A GLOSS ON A THEOREM OF FURSTENBERG

LESTER E. DUBINS

Abstract. Certain refinements are offered for Furstenberg's ergodic-theoretic version of Szemeredi's theorem.

Furstenberg [1977] has proven a significant generalization of a theorem of Poincaré, which, with no real loss, can be formulated thus: If \( k \) is a positive integer and \( B_1, B_2, \ldots \) is a stationary sequence of events of positive probability in a countably additive probability space, then there is a \( k \)-progression, \( K \), such that \( \bigcap_{k \in K} B_k \) has positive probability. (A \( k \)-progression is a set of \( k \) integers of the form \( \{a, a + b, a + 2b, \ldots, a + (k - 1)b\} \) with \( a > 0, b > 0 \).

The present paper observes that neither the hypothesis of countable additivity nor of stationarity is needed. Moreover, the probability of \( B_K \) can be bounded from below by a \( \delta > 0 \) which depends only on \( k \) and \( p = P(B_1) \). These facts are immediate corollaries to:

**Theorem 1.** Let \( p > 0 \) and let \( k \) be a positive integer. Then there is a \( \delta > 0 \) and a positive integer \( n \) such that, for every \( n \)-tuple of events \( B_1, \ldots, B_n \) of average probability at least \( p \), there is a \( k \)-progression \( K \subset \{1, \ldots, n\} \) for which \( \bigcap_{i \in K} B_i \) has probability at least \( \delta \).

This form of Furstenberg's theorem follows by an argument which he chose not to provide in [1977]. Indeed, it is a simple consequence of Szemeredi's theorem [1975] on the existence of arbitrarily long arithmetic sequences in each set of integers of positive density. But it is convenient first to provide a trivial lemma.

**Lemma 1.** Let \( B_1, \ldots, B_n \) be events of average probability at least \( p \) and let \( l \) be a positive integer less than \( n \). Then there is a subset \( X \) of \( \{1, \ldots, n\} \) of cardinality \( l \) such that

\[
P\left( \bigcap_{i \in X} B_i \right) \geq \left( p - \frac{l}{n} \right) \binom{n}{l}.
\]

**Proof of Lemma 1.** Let \( Y \) be the number of \( B \) that occur. Since \( Y \) is at most \( n \) on the event \( Y \geq l \) and is at most \( l - 1 \) on its complement, the following inequality (sharp) is easily obtained.

\[
P(Y \geq l) \geq \left( \frac{PY}{n} - \frac{l - 1}{n} \right) \left( 1 - \frac{l - 1}{n} \right)^{-1}.
\]
(In (2), the precision (expectation) of \( Y \) is designated by \( PY \) as accords with a notational innovation of de Finetti.)

For the purposes of this note, this weaker inequality suffices:

\[
P(Y > t) \geq \frac{PY}{n} - \frac{l}{n}.
\]

Plainly, the event \( Y > t \) is the union of the events \( \cap B_i (i \in X) \) as \( X \) ranges over \([n]^l\), the subsets of \( \{1, \ldots, n\} \) of cardinality \( l \). Therefore,

\[
P(Y > t) \leq \sum P(\cap B_i (i \in X)) \leq \left( \frac{n}{l} \right) \max P(\cap B_i (i \in X)),
\]

as \( X \) ranges over \([n]^l\). So, for some \( X \in [n]^l \),

\[
P(\cap B_i (i \in X)) \geq P(Y > t) / \left( \frac{n}{l} \right) \geq \left( \frac{PY}{n} - \frac{l}{n} \right) / \left( \frac{n}{l} \right)
\]

\[
\geq \left( p - \frac{l}{n} \right) / \left( \frac{n}{l} \right),
\]

where the second inequality obtains in view of (3), and the third by hypothesis. \( \square \)

Let \( \gamma_k(n) \) be the least integer \( l \) such that, if \( X \) is a subset of \( \{1, \ldots, n\} \) of cardinality \( l \), then \( X \) includes a \( k \)-progression. Szemeredi [1975] has shown that \( \gamma_k(n)/n \) converges to 0 as \( n \to \infty \).

PROOF OF THEOREM 1. By Szemeredi’s theorem, there is an \( n = n(\gamma_2(n)) \) such that \( \gamma_2(n) < np/2 \). For \( l = \gamma_k(n) \), let \( \delta = p/2(\gamma_2(n)) \). That \( (\delta, n) \) satisfies Theorem 1 can be verified, thus. Let \( B_1, \ldots, B_n \) be events of average possibility at least \( p \). By Lemma 1, there is an \( X \subset \{1, \ldots, n\} \) of cardinality \( l \) such that (1) holds. Since \( l/n < p/2 \), the right-hand side of (1) is at least \( \delta \). So \( \cap B_i (i \in X) \) has probability no less than \( \delta \).

Since \( X \) is of cardinality \( \gamma_k(n) \), \( X \) includes a \( k \)-progression, \( K \). Plainly, \( \cap B_i (i \in K) \) includes \( \cap B_i (i \in X) \). So it, too, has probability no less than \( \delta \). \( \square \)

REFERENCES
