Math 155 (Lecture 28)

November 8, 2011

In the last lecture, we gave a combinatorial proof of the following result:

Theorem 1 (Infinite Version of van der Waerden's Theorem)**.** *Let C be a finite set of colors, and let $f : \mathbf{Z} \to C$ be a coloring of the integers. Then there exist arbitrarily large arithmetic progressions.*

In this lecture, we will describe another proof of this result, due to Furstenberg and Weiss, which uses ideas from topological dynamics.

Notation 2. Fix a finite set C in what follows. Let $Y = C^{\mathbf{Z}}$ denote the set of all colorings of the integers. We define a metric on Y as follows: given a pair of colorings $y, y' : \mathbf{Z} \to C$, we let

$$d(y, y') = \begin{cases} 0 & \text{if } y = y' \\ \frac{1}{n+1} & otherwise. \end{cases}$$

where n denotes the smallest nonnegative integer such that $y(n) \neq y'(n)$ or $y(-n) \neq y'(-n)$.

As a topological space, Y is given by the product

$$\prod_{n \in \mathbf{Z}} C.$$

By Tychanoff's theorem, Y is compact. We define a "translation" operator $T : Y \to Y$, given by $Ty(n) = y(n+1)$. This is a homeomorphism from Y to itself. We let T^n denote the nth power of T (where n is a positive or negative integer). Each T^n is a continuous map from a compact metric space to itself, and therefore uniformly continuous.

By construction, we have $d(y, y') < 1$ if and only if $y(0) = y'(0)$. To say that a coloring y has a monochromatic arithmetic progression of length $k+1$ is to say that there exists integers n and q such that

$$y(q) = y(q+n) = y(q+2n) = \cdots = y(q+kn).$$

This is equivalent to saying that

$$d(T^q y, T^{q+in} y) < 1$$

for $0 \leq i \leq k$.

Given a coloring $y \in Y$, we define the *orbit* of y to be the set

$$O(y) = \{T^n y : n \in \mathbf{Z}\} \subseteq Y.$$

We can then state Theorem 1 as follows:

($*$) For every coloring $y \in Y$ and every $k \geq 0$, there exists a coloring $z \in O(y)$ and an integer $n \geq 1$ such that $d(z, T^{ni} z) < 1$ for $0 \leq i \leq k$.

Let $\overline{O(y)}$ denote the closure of $O(y)$. To prove ($*$), it suffices to verify the following slightly weaker assertion:

($*'$) There exists an integer $n \geq 1$ and a coloring $z \in \overline{O(y)}$ such that $d(z, T^{ni}z) < 1$ for $0 \leq i \leq k$.

To see that ($*'$) implies ($*$), note that if n is fixed then the set

$$\{z \in Y : d(z, T^{ni}z) < 1 \text{ for } 1 \leq i \leq k\}$$

is open in Y. If this set intersects $\overline{O(y)}$, it must also intersect $O(y)$.

Assertion ($*'$) is a little easier to work with because the $X = \overline{O(y)}$ is a closed subset of Y, and therefore compact. We have reduced the proof of Theorem 1 to the following purely topological assertion:

Proposition 3. *Let X be a nonempty compact metric space, let $T : X \to X$ be a homeomorphism. For every integer $k \geq 0$ and every positive real number ϵ, there exists a point $x \in X$ and an integer $n \geq 1$ such that $d(x, T^{ni}x) < \epsilon$ for $1 \leq i \leq k$.*

The first step in the proof of Proposition 3 is to arrange that the action of T on X is particularly simple.

Definition 4. Let $T : X \to X$ be as in Theorem 3. We say that a closed subset $Z \in X$ is *minimal* if Z is nonempty, $T(Z) = Z$, and $Z = \overline{O(z)}$ for each $z \in Z$.

Lemma 5. *Let $T : X \to X$ be as in Theorem 3. Then there exists a minimal closed subset $Z \subseteq X$.*

Proof. Since X is a compact metric space, there exists a countable basis for the topology of X, given by open sets $U_1, U_2, \dots \subseteq X$. We define a decreasing sequence of closed subsets

$$X = X_0 \supseteq X_1 \supseteq X_2 \supseteq \cdots$$

as follows. Assume that X_i has been defined. If $X_i \subseteq \bigcup_n T^n(U_i)$, set $X_{i+1} = X_i$. Otherwise, set $X_{i+1} = X_i - \bigcup T^n(U_i)$. By construction, each X_i is nonempty and invariant under T. Since X is compact, the intersection $Z = \bigcap X_i$ is also nonempty (and invariant under T). We claim that Z is minimal. Suppose otherwise: then there exists $z \in Z$ such that $O(z)$ is not dense in Z. Then we can find an open set U_i such that $Z \cap U_i \neq \emptyset$ and $O(z) \cap U = \emptyset$. Then

$$z \in X_i - \bigcup_n T^n(U_i)$$

, so that by construction $X_{i+1} = X_i - \bigcup_n T^n U_i$. Since $Z \subseteq X_{i+1}$, we have $Z \cap U_i = \emptyset$, contradicting our assumption. □

By virtue of Lemma 5, we can replace X by a minimal closed subset $Z \subseteq X$ and thereby reduce to the case where X is itself minimal. We now proceed by induction on k. If $k = 0$, we can take x to be any point of X. Assume therefore that $k > 0$ and that the Proposition is known for $k - 1$. We regard k as fixed in the arguments which follow. For each $\epsilon > 0$, the inductive hypothesis implies that there exists a point $x \in X$ and an integer $n > 0$ such that $d(x, T^{ni}x) < \epsilon$ for $i = 1, 2, \dots, k - 1$. Taking $y = T^{-n}x$, we have obtain the following weak version of Proposition 3:

(a) For each $\epsilon > 0$, there exists an integer $n > 0$ and a pair of points $x, y \in X$ such that $d(x, T^{ni}y) < \epsilon$ for $1 \leq i \leq k$.

Now fix $\epsilon > 0$, and apply (a) to choose $x, y \in X$ and an integer n such that

$$d(x, T^{ni}y) < \frac{\epsilon}{2}$$

for $1 \leq i \leq k$.

Lemma 6. *For each $x \in X$ and each $\epsilon > 0$, there exists a point $y \in X$ and an integer $n > 0$ such that $d(x, T^{ni}y) < \epsilon$ for $1 \leq i \leq k$.*

Proof. Let B denote an open ball of radius $\frac{\epsilon}{2}$ around X. Then $\bigcup_{m\in\mathbf{Z}} T^m B$ is an open T-invariant subset of X. Since X is minimal, we must have $X = \bigcup_{m\in\mathbf{Z}} T^m B$. Since X is compact, we have $X = \bigcup_{-M\leq m\leq M} T^m B$ for some $M \gg 0$. Each of the maps T^m is uniformly continuous. We may therefore choose a constant δ such that, whenever $d(z,z') < \delta$, we have $d(T^m z, T^m z') < \frac{\epsilon}{2}$ for $-M \leq m \leq M$. Applying (a), we can find points $x', y' \in X$ and an integer $n > 0$ such that $d(x', T^{ni} y') < \delta$ for $1 \leq i \leq k$. We have $x' \in T^m B$ for some $-M \leq m \leq M$. Set $y = T^{-m} y'$. Then for $1 \leq i \leq k$, we have

$$d(x, T^{ni} y) = d(x, T^{-m} x') + d(T^{-m} x', T^{-m} T^{ni} y') < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

□

Now fix a $\epsilon > 0$ and choose an arbitrary point $x_0 \in X$. Choose any positive real number $\delta_0 < \epsilon$. Using Lemma 6, we can find a point $x_1 \in X$ and an integer n_0 such that

$$d(x_0, T^{in_0} x_1) < \delta_0$$

for $1 \leq i \leq k$. Since the functions $T^{n_0}, T^{2n_0}, \ldots, T^{kn_0}$ are uniformly continuous, we can choose a positive real number $\delta_1 \leq \delta_0$ such that $d(y,z) < \delta_0 - d(x_0, T^{in_0} x_1)$ for $1 \leq i \leq k$. Applying Lemma 6 again, we can choose a point $x_2 \in X$ and an integer $n_1 > 0$ such that

$$d(x_1, T^{in_1} x_2) < \delta_1$$

for $1 \leq i \leq k$. Note that this implies that

$$d(T^{in_0} x_1, T^{i(n_0+n_1)} x_2) < \delta_0 - d(x_0, T^{in_0} x_1),$$

so that by the triangle inequality we get

$$d(x_0, T^{i(n_0+n_1)} x_2) < \delta_0$$

for $1 \leq i \leq k$. The functions $T^{n_1}, T^{2n_1}, \ldots, T^{kn_1}$ are again uniformly continuous. We may therefore choose a positive real number $\delta_2 \leq \delta_0$ such that $d(y,z) < \delta_2$ implies that $d(T^{in_1} y, T^{in_1} z) < \delta_1 - d(x_1, T^{in_1} x_2)$ for $i \leq i \leq k$. We may now apply Lemma 6 again to choose $n_2 > 0$ and a point $x_3 \in X$ such that

$$d(x_2, T^{in_2} x_3) < \delta_2$$

for $1 \leq i \leq k$. Arguing as above, we see that this implies

$$d(x_1, T^{i(n_1+n_2)} x_3) < \delta_1$$

and hence

$$d(x_0, T^{i(n_0+n_1+n_2)} x_3) < \delta_0$$

for $1 \leq i \leq n$. Proceeding in this way, we obtain a sequence of points $x_0, x_1, \ldots$ such that for $a < b$ and $1 \leq i \leq k$, we have

$$d(x_a, T^{i(n_a+n_{a+1}+\cdots+n_{b-1})} x_b) < \delta_a \leq \delta_0.$$

Since X is compact, the sequence $x_0, x_1, \ldots$ has a convergent subsequence. In particular, we can choose $a < b$ such that $d(x_a, x_b) < \epsilon - \delta_0$. The triangle inequality then gives

$$d(x_b, T^{i(n_a+\cdots+n_{b-1}} x_b) < \epsilon$$

for $1 \leq i \leq k$, which completes the proof of Proposition 3.