

Definition *Symbolic Flow*

Let $X = \Lambda^{\mathbb{Z}}$, where Λ is a finite list of symbols $1, 2, ..., q$ and $x \in X$ is viewed as a function $\mathbb{Z} \to \Lambda$, i.e. a *partition* of $\mathbb{Z}$. Endow X with the metric

$$d(x, x') = \inf_{k>0} \{ \frac{1}{1+k} \mid x(i) = x'(i) \, \forall -k < i < k \}$$

Note that if $d(x, x') < 1$, then we have $x(0) = x'(0)$.

Claim. X is compact-Hausdorff in the metric topology.

Proof. The metric topology on X is strictly coarser than the product topology (product of discrete topologies on Λ), since an open ball about a point $x_0 \in X$ of radius ε can be written as

$$B_\varepsilon(x_0) = \prod_i U_i$$

where $U_i = \Lambda$ for all but finitely many i. Tychonoff's theorem ensures that X is compact in the product topology, since $|\Lambda| < \infty$. Compactness in the metric topology follows. Hausdorff is immediate from the metric. $\square$

Definition Let $T : X \to X$ be the *shift operator*, which, true to its name, is given by $Tx(n) = x(n+1)$.

Claim. $T : X \to X$ is a self-homeomorphism of X.

Proof. Clearly T is a bijection, so it remains to show that T and its inverse are continuous. To prove that T is continuous in the metric topology on $X = \Lambda^{\mathbb{Z}}$, we must show that $\forall \varepsilon > 0$, $\exists \delta_\varepsilon$ such that $d(x, x') < \delta_\varepsilon$ implies $d(Tx, Tx') < \varepsilon$. Let $m > \frac{1}{\varepsilon}$ be a positive integer, and set $\delta_\varepsilon = \frac{1}{m} < \varepsilon$. It is an easy exercise to check that this choice of δ_ε works. A similar method yields continuity for T^{-1}. □

2 Van der Waerden's Theorem and Szemeredi's Theorem

Definition Let $T_1, T_2, ..., T_l$ be l continuous, commuting maps $X \to X$. A point $x \in X$ is said to be *multiply recurrent* if $\exists \{n_k\}_{k\in\mathbb{N}}$ with $n_k \to \infty$ such that $T_i^{n_k} x \to x$ for all $i = 1, 2, ..., l$ as $k \to \infty$.

We can now state (MBR), which will yield a proof of Van der Waerden's theorem.

Theorem 2. *Multiple Birkhoff Recurrence Theorem*

If $T_1, ..., T_l$ satisfy the requirements given above and X is a compact metric space, then there exists a multiply recurrent point.

Corollary 3. *Van der Waerden's Theorem*

If $B_1, B_2, ..., B_q$ comprise a partition of the integers, then for any l, at least one of the partitions B_j contains an $l+1$ progression.

Proof. Let $\Lambda = \{1, 2, 3, ..., q\}$ so that any given partition of the integers is represented by an element $x_0 \in X = \Lambda^{\mathbb{Z}}$. Let T be the shift operator. Then, let $Y = \{T^m x_0\}_{m\in\mathbb{Z}}$, the so-called <u>orbit closure</u> of the point x_0. Note that by construction, $\overline{Y}$ is a closed subspace of a compact space X, so that $\overline{Y}$ is compact.

Let $T_1 = T$, and $T_i = T^i$ for $i = 1, 2, ..., l$. Then, the operators $T_1, T_2, ..., T_l$ all commute, so by the previous theorem, there exists a multiply recurrent point $y \in \overline{Y}$ with respect to this system of operators, i.e. $\exists$ an unbounded sequence of integers n_k such that $T_i^{n_k} y \to y$ for all $i = 1, 2, ..., l$. Recalling the definition of a limit point in a metric space, we see that $\exists n = n_k$ for some $k \in \mathbb{N}$ such that $d(y, T_i^n y) < 1$ for all i. As was noted, this means that $y(0) = T_i^n y(0) = y(ni)$ for $i = 1, 2, ..., l$, so we have established that $\exists y \in \overline{Y}$ such that $y(0) = y(n) = y(2n) = ... = y(ln)$.

Finally, notice that $y \in \overline{Y}$ means that y is a limit point of Y, so that we can select some $y' \in Y$ sufficiently close to y (the choice $d(y, y') < \frac{1}{ln}$ works), with $y' = T^{-m} x_0$ for some $m \in \mathbb{Z}$, so that $y(0) = y'(0) = y(n) = y'(n) = ... = y(ln) = y'(ln)$. So, we have $x_0(m) = x_0(m+n) = x_0(m+2n) = ... = x_0(m+ln)$. Put another way, we have found an $(l+1)$-progression which lies entirely in one of the partitions we started with. □

We now proceed with the machinery necessary to prove Szemeredi's theorem in the integers. Below, TA will refer to the image of the set A under the transformation T.

Theorem 4. *Let (X, β, μ) be a measure space with $\mu(X) < \infty$, and suppose $T_1, ..., T_l$ are l commuting transformations of X which preserve the measure μ. Then, for any $A \in \beta$ with $\mu(A) > 0$, there exists $n \in \mathbb{N}$ such that*

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

Definition Let $\xi \in \Lambda^{\mathbb{Z}}$, and $\lambda \in \Lambda$. The *upper banach density* $BD_\xi(\lambda)$ of λ in ξ is defined as

$$BD_\xi(\lambda) = \limsup_{k\to\infty} \frac{|\{n \mid \xi(n) = \lambda, n \in I_k\}|}{|I_k|}$$

where the supremum is over sequences of intervals $I_k \subset \mathbb{Z}$ with $|I_k| \to \infty$ as $k \to \infty$.

We say that a subset $S \in \mathbb{Z}$ has upper banach density defined in the system $\Lambda = \{0,1\}$ with $\lambda = 1$ and $\xi = 1_S$, the indicator function of S.

Lemma 5. *Let $\Omega = \Lambda^{\mathbb{Z}}$, and let $\xi \in \Omega$. Define $X(\xi) = \{T^m \xi\}_{m \in \mathbb{Z}}$ to be the orbit closure of ξ. For $\lambda \in \Lambda$, define $A(\lambda) = \{\omega(0) = \lambda \mid \omega \in \Omega\}$. Then, if $BD_\xi(\lambda) > 0$, there is a T-invariant probability measure μ on $X(\xi)$ such that $\mu(A(\lambda)) > 0$.*

Proof. I will sketch a proof of the direction we need for Szemeredi's theorem, but the converse holds as well. Let $BD_\xi(\lambda) > 0$, and $A(\lambda) = \{\omega(0) = \lambda \mid \omega \in \Omega\}$. Let $f : X(\xi) \to \mathbb{C}$ be a bounded, continuous function, and suppose $\{I_k\}_{k \in \mathbb{N}}$ is the sequence of intervals in $\mathbb{Z}$ for which we obtain $BD_\xi(\lambda)$. Now, define

$$L(f) = \lim_{k \to \infty} \frac{1}{|I_k|} \sum_{n \in I_k} f(T^n \xi)$$

For any $S \subset X$, let 1_S be the indicator function of S. Then, define $\mu(S) = L(1_S)$. As long as L is well defined, μ is a valid probability measure ($\mu(X(\xi)) = 1$). So, we must check $\mu(A(\lambda)) > 0$. But, we see that

$$\begin{aligned} \mu(A(\lambda)) = L(1_{A(\lambda)}) &= \lim_{k \to \infty} \frac{1}{|I_k|} \sum_{n \in I_k} 1_{A(\lambda)}(T^n \xi) \\ &= \lim_{k \to \infty} \frac{|\{n \mid \xi(n) = \lambda, n \in I_k\}|}{|I_k|} = BD_\xi(\lambda) > 0 \end{aligned}$$

It remains to show that $L(f)$ is well-defined for indicator functions. By passing to a subsequence of $\{I_k\}$ and using a diagonal argument, we can show that $L(f)$ exists for a countable collection of continuous functions. Then, by the separability of this function space, we can choose L for the countable-dense subset. $\square$

We are now in a position to prove the main result of this lecture.

Theorem 6. *Szemeredi's Theorem in the Integers*

Let $S \subset \mathbb{Z}$ have positive upper banach density in the integers. Then, S contains arithmetic progressions of arbitrary length.

Proof. Let $\Lambda = \{0,1\}$, and denote $\Omega = \Lambda^{\mathbb{Z}}$. As before, let T be the shift operator. Define the point $1_S \in \Omega$ to be the indicator function of S, so that $1_S(n) = 1$ if $n \in S$, and $1_S(n) = 0$ otherwise. Assume that $1 \in \Lambda$ occurs in 1_S with positive upper banach density. Now define X to be the orbit closure of 1_S. Define $A(1) = \{\omega(0) = 1 \mid \omega \in \Omega\}$. By Lemma 5, we see that there exists a T-invariant measure μ on X with $\mu(A(1)) > 0$. Now, denote $T_i = T^i$ for $i = 1, 2, ..., l$, where l is an arbitrary positive integer. For brevity, let $A' = A(1) \cap X$. We see that $T_1, ..., T_l$ are commuting transformations, so that by Theorem 4, $\exists n \in \mathbb{N}$ such that

$$\mu(A' \cap T^{-n} A' \cap ... \cap T^{-ln} A') > 0$$

This means that there exists $\omega \in A' \cap T^{-n} A' \cap ... \cap T^{-ln} A'$. Because $\omega \in A(1)$, we have $\omega(0) = 1$. Since $T^{jn} \omega \in A(1)$, we see that $\omega(jn) = 1$ for $j = 1, 2, ..., l$. Furthermore, we have $\omega \in X$, so that $\omega = T^m 1_S$ for some $m \in \mathbb{Z}$. This means that $T^m 1_S(0) = T^m 1_S(n) = ... = T^m 1_S(ln) = 1$, i.e. $1_S(m) = 1_S(m+n) = ... = 1_S(m+ln) = 1$. This is the desired result. $\square$

Corollary 7. *If $S \subset \mathbb{Z}$ has positive upper banach density, then for any positive integer l, the subset S contains infinitely many l-progressions.*

Proof. Fix l- then by Szemeredi's theorem, we know that S contains an l-progression $P_1 \subset S$. Denoting $S_1 = S - P_1$, we see that S_1 has positive upper banach density, so that S_1 contains an l-progression $P_2 \subset S_1$. Since we may iterate this process arbitrarily many times, we see that S contains an infinite sequence of l-progressions $P_1, P_2, ...$ $\square$

3 Applications

Definition Let $H : \mathbb{R}^{2n} \to \mathbb{R}$ be smooth, and let $X \in \mathbb{R}^{2n}$. A *hamiltonian system* is a first order differential system of equations of the form

$$J = \begin{pmatrix} 0_n & I_n \\ -I_n & 0_n \end{pmatrix}$$

$$\frac{dX}{dt} = J\nabla H(X)$$

where I_n is the $n \times n$ identity matrix, and 0_n is the $n \times n$ zero matrix.

Claim. Let $X(t)$ be the (unique) solution corresponding to $X(t_0) = X_0$ for given t_0. Then, $\frac{dH}{dt}(X(t)) = 0$.

Proof. $\frac{dH}{dt} = \nabla H \cdot \frac{dX}{dt} = \nabla H \cdot J\nabla H = 0$ since $Jv \cdot v = 0$ for all $v \in \mathbb{R}^{2n}$. □

H is a so-called *integral of the motion.* This corresponds to the conservation of energy in classical mechanics. Think of solutions $X(t)$ as living on submanifolds of $\mathbb{R}^{2n}$ defined as level surfaces of the Hamiltonian.

Suppose $X_0 \in \mathbb{R}^{2n}$ is an initial condition for this system. Let $T_t : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ act on X_0 by sending it to the (unique!) value $X(t)$ of a solution X which satisfies $X(0) = X_0$. Then, if we let $\mu : \beta \to \mathbb{R}_+$ denote the volume measure of the phase space $\mathbb{R}^{2n}$, we see that T preserves the measure μ due to the theorem of Liouville. So, we can apply Poincare recurrence to conclude that if $X(t)$ is a solution on a level surface S_c of finite volume, then almost all such $X(t)$ tend to periodic solutions.

An interesting consequence: suppose you burn a piece of paper in a closed room with no ventilation, so that the phase space of the system has finite volume. Then there exist arbitarily small perturbations of the initial conditions such that after some (extremely long) time, the paper will reassemble to within arbitrarily small precision. This may seem to contradict the Second Law of Thermodynamics, but the time scales involved in this demonstration are so long (exponential dependence on the number of degrees of freedom, so on the order of Avogadro's number) that there is ample time for Maxwell's demon to take effect and reverse the increase in entropy.

There is another simple application of the method of recurrent points, that of 'Diophantine Inequalities'.

Lemma 8. *Diophantine Inequalities*
Let $p \in \mathbb{R}[X]$ with $p(0) = 0$. Then, for any $\varepsilon > 0$, we can find $m, n \in \mathbb{Z}$ such that

$$|p(n) - m| < \varepsilon$$

Proof. The result follows from the Birkhoff recurrence theorem applied to an appropriately chosen continuous transformation of the d-dimensional torus $T^d \cong [0, 1)^d$. □

Definition A subset $S \subset \mathbb{Z}$ is called an *IP system* (for Infinite Paralelopiped) if there exists an infinite sequence $\{p_k\} \subset S$ such that the set of all finite sums $p_{i_1} + p_{i_2} + ... + p_{i_n}$ is contained in S.

Theorem 9. *(due to Hindeman)*
If $S \subset \mathbb{Z}$ is of positive upper banach density, then S is an IP system.

4 Bibliography

I have made prolific use of Furstenberg's *Recurrence in Ergodic Theory and Combinatorial Number Theory*, and also lecture notes from a class taught by Terrance Tao, available on his blog at terrytao.wordpress.com.