COLLOQUIUM MATHEMATICUM

VOL. 84/85 2000 PART 2

ASPECTS OF UNIFORMITY IN RECURRENCE

BY

VITALY BERGELSON (COLUMBUS, OH),
BERNARD HOST (CHAMPS SUR MARNE),
RANDALL McCUTCHEON (COLLEGE PARK, MD)
AND FRANÇOIS PARREAU (VILLETANEUSE)

Abstract. We analyze and cite applications of various, loosely related notions of uniformity inherent to the phenomenon of (multiple) recurrence in ergodic theory. An assortment of results are obtained, among them sharpenings of two theorems due to Bourgain. The first of these, which in the original guarantees existence of sets $\{x, x+h, x+h^2\}$ in subsets E of positive measure in the unit interval, with lower bounds on h depending only on $m(E)$, is expanded to the case of arbitrary finite polynomial configurations in subsets of positive measure in cubes of $\mathbb{R}^n$. The second is a direct computation of a lower bound, uniform in a and b and depending only on $\int f$, for $\int f(x)f(x+at)f(x+bt)\,dx\,dt$, where $0 \le f \le 1$ is a function on the 1-torus. Our methodology parallels that of Bourgain, who originally considered the case $a = 1$, $b = 2$.

1. Introduction. In 1974 Szemerédi proved the following theorem ([S1]).

THEOREM Sz. *For any $l \in \mathbb{N}$ and $\varepsilon > 0$ there exists a constant $N = N(l, \varepsilon)$ such that if a set $S \subset \{1, \ldots, N\}$ satisfies $|S| \ge \varepsilon N$ then S contains an l-term arithmetic progression.*

Szemerédi's proof used combinatorial methods. In 1977 Furstenberg gave an ergodic-theoretic proof of the following multiple recurrence theorem ([F]).

THEOREM F1. *For any probability measure preserving system $(X, \mathcal{B}, \mu, T)$, for any $A \in \mathcal{B}$ with $\mu(A) > 0$, and for any $l \in \mathbb{N}$ there exists $n \in \mathbb{N}$ such that*

$$\mu(A \cap T^{-n}A \cap \ldots \cap T^{-(l-1)n}A) > 0.$$

Furstenberg also gave a correspondence principle linking the fields of density combinatorics and recurrence in ergodic theory. Using this principle, he was able to obtain a new proof of Theorem Sz via Theorem F1. Conversely, Theorem F1 can easily be derived from Theorem Sz and, for example, the ergodic theorem.

Consider the following ostensible strengthening of Theorem F1:

2000 *Mathematics Subject Classification*: 37A, 28D.

THEOREM F2. *For any $l \in \mathbb{N}$ and $\varepsilon > 0$ there exist $M = M(l,\varepsilon)$ and $\delta = \delta(l,\varepsilon) > 0$ such that for any system $(X,\mathcal{B},\mu,T)$ and any $A \in \mathcal{B}$ with $\mu(A) \geq \varepsilon$ there exists n, with $1 \leq n \leq M$, such that*

$$\mu(A \cap T^{-n}A \cap \ldots \cap T^{-(l-1)n}A) \geq \delta.$$

Proof. Let $M = N(l,\varepsilon/2)$ as in Theorem Sz and let J be the number of distinct arithmetic progressions of length l in $\{1,\ldots,M\}$. Put $\delta = \varepsilon/(2J)$. Suppose we are given a probability measure preserving system $(X,\mathcal{B},\mu,T)$ and a set $A \in \mathcal{B}$ with $\mu(A) \geq \varepsilon$. Let

$$f(x) = \frac{1}{M}\sum_{n=1}^{M} 1_{T^{-n}A}(x).$$

Then $\int f\,d\mu \geq \varepsilon$, so letting $B = \{x : f(x) \geq \varepsilon/2\}$, we have $\mu(B) \geq \varepsilon/2$. For every $x \in B$, the set $E_x = \{n : 1 \leq n \leq M,\ x \in T^{-n}A\}$ satisfies $|E_x| \geq (\varepsilon/2)M$ and hence contains an arithmetic progression $I = \{k, k+n, \ldots, k+(l-1)n\}$, which implies that $x \in \bigcap_{t\in I} T^{-t}A$. Hence

$$B \subset \bigcup_{I}\bigcap_{t\in I} T^{-t}A,$$

so that for some I, $\mu(\bigcap_{t\in I} T^{-t}A) \geq \delta$. Since T preserves μ, we are done. ■

We say that F2 is a *uniform* version of F1. Other recurrence theorems admit uniform formulations as well. Although the uniformity of recurrence seems to be a rather transparent (and hence neglected) footnote to the phenomenon of recurrence, there are reasons to pay attention to it. Uniform versions of recurrence theorems can be stronger tools for application than their non-uniform counterparts, and some confusion can result when this is left unaccounted for. At the end of Section 2, an expository section devoted mostly to various examples of "uniform" formulations, we give an example of a multi-parameter multiple recurrence result that is a simple consequence of (the uniform formulation of!) the single parameter case.

Although the constants of uniformity can be shown to exist based on general principles, estimating them is another matter. For the constant $N(l,\varepsilon)$ of Theorem Sz, the original proof of Szemerédi gave an enormous estimate, while Furstenberg's argument gave none at all. The best estimates for this class of problems in simple cases had been achieved via harmonic analysis, an approach dating back to K. Roth, who in [R] showed that $N(3,\varepsilon) \leq \exp\exp(C/\varepsilon)$ for an absolute constant C, but it was only recently that W. T. Gowers ([G3]) was able to show, via highly non-trivial extensions of these methods, that $N(l,\varepsilon) \leq \exp\exp(\delta^{-\exp\exp(k+9)})$. In Section 4 we adopt a slightly different approach. Sticking to three-term configurations, but of the form $\{x, x+an, x+bn\}$, we adopt a specialized formulation of the corresponding generalized Roth theorem and, using methods developed by

Bourgain in [Bo2], achieve in principle estimates that are uniform over all choices (a, b). For more information concerning the best known estimates of various constants of uniformity, the reader is referred to [G1], [G2], [Bo2], [H-B] and [S2].

In Sections 3 and 5, we examine various matters having the flavor of uniformity in $\mathbb{R}$.

In Section 3, we state a finitary version (Theorem 3.2) of a polynomial Szemerédi theorem (see [BL]) and use it to obtain polynomial configurations in large subsets of $\mathbb{R}^l$. The following special case of this result was considered by Bourgain. Let m denote Lebesgue measure on $\mathbb{R}$.

Theorem Bo ([Bo2]). *Given $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that if S is a measurable subset of $[0, N]$, $N \geq 1$, with $m(S) \geq \varepsilon N$, then there exist $x, h \in \mathbb{R}$ with $h > \delta N^{1/2}$ such that $\{x, x + h, x + h^2\} \subset S$.*

We reprove this result (Theorem 3.4), with the exception of finding a lower bound on δ, which Bourgain's method achieves. However, our method allows for generalization to arbitrary finite families of polynomials. For example: given $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that if S is a measurable subset of $[0, N]^2$, $N \geq 1$, with $\lambda(S) \geq \varepsilon N^2$ (λ denotes Lebesgue measure), then there exist $(x, y) \in \mathbb{R}^2$ and $h \in \mathbb{R}$ with $h > \delta N^{1/17}$ such that

$$\{(x, y), (x + h^2, y), (x + h, y + h^3), (x + \sqrt{2}\, h^{17}, y + 5h^4)\} \subset S.$$

Section 5 deals with sets of recurrence in $\mathbb{R}$.

Definition 1.5. A subset R of a topological (semi)group G is called a *set of recurrence* if for any measurable measure preserving flow $(X, \mathcal{B}, \mu, \{T_g\}_{g \in G})$, where $\mu(X) = 1$, and any $A \in \mathcal{B}$ with $\mu(A) > 0$ there exists $g \in R$ such that $\mu(A \cap T_g^{-1} A) > 0$.

The measurability condition in Definition 1.5 is that the map $(x, g) \mapsto T_g x$ from $X \times G$ to X should be measurable, where we are taking the σ-algebra on G to be the σ-algebra of Borel sets.

One can show ([Fo], see also [BH]) that any set of recurrence R in $\mathbb{Z}$ is actually a set of *uniform recurrence*, in the sense that for any $\varepsilon > 0$ there exist a finite subset $R' \subset R$ and a number $\delta > 0$ having the property that for any probability measure preserving system $(X, \mathcal{B}, \mu, T)$ and any $A \in \mathcal{B}$ with $\mu(A) > \varepsilon$ there exists $n \in R'$ such that $\mu(A \cap T^{-n} A) > \delta$.

Remark 1.6. Appropriate modifications of the arguments in either [Fo] or [BH] yield similar uniformities in the case of multiple recurrence. To be precise, define a *set of k-recurrence* to be a set R having the property that for any k commuting measure preserving tranformations $T_1, \ldots, T_k$ of a probability measure space $(X, \mathcal{B}, \mu)$ and any $A \in \mathcal{B}$ with $\mu(A) > 0$ there exists $n \in R$ such that $\mu(A \cap T_1^{-n} A \cap \ldots \cap T_k^{-n} A) > 0$. Then for any set of k-recurrence R and any $\varepsilon > 0$ there exists a finite subset $R' \subset R$ and a number $\delta > 0$ having the property that for any commuting probability measure preserving system $(X, \mathcal{B}, \mu, T_1, \ldots, T_k)$ and any $A \in \mathcal{B}$ with $\mu(A) > \varepsilon$ there exists $n \in R'$ such that $\mu(A \cap T_1^{-n} A \cap \ldots \cap T_k^{-n} A) > \delta$.

The situation changes dramatically when one deals with $\mathbb{R}$-actions. We show in Theorem 5.3 below that if $\{r_n : n \in \mathbb{N}\} \subset \mathbb{R}$ is a set of recurrence and if the set $\{1, r_1, r_2, \ldots\}$ is linearly independent over $\mathbb{Q}$ (for example, $r_n = n^\alpha$ for α suitably chosen), then for any aperiodic flow $(X, \mathcal{B}, \mu, \{T_t\}_{t \in \mathbb{R}})$, $N \in \mathbb{N}$, and $\varepsilon > 0$ there exists a set $A \in \mathcal{B}$ with $\mu(A) > 1/2 - \varepsilon$ such that $\mu(A \cap T_{r_n}^{-1} A) = 0$, $1 \le n \le N$.

2. Four formulations of the multi-dimensional Szemerédi theorem. In this section, we show the equivalence of four "uniform" formulations of the following multiple recurrence theorem of Furstenberg and Katznelson from [FK].

THEOREM FK. *Suppose that $l \in \mathbb{N}$ and $(X, \mathcal{B}, \mu)$ is a probability measure space with commuting measure preserving transformations $T_1, \ldots, T_l$. For any $A \in \mathcal{B}$ with $\mu(A) > 0$,*

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

Here are the four versions. (i) and (iv) are certainly standard, (ii) corresponds to the case $R = \mathbb{N}$ in the remark at the end of the introduction, and (iii) seems to be new. The proofs of all four implications are straightforward. We include them as a service to the reader.

THEOREM 2.1. *The following are equivalent*:

(i) *For any $l \in \mathbb{N}$ and $\varepsilon > 0$ there exists a constant $N_0 = N_0(l, \varepsilon) \in \mathbb{N}$ having the property that for any subset $E \subset \{1, \ldots, N\}^l$ satisfying $|E| \ge \varepsilon N^l$, E contains a configuration of the form*

$$\{(x_1, x_2, x_3, \ldots, x_l), (x_1 + n, x_2, x_3, \ldots, x_l), (x_1, x_2 + n, x_3, \ldots, x_l), \ldots, (x_1, x_2, x_3, \ldots, x_l + n)\}.$$

(ii) *For any $l \in \mathbb{N}$ and for any $\varepsilon > 0$, there exist constants $N_1 = N_1(l, \varepsilon)$ and $\beta = \beta(l, \varepsilon)$ such that, for any l commuting measure preserving transformations $T_1, \ldots, T_l$ of a probability measure space $(X, \mathcal{B}, \mu)$ and any $A \in \mathcal{B}$ with $\mu(A) > \varepsilon$, there exists a positive integer $n \le N_1$ such that*

$$\mu(A \cap T_1^{-n} A \cap \ldots \cap T_l^{-n} A) \ge \beta.$$

(iii) *For any $l \in \mathbb{N}$ and for any $\varepsilon > 0$, there exist constants $N_2 = N_2(l, \varepsilon)$ and $\gamma = \gamma(l, \varepsilon)$ such that, for any l commuting measure preserving transfor-*

mations $T_1,\dots,T_l$ *of a probability measure space* $(X,\mathcal{B},\mu)$, *any* $A\in\mathcal{B}$ *with* $\mu(A)>\varepsilon$, *and any* $N>N_2$ *one has*

$$\frac{1}{N}\sum_{n=1}^{N}\mu(A\cap T_1^{-n}A\cap\ldots\cap T_l^{-n}A)\ge\gamma.$$

(iv) *Let* $l\in\mathbb{N}$ *and let* $\mathbb{T}^l$ *be the torus with normalized Lebesgue measure* λ. *For any* $\varepsilon>0$, *there exists a constant* $\delta=\delta(l,\varepsilon)>0$ *such that*, *for any measurable function* $f:\mathbb{T}^l\to[0,1]$ *with* $\int f\,d\lambda\ge\varepsilon$,

$$\int_{\mathbb{T}^l}\int_0^{1/2} f(x)f(x+t\mathbf{e}_1)\ldots f(x+t\mathbf{e}_l)\,dt\,d\lambda(x)\ge\delta.$$

Remark. The vectors $\mathbf{e}_i$ which appear in (iv) correspond to the ith coordinate vectors in $\mathbb{R}^l$, and the addition there is mod 1.

Proof (of Theorem 2.1). (i)⇒(ii). Let $N_1=N_0(\varepsilon/2)$ and let $\beta=\varepsilon/(2N_1^{l+1})$. For $v=(v_1,\dots,v_l)\in\mathbb{N}^l$, let $A_v=T_1^{-v_1}\ldots T_l^{-v_l}A$. Let

$$f=\frac{1}{N_1^l}\sum_{v\in\{1,\dots,N_1\}^l}1_{A_v}\quad\text{and}\quad B=\{x:f(x)\ge\varepsilon/2\}.$$

As $\mu(A_v)=\mu(A)$ for every v, we have $\int_X f\,d\mu=\mu(A)\ge\varepsilon$. It follows that $\mu(B)\ge\varepsilon/2$. For $x\in B$,

$$|\{v\in\{1,\dots,N_1\}^l:x\in A_v\}|\ge\frac{\varepsilon}{2}N_1^l,$$

which implies that this set contains an "l-simplex" $\{v,v+\mathbf{e}_1,\dots,v+\mathbf{e}_l\}$. Hence

$$B\subset\bigcup_{v,n}(A_v\cap A_{v+n\mathbf{e}_1}\cap\ldots\cap A_{v+n\mathbf{e}_l}),$$

where the union is taken over all v,n such that the simplex is contained in $\{1,\dots,N_1\}^l$. There are fewer than N_1^{l+1} such simplices and thus one of the intersections

$$A_v\cap A_{v+n\mathbf{e}_1}\cap\ldots\cap A_{v+n\mathbf{e}_l}$$

(clearly $1\le n\le N_1$) has measure at least $\varepsilon/(2N_1^{l+1})=\beta$. Therefore,

$$\mu(A\cap T_1^{-n}A\cap\ldots\cap T_l^{-n}A)=\mu(A_v\cap A_{v+n\mathbf{e}_1}\cap\ldots\cap A_{v+n\mathbf{e}_l})\ge\beta.$$

(ii)⇒(iii). Let $K=N_1(l,\varepsilon)$ and $\beta=\beta(l,\varepsilon)$. It is enough to show that for all $M\in\mathbb{N}$ we have

$$\frac{1}{MK^2}\sum_{n=1}^{MK^2}\mu(A\cap T_1^{-n}A\cap\ldots\cap T_l^{-n}A)\ge\frac{\beta}{K^2}.$$

For any $m\in\mathbb{N}$, by substituting T_j^m for T_j in (ii), $1\le j\le m$, we may find at least one $n\in\{1,\dots,K\}$ such that $\mu(A\cap T_1^{-mn}A\cap\ldots\cap T_l^{-mn}A)\ge\beta$.

Now, the products mn, where $M(K-1) < m \le MK$ and $1 \le n \le K$, are pairwise distinct: indeed, if $1 \le n < n' \le K$, then $MKn \le MK(n'-1) \le M(K-1)n'$. Therefore

$$\sum_{n=1}^{MK^2} \mu(A \cap T_1^{-n} A \cap \ldots \cap T_l^{-n} A)$$
$$\ge \sum_{m=M(K-1)+1}^{MK} \sum_{n=1}^{K} \mu(A \cap T_1^{-mn} A \cap \ldots \cap T_l^{-mn} A) \ge M\beta.$$

(iii)⇒(iv). Let $\delta = \frac{1}{2}(\varepsilon/2)^{l+1}\gamma(l, \varepsilon/2)$. Suppose $f : \mathbb{T}^l \to [0,1]$ with $\int f\, d\lambda \ge \varepsilon$. Since f is bounded, the map

$$t \mapsto I(t) = \int_{\mathbb{T}^l} f(x) f(x + t\mathbf{e}_1) \ldots f(x + t\mathbf{e}_l)\, d\lambda(x)$$

is continuous and

$$\int_0^{1/2} I(t)\, dt = \lim_{N\to\infty} \frac{1}{2N} \sum_{n=1}^{N} I\left(\frac{n}{N}\right). \tag{2.1}$$

Let $A = \{x : f(x) \ge \varepsilon/2\}$. Then $\lambda(A) \ge \varepsilon/2$. Fix N and let

$$T_j(x) = x + \frac{1}{2N}\mathbf{e}_j \pmod 1, \qquad j = 1, \ldots, l;\ x \in \mathbb{T}^l.$$

Then $T_1, \ldots, T_l$ are commuting and measure preserving. Furthermore,

$$I\left(\frac{n}{N}\right) \ge \left(\frac{\varepsilon}{2}\right)^{l+1} \lambda(A \cap T_1^{-n} A \cap \ldots \cap T_l^{-n} A), \qquad 1 \le n \le N.$$

It follows that for N sufficiently large,

$$\frac{1}{2N} \sum_{n=1}^{N} I\left(\frac{n}{2N}\right) \ge \frac{1}{2}\left(\frac{\varepsilon}{2}\right)^{l+1} \frac{1}{N} \sum_{n=1}^{N} \lambda(A \cap T_1^{-n} A \cap \ldots \cap T_l^{-n} A)$$
$$\ge \frac{1}{2}\left(\frac{\varepsilon}{2}\right)^{l+1} \gamma\left(l, \frac{\varepsilon}{2}\right) = \delta,$$

which together with (2.1) implies that $\int_0^{1/2} I(t)\, dt \ge \delta$, as desired.

(iv)⇒(i). Let $\delta = \delta(l, \varepsilon(1/4)^l)$ and choose N so that $N > 2/\delta$. Suppose now that $E \subset \{1, \ldots, N\}^l$ with $|E| \ge \varepsilon N^l$. Let

$$S = \frac{1}{2N}E - \left(0, \frac{1}{4N}\right)^l = \left\{\frac{1}{2N}\mathbf{u} - \mathbf{v} : \mathbf{u} \in E,\ \mathbf{v} \in \left(0, \frac{1}{4N}\right)^l\right\} \subset \mathbb{T}^l.$$

Then $\lambda(S) \geq \varepsilon(1/4)^l$, so by (iv) we have

$$\int_{\mathbb{T}^l} \int_0^{1/2} 1_S(x) 1_S(x + t\mathbf{e}_1) \ldots 1_S(x + t\mathbf{e}_l)\, d\lambda(x)\, dt \geq \delta.$$

In particular, for some t with $\delta/2 \leq t \leq 1/2$ and some $x \in \mathbb{T}^l$ we have

$$\{x, x + t\mathbf{e}_1, \ldots, x + t\mathbf{e}_l\} \subset S.$$

Letting $\mathbf{e}_0 = 0$, we may write $x + t\mathbf{e}_j = \frac{1}{2N}\mathbf{u}_j - \mathbf{v}_j$, where $\mathbf{u}_j \in E$ and $\mathbf{v}_j \in (0, 1/(4N))^l$, $0 \leq j \leq l$. For $j = 1, \ldots, l$, we have $t\mathbf{e}_j = \frac{1}{2N}(\mathbf{u}_j - \mathbf{u}_0) + (\mathbf{v}_0 - \mathbf{v}_j)$, from which we get $\mathbf{u}_j - \mathbf{u}_0 = 2N(t\mathbf{e}_j + (\mathbf{v}_j - \mathbf{v})) = 2Nt\mathbf{e}_j + \mathbf{c}_j$, where $\mathbf{c}_j \in (-1/2, 1/2)^l$. It follows that $\mathbf{u}_j - \mathbf{u}_0 = [2Nt + 1/2]\mathbf{e}_j$ (where $[z]$ denotes the greatest integer $\leq z$). Letting $n = [2Nt + 1/2]$, we have

$$\{\mathbf{u}_0, \mathbf{u}_0 + n\mathbf{e}_1, \ldots, \mathbf{u}_0 + n\mathbf{e}_l\} \subset E. \blacksquare$$

We now come to the following two-parameter multiple recurrence theorem, which was alluded to in the introduction as a substantiating reason for making explicit formulations of uniform versions of multiple recurrence results. A special case of Theorem 2.2 below has been given a proof ([La]) parallel in methodology (and therefore of comparable difficulty) to the proof of Theorem Sz given in [FKO]. Observe, however, that no such parallel proof is necessary, for the two-parameter theorem is a consequence of the uniform version 2.1(iii) of the one-parameter theorem.

THEOREM 2.2. *Let $(X, \mathcal{B}, \mu)$ be a probability space and suppose that T and S are commuting measure preserving transformations of X. Then for every $l \in \mathbb{N}$ and every $A \in \mathcal{B}$ with $\mu(A) > 0$ one has*

$$\liminf_{M,N\to\infty} \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \mu\Big(\bigcap_{0\leq i,j\leq l-1} (T^{in}S^{jm})^{-1}A \Big) > 0.$$

P r o o f. By Theorem 2.1(iii) there exist $N_0 \in \mathbb{N}$ and $\gamma_1 > 0$ having the property that for every $N \geq N_0$ we have

$$\frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-n}A \cap \ldots \cap T^{-(l-1)n}A) > \gamma_1. \tag{2.2}$$

Similarly, there exist $M_0 \in \mathbb{N}$ and $\gamma_2 > 0$ having the property that for every set $B \in \mathcal{B}$ with $\mu(B) \geq \gamma_1/2$ and every $M \geq M_0$ we have

$$\frac{1}{M} \sum_{m=1}^{M} \mu(B \cap S^{-m}B \cap \ldots \cap S^{-(l-1)m}B) > \gamma_2. \tag{2.3}$$

Suppose now that $N > N_0$ and $M > M_0$. Let

$$H = \{n : 1 \leq n \leq N_0,\ \mu(A \cap T^{-n}A \cap \ldots \cap T^{-(l-1)n}A) \geq \gamma_1/2\}.$$

By (2.2) we have $|H| \geq N\gamma_1/2$. Hence

$$\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\mu\Big(\bigcap_{0\leq i,j\leq l-1}(T^{in}S^{jm})^{-1}A\Big)$$

$$= \frac{1}{N}\sum_{n=1}^{N} g\Big(\frac{1}{M}\sum_{m=1}^{M}\mu\Big(\bigcap_{0\leq j\leq l-1}(S^{-jm}(A\cap T^{-n}A\cap\ldots\cap T^{-(l-1)n}A))\Big)\Big)$$

$$\geq \frac{1}{N}\sum_{n\in H}\Big(\frac{1}{M}\sum_{m=1}^{M}\mu\Big(\bigcap_{0\leq j\leq l-1}(S^{-jm}(A\cap T^{-n}A\cap\ldots\cap T^{-(l-1)n}A))\Big)\Big)$$

$$\geq \frac{1}{N}\sum_{n\in H}\gamma_2 = \frac{|H|\gamma_2}{N} \geq \frac{\gamma_1\gamma_2}{2}. \ \blacksquare$$

It is easy to see in the above proof that N_0 actually need not depend on A, but only on $\mu(A)$. The proof would then give a uniform conclusion serving as the initial case in an inductive scheme yielding the following multi-parameter Szemerédi theorem:

Theorem 2.3. *Let $(X,\mathcal{B},\mu)$ be a probability space and suppose that $\{T_j\}_{j=1}^{t}$ are commuting measure preserving transformations of X. Then for every $l\in\mathbb{N}$ and every $A\in\mathcal{B}$ with $\mu(A)>0$ one has*

$$\liminf_{N_1,\ldots,N_t\to\infty}\frac{1}{N_1\ldots N_t}\sum_{n_1=1}^{N_1}\ldots\sum_{n_t=1}^{N_t}\mu\Big(\bigcap_{0\leq j_1,\ldots,j_t\leq l}(T_1^{j_1n_1}\ldots T_t^{j_tn_t})^{-1}A\Big) > 0.$$

3. A polynomial Szemerédi type theorem for $\mathbb{R}^l$. Our plan in this section is to generalize Theorem Bo from the introduction. This involves giving a polynomial version of formulation (iv) from Theorem 2.1. This is accomplished via a polynomial version (Theorem 3.2) of formulation (i) from Theorem 2.1. We leave it to the reader to show that polynomial versions of formulations (ii) and (iii) could be given as well. However, due to non-linearity, we see no simple way of obtaining the implication (iv)⇒(i) for the polynomial case.

We use the following polynomial Szemerédi theorem.

Theorem 3.1 ([BL]). *Let $k,l\in\mathbb{N}$ and suppose that $S\subset\mathbb{Z}^l$ is a set of positive upper Banach density*

$$d^*(S) = \limsup_{N_i-M_i\to\infty,\,1\leq i\leq l}\frac{|S\cap\prod_{i=1}^{l}\{M_i+1,\ldots,N_i\}|}{\prod_{i=1}^{l}(N_i-M_i)} > 0$$

and let $p_{i,j}(n)\in\mathbb{Z}[n]$ with $p_{i,j}(0)=0$, $1\leq i\leq k$, $1\leq j\leq l$. Then there exist $n\in\mathbb{N}$ and $(u_1,\ldots,u_l)\in\mathbb{Z}^l$ such that

$$(u_1+p_{i,1}(n),\ldots,u_l+p_{i,l}(n))\in S,\qquad 1\leq i\leq k.$$

Unfortunately, Theorem 3.1 is not quite a polynomial version of (i) from Theorem 2.1. The following is. Its derivation from Theorem 3.1 is based on a very simple idea: if it were possible to have arbitrarily large blocks of density ε, each of which does not contain a polynomial configuration of a given type, then one could build a set of upper Banach density ε not containing such a configuration. As easy as this statement is to believe, the proof is rather technical, so we defer it to an appendix.

THEOREM 3.2. *Let $\varepsilon > 0$, $k, l \in \mathbb{N}$, and let $p_{i,j}(n) \in \mathbb{Z}[n]$ with $p_{i,j}(0) = 0$, $1 \le i \le k$, $1 \le j \le l$. Then there exists $N = N(\varepsilon) \in \mathbb{N}$ such that for every set $S \subset \{1, \ldots, N\}^l$ satisfying $|S| \ge \varepsilon N^l$ there exists $(u_1, \ldots, u_l) \in S$ and $n \in \mathbb{N}$, $1 \le n \le N$, such that*

$$(u_1 + p_{i,1}(n), \ldots, u_l + p_{i,l}(n)) \in S, \quad 1 \le i \le k.$$

Proof. See appendix.

We need the following lemma, a more general version of which is given in Lemma 3.5 below.

LEMMA 3.3. *Let $0 < \varepsilon < 1$ and suppose that $A \subset [0, \varepsilon/2]$ is a finite set. Suppose that $S \subset [0, 1]$ is a measurable set with $m(S) \ge \varepsilon$. Then*

$$m(\{x : |(x + A) \cap S|/|A| \ge \varepsilon/4\}) \ge \varepsilon/4.$$

Proof. Take $l = 1$ in Lemma 3.5 below. ∎

We now see how Theorem 3.2 and Lemma 3.3 combine to give Theorem Bo. Let us outline the basic strategy of the proof, as it is the same for Theorems 3.6 and 3.7 below.

We seek a configuration $\{x, x + h, x + h^2\}$ in a measurable set S of "density" ε (relative to the interval in which S lies). We choose by Theorem 3.2 a square in $\mathbb{N}^2$ so big that any subset of it having density $\varepsilon/4$ (based on counting measure) contains a configuration $\{(s, t), (s + n, t), (s, t + n^2)\}$. We then construct a linear map from the square into $\mathbb{R}$ taking the configurations we can find in the square to the configurations we seek in $\mathbb{R}$. The proof is completed by "sliding" the image of this linear map around in $\mathbb{R}$ until it intersects S with density $\varepsilon/4$, which is possible by Lemma 3.3. The preimage of S under the map must now contain a configuration $\{(s, t), (s + n, t), (s, t + n^2)\}$, so that in particular S contains a configuration $\{x, x + h, x + h^2\}$.

THEOREM 3.4 ([Bo2]). *Given $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that if S is a measurable subset of $[0, N]$, $N \ge 1$, with $m(S) \ge \varepsilon N$, then there exist $x, h \in \mathbb{R}$ with $h > \delta N^{1/2}$ such that $\{x, x + h, x + h^2\} \subset S$.*

Proof. Let $\varepsilon > 0$. By Theorem 3.2 there exists $L \in \mathbb{N}$ having the property that every set $E \subset \{1, \ldots, L\}^2$ satisfying $|E| > (\varepsilon/4)L^2$ contains a configuration of the form $\{(s, t), (s + n, t), (s, t + n^2)\}$, where $1 \le n \le L$.

Fix δ with $0 < \delta < \varepsilon/(4L)$. Suppose now that $N > 1$ and $S \subset [0, N]$ is a measurable set with $m(S) \geq \varepsilon N$. Let $S' = (1/N)S = \{(1/N)s : s \in S\}$. Then $S' \subset [0,1]$ and $m(S) \geq \varepsilon$.

Choose a number $M \in \mathbb{R}$ which is not a rational multiple of $N^{1/2}$ such that $\delta L < 1/M < \varepsilon/4$. Consider the map $f : \{1, \ldots L\} \times \{1, \ldots, L\} \to \mathbb{R}$ given by

$$f(s,t) = \frac{s}{N^{1/2}LM} + \frac{t}{L^2M^2}.$$

Let $A = f(\{1, \ldots, L\}^2)$. Since

$$f(s,t) \leq f(L,L) = \frac{1}{M} + \frac{1}{LM^2} < \frac{\varepsilon}{2}$$

for all $1 \leq s,t \leq L$, we have $A \subset [0, \varepsilon/2]$. Furthermore, since M is not a rational multiple of $N^{1/2}$, one may show that $|A| = L^2$.

Using part of the strength of Lemma 3.3, we see that there exists $x' \in \mathbb{R}$ such that $|(x' + A) \cap S'| \geq (\varepsilon/4)L^2$. Let

$$E = \{(s,t) \in \{1, \ldots, L\}^2 : x' + f(s,t) \in S'\}.$$

Then $|E| \geq (\varepsilon/4)L^2$, so that E contains a configuration $\{(s,t), (s+n,t), (s, t+n^2)\}$, where $1 \leq n \leq L$. Let

$$x = N\left(x' + \frac{s}{N^{1/2}LM} + \frac{t}{L^2M^2}\right) \quad \text{and} \quad h = nN^{1/2}/LM.$$

One may now check that

$$\{x, x+h, x+h^2\} \subset S.$$

Furthermore, $h \geq \frac{1}{LM}N^{1/2} > \delta N^{1/2}$, as required. ∎

As promised, here is the stronger version of Lemma 3.3.

Lemma 3.5. *Suppose that* $1 > \varepsilon > 0$, $l \in \mathbb{N}$, *and* $A \subset [0, \varepsilon/(2l)]^l$ *is a finite set. Suppose that* $S \subset [0,1]^l$ *is a measurable set with* $\lambda(S) \geq \varepsilon$. *Then*

$$\lambda(\{x \in [0,1]^l : |(x+A) \cap S|/|A| \geq \varepsilon/4\}) \geq \varepsilon/4.$$

Proof. Since $A \subset [0, \varepsilon/(2l)]^l$ and $\lambda(S) \geq \varepsilon$, for each $a \in A$ we have

$$\int_{[0,1]^l} 1_S(a+x)\,dx \geq \lambda(S) - l\frac{\varepsilon}{2l} > \frac{\varepsilon}{2}.$$

Therefore,

$$\int_{[0,1]^l} |(x+A) \cap S|\,dx = \int_{[0,1]^l} \sum_{a \in A} 1_S(a+x)\,dx = \sum_{a \in A} \int_{[0,1]^l} 1_S(a+x)\,dx \geq |A|\frac{\varepsilon}{2},$$

so that

$$\int_{[0,1]^l} \frac{|(x+A)\cap S|}{|A|}\,dx \ge \frac{\varepsilon}{2}.$$

It follows that if we let

$$U = \{x \in [0,1]^l : |(x+A)\cap S|/|A| \ge \varepsilon/4\},$$

then

$$\int_{U} \frac{|(x+A)\cap S|}{|A|}\,dx \ge \frac{\varepsilon}{4}.$$

Therefore, $\lambda(U) \ge \varepsilon/4$. ■

We are almost ready to give the main result of this section. As the proof is very technical, we choose to do another special case here that is more indicative of the ideas necessary, relegating the proof of the general case to the appendix.

THEOREM 3.6. *Given $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that if S is a measurable subset of $[0,N]^2$, $N \ge 1$, with $\lambda(S) \ge \varepsilon N^2$, then there exist $(x,y) \in \mathbb{R}^2$ and $h \in \mathbb{R}$ with $h > \delta N^{1/3}$ such that*

$$\{(x,y), (x+h^2-\sqrt{7}\,h, y), (x+h, y+3h^3+\pi h)\} \subset S.$$

Proof. Let $\varepsilon > 0$. By Theorem 3.2 there exists $L \in \mathbb{N}$ having the property that every set $E \subset \{1,\dots,L\}^5$ satisfying $|E| > (\varepsilon/4)L^5$ contains a configuration of the form

$$\{(s,t,u,v,w), (s+n^2, t+n, u, v, w), (s,t,u+n, v+n^3, w+n)\},$$

where $1 \le n \le L$. Fix a number δ with $0 < \delta < \frac{1}{7}\frac{\varepsilon}{8L}$. Suppose that $N > 1$ and $S \subset [0,N]^2$ is a measurable set with $\lambda(S) \ge \varepsilon N^2$. Let $S' = (1/N)S$. Then $S' \subset [0,1]^2$ and $\lambda(S) \ge \varepsilon$.

Choose a number $M \in \mathbb{R}$ which is not a root of any polynomial whose coefficients are taken from $\mathbb{Q}[N^{1/3}, N^{2/3}, \sqrt{7}, \pi]$ (this is the countable field generated by $\mathbb{Q}$ and these 4 numbers) such that $\delta L < 1/M < \varepsilon/56$. Consider the map

$$f(s,t,u,v,w) = s\left(\frac{1}{L^2M^2N^{1/3}}, 0\right) + t\left(\frac{-\sqrt{7}}{N^{2/3}LM}, 0\right) + u\left(\frac{1}{LMN^{2/3}}, 0\right) + v\left(0, \frac{3}{L^3M^3}\right) + w\left(0, \frac{\pi}{LMN^{2/3}}\right), \quad 1 \le s,t,u,v,w \le L.$$

Then f is a linear map whose coefficients are determined by the monomial expressions appearing in the formulation of the theorem, namely $p_1(h) = h^2$, $p_2(h) = -\sqrt{7}\,h$, $p_3(h) = h$, $p_4(h) = 3h^3$, and $p_5(h) = \pi h$. For example, the coefficient of s is precisely $\left(\frac{1}{N}p_1\left(\frac{N^{1/3}}{LM}\right), 0\right)$. The coefficient of t is

$\big(\frac{1}{N}p_2\big(\frac{N^{1/3}}{LM}\big),0\big)$. The coefficient of u is $\big(0,\frac{1}{N}p_3\big(\frac{N^{1/3}}{LM}\big)\big)$, and so forth. It will become clear presently why the choices are made in this fashion.

Let $A=(\varepsilon/8,\varepsilon/8)+f(\{1,\dots,L\}^5)$. One checks that the coordinates of $f(s,t,u,v,w)$ are each not greater in absolute value than $7/M$, which is less than $\varepsilon/8$, $1\le s,t,u,v,w\le L$. Therefore $A\subset[0,\varepsilon/4]^2$. Furthermore, since M is not a rational combination of the elements in the set $\{1,N^{1/3},N^{2/3},\sqrt{7},\pi\}$, one may show that $|A|=L^5$.

By Lemma 3.5, there exists $x'\in\mathbb{R}^2$ such that $|(x'+A)\cap S'|\ge(\varepsilon/4)L^5$. Let

$$E=\{(s,t,u,v,w)\in\{1,\dots,L\}^5 : x'+f(s,t,u,v,w)\in S'\}.$$

Then $|E|\ge(\varepsilon/4)L^5$, so E contains a configuration

$$\{(s,t,u,v,w),(s+n^2,t+n,u,v,w),(s,t,u+n,v+n^3,w+n)\},$$

where $1\le n\le L$.

Let $(x,y)=N(x'+f(s,t,u,v,w))$ and put $h=nN^{1/3}/(LM)$. One may now check that

$$\{(x,y),(x+h^2-\sqrt{7}\,h,y),(x+h,y+3h^3+\pi h)\}\subset S.$$

Furthermore, $h\ge\frac{1}{LM}N^{1/3}>\delta N^{1/3}$, as required. ■

We now come to the primary result of this section, Theorem 3.7. Bourgain mentions in [Bo2] that his method may be modified to prove the cases $l=1$, $k=2$, $p_1(y)=y$, $p_2(y)=y^t$, $t\in\mathbb{N}$, though only $t=2$ (i.e. our Theorem 3.4) is carried out explicitly.

THEOREM 3.7. *Let $\varepsilon>0$, $k,l\in\mathbb{N}$, and let $p_{i,j}(x)\in\mathbb{R}[x]$ with $p_{i,j}(0)=0$, $1\le i\le k$, $1\le j\le l$. Let $t=\max_{1\le i\le k,\,1\le j\le l}\deg p_{i,j}$ and write $\mathbf{p}_i(x)=(p_{i,1}(x),\dots,p_{i,l}(x))$. There exists $\delta>0$ having the property that for any $N>1$, if $f\in L^\infty(\mathbb{R}^l)$, $0\le f\le1$, and $\int_{[0,N]^l}f\,d\lambda>\varepsilon N^l$ then*

$$\int_{[0,N]^l}\int_0^{N^{1/t}} f(\mathbf{u})f(\mathbf{u}+\mathbf{p}_1(y))f(\mathbf{u}+\mathbf{p}_2(y))\dots f(\mathbf{u}+\mathbf{p}_k(y))\,dy\,d\lambda(\mathbf{u})\ge\delta N^{l+1/t}.$$

Proof. See appendix.

COROLLARY 3.8. *Let $\varepsilon>0$, $k,l\in\mathbb{N}$, and let $p_{i,j}(x)\in\mathbb{R}[x]$ with $p_{i,j}(0)=0$, $1\le i\le k$, $1\le j\le l$. Let $t=\max_{1\le i\le k,\,1\le j\le l}\deg p_{i,j}$. There exists $\delta>0$ having the property that for any $N>1$, if $S\subset[0,N]^l$ is a measurable set with $\lambda(S)\ge\varepsilon N^l$ then there exist $(x_1,\dots,x_l)\in\mathbb{R}^l$ and $h\in\mathbb{R}$ with $h\ge\delta N^{1/t}$ such that*

$$\begin{aligned}\{(x_1,\dots,x_l),&(x_1+p_{1,1}(h),\dots,x_l+p_{1,l}(h)),\\ &(x_1+p_{2,1}(h),x_2+p_{2,2}(h),\dots,x_l+p_{2,l}(h)),\dots,\\ &(x_1+p_{k,1}(h),\dots,x_l+p_{k,l}(h))\}\subset S.\end{aligned}$$

4. A generalized Roth theorem with estimates. Our small contribution to the question of estimating the uniform constants of recurrence is the following theorem:

THEOREM 4.1. *For any $\varepsilon > 0$ there exists a constant $\delta = \delta(\varepsilon) > 0$ (which may be explicitly identified) such that for any integers a, b and every measurable function $f : \mathbb{T} \to [0,1]$ with $\int f(x)\,dx \ge \varepsilon$,*

$$\int_{\mathbb{T}^2} f(x)f(x+at)f(x+bt)\,dx\,dt \ge \delta.$$

Note that as the bounds we find on δ are not especially impressive, we will satisfy ourselves with finding a recurrence relation from which they may be computed.

In [Bo1], Bourgain gives a proof of Theorem 4.1 restricted to the case $a = 1$, $b = 2$ by harmonic analysis, which yields Roth's theorem. We follow the main arguments of [Bo1], with some modification in order to find a constant which does not depend on a and b. We then indicate how Theorem 4.1 may be used to obtain explicit gap estimates for a generalized Roth theorem and for the special case $l = 2$ of Theorem 2.1(ii) in which T_1 and T_2 are both powers of the same transformation T.

When $a, b \in \mathbb{Z}$ are fixed and $f : \mathbb{T} \to [0,1]$ is measurable we define

$$J(f) = \int_{\mathbb{T}^2} f(x)f(x+at)f(x+bt)\,dx\,dt.$$

If $a = 0$, $b = 0$ or $a = b$, it is elementary to check that $J(f) \ge (\int f(x)\,dx)^2$. So, we henceforth suppose that a and b are non-zero and distinct.

As usual we denote the Fourier transform of f by $\widehat{f}$, that is, for $n \in \mathbb{Z}$, $\widehat{f}(n) = \int f(x)\exp(-2\pi inx)\,dx$. Also we denote by $f * g$ the convolution $f * (x) = \int f(x-y)g(y)\,dy$. For $f, g \in L^1$ we have $\widehat{f * g} = \widehat{f}\widehat{g}$. Finally we have Parseval's identity: if $f \in L^2$ then $\widehat{f} \in l^2$ and $\|f\|_2 = \|\widehat{f}\|_2$. The following norm plays a crucial role here: let, for any bounded measurable function f on $\mathbb{T}$,

$$|||f||| = \|\widehat{f}\|_4 = \Big(\sum_{n\in\mathbb{Z}} |\widehat{f}(n)|^4\Big)^{1/4}.$$

By the Cauchy–Schwarz inequality we have $|||f||| \le \|f\|_2^2$. More generally, $\|f * g\|_2 = \|\widehat{f}\,\widehat{g}\|_2 \le \|\widehat{f}\|_4\|\widehat{g}\|_4 = |||f||| \cdot |||g|||$. We will use these facts repeatedly.

LEMMA 4.2. *For every measurable $f, g : \mathbb{T} \to [0,1]$,*

$$|J(f) - J(g)| \le 3|||f - g|||.$$

Proof. Let $h(x)=f(x)-g(x)$ and notice that

$$\begin{aligned}f(x)f(x+at)f(x+bt)-g(x)g(x+at)g(x+bt)\\ =h(x)g(x+at)g(x+bt)+f(x)h(x+at)g(x+bt)\\ +f(x)f(x+at)h(x+bt).\end{aligned}$$

Define

$$J(f_1,f_2,f_3)=\iint_{\mathbb{T}^2} f_1(x)f_2(x+at)f_3(x+bt)\,dx\,dt.$$

We must show that $|J(h,g,g)+J(f,h,g)+J(f,f,h)|\le 3|||h|||$. First, we show $|J(f,f,h)|\le |||h|||$:

$$\begin{aligned}J(f,f,h)^2 &\le \|f\|_2^2\int_{\mathbb{T}}\Big(\int_{\mathbb{T}} f(x+at)h(x+bt)\,dt\Big)^2dx\\
&=\|f\|_2^2\int_{\mathbb{T}^3} f(x+as)f(x+at)h(x+bs)h(x+bt)\,dx\,ds\,dt\\
&=\|f\|_2^2\int_{\mathbb{T}^3} f(y)f(y+au)h(z)h(z+bu)\,dy\,dz\,du\\
&=\|f\|_2^2\int_{\mathbb{T}^2} f(y)f(y+au)\Big(\int_{\mathbb{T}} h(z)h(z+bu)\,dz\Big)\,dy\,du\\
&\le \int_{\mathbb{T}}\Big|\int_{\mathbb{T}} h(z)h(z+bu)\,dz\Big|\,du\\
&=\int_{\mathbb{T}}|h*\check{h}(bu)|\,du=\|h*\check{h}\|_1\le\|h*\check{h}\|_2,\end{aligned}$$

where $\check{h}(x)=h(-x)$. We have $\overline{\widehat{h}}=\widehat{\check{h}}$, whence

$$J(f,f,h)^2\le\|h*\check{h}\|_2=\Big(\sum_{j\in\mathbb{Z}}(\widehat{h}(j)\widehat{\check{h}}(j))^2\Big)^{1/2}=\Big(\sum_{j\in\mathbb{Z}}|\widehat{h}(j)|^4\Big)^{1/2}=|||h|||^2.$$

Similar arguments show that $|J(f,h,g)|\le |||h|||$ and $|J(h,g,g)|\le |||h|||$, completing the proof. ■

We shall need auxiliary "kernels", i.e. continuous functions k on $\mathbb{T}$ such that

(4.1) $$k\ge 0,\quad \widehat{k}\ge 0\quad\text{and}\quad \int_{\mathbb{T}} k(x)\,dx=1.$$

In the following lemma, $\|\cdot\|$ denotes the distance to the nearest integer. Also we will use the fact that if $f,g\in L^\infty(\mathbb{T})$ then

(4.2) $$\widehat{fg}(n)=\sum_{k\in\mathbb{Z}}\widehat{f}(n-k)\widehat{g}(k).$$

Lemma 4.3. *Given a finite set* $S \subset \mathbb{Z}$ *and* η, *with* $0 < \eta \le 1/2$, *there exists a kernel* k *satisfying* (4.1) *which vanishes outside the set*

$$A(S,\eta) = \{t \in \mathbb{T} : \|jt\| \le \eta \text{ for all } j \in S\} \tag{4.3}$$

and such that moreover

$$\|\widehat{k}\|_1 = \|k\|_\infty \le \eta^{-|S|}.$$

Proof. Consider the standard "triangle" kernel k_0 given by $k_0(t) = (\eta - |t|)/\eta^2$ if $-\eta \le t \le \eta \pmod 1$ and $k_0(t) = 0$ otherwise. One easily checks that $\widehat{k}_0(j) \ge 0$ for every $j \in \mathbb{Z}$ and that $\widehat{k}_0(0) = 1$. Put $k_1(t) = \prod_{j\in S} k_0(jt)$, so that $k_1(t) = 0$ outside $A(S,\eta)$. Repeated application of (4.2) shows that $\widehat{k}_1(j) \ge 0$ for all $j \in \mathbb{Z}$ and that moreover $\widehat{k}_1(0) \ge 1$. Now just let

$$k(t) = k_1(t)/\widehat{k}_1(0), \quad t \in \mathbb{T}. \ \blacksquare$$

For $f : \mathbb{T} \to [0,1]$ let f_t denote the function $f_t(x) = f(x+t)$. Set

$$d(t) = \max\{|||f - f_t|||, |||f - f_{at}|||, |||f - f_{bt}|||\} \tag{4.4}$$

and for $\delta > 0$ put

$$B(\delta) = \{t \in \mathbb{T} : d(t) < \delta\}. \tag{4.5}$$

Lemma 4.4. *For every* $\delta > 0$, *there exists* $\nu = \nu(\delta) > 0$ (*which may be identified explicitly*), *having the property that for all* $a, b \in \mathbb{Z}$ *and every* $f : \mathbb{T} \to [0,1]$ *with* $\int f(x)\,dx \le 1$ *there exists a kernel* k *satisfying* (4.1) *which vanishes outside* $B(\delta)$ (*where* $B(\delta)$ *is given by* (4.4) *and* (4.5)) *and such that* $\|k\|_\infty \le 1/\nu$. *In particular since* $\int_{\mathbb{T}} k(x)\,dx = 1$ *we have*

$$\lambda(B(\delta)) \ge \nu(\delta).$$

Proof. Let $\nu = (4\pi\sqrt{2}/\delta^2)^{-96/\delta^4}$ and suppose a, b and f are given. Let

$$S = \{j \in \mathbb{Z} : |\widehat{f}(j)|^2 \ge \delta^4/32\} \quad \text{and} \quad S' = S \cup aS \cup bS.$$

Since $\|f\|_2 \le 1$, we have $|S| \le 32/\delta^4$ and $|S'| \le 96/\delta^4$. If $t \in A(S', \delta^2/(4\pi\sqrt{2}))$ (see (4.3)) and $c = 1$, a or b we have

$$\begin{aligned} |||f - f_{ct}|||^4 &= \sum_{j\in\mathbb{Z}} |\widehat{f}(j) - e^{-2\pi ijct}\widehat{f}(j)|^4 = \sum_{j\in\mathbb{Z}} |\widehat{f}(j)|^4\,|1 - e^{2\pi ijct}|^4 \\ &\le 2^4|S|\left(\frac{\delta^4}{32}\right)^2 + 2^4\,\frac{\delta^4}{32}\sum_{j\notin S}|\widehat{f}(j)|^2 \le \frac{\delta^4}{2} + \frac{\delta^4}{2} = \delta^4. \end{aligned}$$

So $A(S', \delta^2/(4\pi\sqrt{2})) \subset B(\delta)$ and, by Lemma 4.3, we can find a kernel k vanishing outside $B(\delta)$ with

$$\|k\|_\infty \le (4\pi\sqrt{2}/\delta^2)^{96/\delta^4} = 1/\nu. \ \blacksquare$$

Proof of Theorem 4.1. Let $\varepsilon > 0$. We define sequences $\{\delta_n\}_{n\geq 0}$ and $\{K_n\}_{n\geq 0}$ by $\delta_0 = 1$ and, for $n \geq 0$,

$$K_n = \nu(\delta_n)^{-1/4}, \qquad \delta_{n+1} = \min\left\{\delta_n,\, \frac{\varepsilon^6}{100K_n},\, \frac{\varepsilon^3}{20}\,\nu\left(\frac{\varepsilon^3}{10K_n}\right)\right\}.$$

Let M be the smallest integer $\geq 50\,\varepsilon^{-6}$ and set $\delta = \delta_M$ (δ has been identified explicitly). Suppose now that $a, b \in \mathbb{Z}$ and $f : \mathbb{T} \to [0,1]$ with $\int f(x)\,dx \geq \varepsilon$. We must show that

$$J(f) = \iint_{\mathbb{T}^2} f(x)f(x+at)f(x+bt)\,dx\,dt \geq \delta.$$

For $n \geq 0$, choose a kernel k_n such that

$$k_n \geq 0, \qquad \widehat{k}_n \geq 0 \quad \text{and} \quad \int_{\mathbb{T}} k_n(x)\,dx = 1,$$

k_n vanishes outside $B(\delta_n)$ (as defined in (4.4) and (4.5)), and $\|k_n\|_\infty \leq 1/\nu(\delta_n)$ (such a k_n exists by Lemma 4.4), and put

$$f_{(n)} = f * k_n.$$

By Young's inequality,

$$\|f_{(n)}\|_p \leq \|f\|_p\|k_n\|_1 \leq 1 \quad \text{ for all } p \geq 1.$$

Meanwhile since

$$\begin{aligned}(f - f_{(n)})(x) &= f(x) - f * k_n(x) = f(x) - \int f(x-t)k_n(t)\,dt \\ &= f(x) - \int f(x+t)k_n(-t)\,dt = \int (f - f_t)(x)k_n(-t)\,dt\end{aligned}$$

we have $f - f_{(n)} = \int (f - f_t)\,d\mu(t)$, where $d\mu = \check{h}d\lambda$. Since $|||\cdot|||$ is convex and $\check{k}_n$ vanishes outside $B(\delta_n)$, Jensen's inequality gives us

$$\begin{aligned}|||f - f_{(n)}||| &\leq \int_{\mathbb{T}} |||f - f_t|||\,d\mu(t) \leq \int_{\mathbb{T}} d(t)\,d\mu(t) \\ &= \int_{\mathbb{T}} d(t)\check{k}_n(t)\,dt \leq \delta_n.\end{aligned} \tag{4.6}$$

Also, since by the Hausdorff–Young inequality $\widehat{k}_n(j) \leq \|\widehat{k}_n\|_\infty \leq \|k_n\|_1 = 1$ for all $j \in \mathbb{Z}$,

$$\begin{aligned}|||k_n||| &= \Big(\sum_{j\in\mathbb{Z}} |\widehat{k}_n(j)|^4\Big)^{1/4} \leq \Big(\sum_{j\in\mathbb{Z}} |\widehat{k}_n(j)|\Big)^{1/4} = \|\widehat{k}_n\|_1^{1/4} \\ &\leq \nu(\delta_n)^{-1/4} = K_n.\end{aligned} \tag{4.7}$$

Now

$$
\begin{aligned}
&\|f_{(n+1)} - f_{(n)}\|_2^2 \\
&\quad = \|\widehat{f}_{(n+1)} - \widehat{f}_{(n)}\|_2^2 \\
&\quad = \|\widehat{f}_{(n+1)}\|_2^2 + \|\widehat{f}_{(n)}\|_2^2 - \sum_{j\in\mathbb{Z}}(-\widehat{f}_{(n+1)}(j)\overline{\widehat{f}}_{(n)}(j) + \overline{\widehat{f}}_{(n)}(j)\widehat{f}_{(n+1)}(j)) \\
&\quad = \|\widehat{f}_{(n+1)}\|_2^2 - \|\widehat{f}_{(n)}\|_2^2 + 2\sum_{j\in\mathbb{Z}}|\widehat{f}_{(n)}(j)|^2 - 2\sum_{j\in\mathbb{Z}}|\widehat{f}(j)|^2\widehat{k}_n(j)\widehat{k}_{n+1}(j) \\
&\quad = \|f_{(n+1)}\|_2^2 - \|f_{(n)}\|_2^2 + 2\sum_{j\in\mathbb{Z}}|\widehat{f}(j)|^2\,\widehat{k}_n(j)(\widehat{k}_n(j) - \widehat{k}_{n+1}(j)) \\
&\quad \le \|f_{(n+1)}\|_2^2 - \|f_{(n)}\|_2^2 + 2\sum_{j\in\mathbb{Z}}|\widehat{f}(j)|^2\widehat{k}_n(j)(1 - \widehat{k}_{n+1}(j)) \\
&\quad \le \|f_{(n+1)}\|_2^2 - \|f_{(n)}\|_2^2 + 2\|\widehat{f}\,\widehat{k}_n\|_2\|\widehat{f} - \widehat{f}\,\widehat{k}_{n+1}\|_2 \\
&\quad \le \|f_{(n+1)}\|_2^2 - \|f_{(n)}\|_2^2 + 2|||f|||^2|||k_n|||\cdot|||f - f_{(n+1)}|||,
\end{aligned}
$$

upon applying twice the Cauchy–Schwarz inequality. By (4.6) and (4.7) and the fact that $|||f|||^2 \le \|f\|_2 \le 1$, the last term is $\le 2K_n\delta_{n+1}$ and according to the definition of δ_{n+1} we get

$$\|f_{(n+1)} - f_{(n)}\|_2^2 \le \|f_{(n+1)}\|_2^2 - \|f_{(n)}\|_2^2 + \varepsilon^6/50.$$

It follows that

$$
\begin{aligned}
\sum_{n=0}^{M-1}\|f_{(n+1)} - f_{(n)}\|_2^2 &\le \|f_{(M)}\|_2^2 - \|f_{(0)}\|_2^2 + M\varepsilon^6/50 \\
&\le 1 + M\varepsilon^6/50 \le M\varepsilon^6/25
\end{aligned}
$$

and therefore we can choose some $n < M$ such that

$$\|f_{(n+1)} - f_{(n)}\|_2 \le \varepsilon^3/5.$$

Next, consider the expression

$$I_n(t) = \int_{\mathbb{T}} f_{(n)}(x)f_{(n)}(x+at)f_{(n)}(x+bt)\,dx, \qquad t\in\mathbb{T}.$$

Decomposing as in the proof of Lemma 4.2 and employing Cauchy–Schwarz, we get

$$|I_n(t) - I_{n+1}(t)| \le 3\|f_{(n)} - f_{(n+1)}\|_2 \le 3\varepsilon^3/5. \tag{4.8}$$

On the other hand,

$$
\begin{aligned}
&|I_n(0) - I_n(t)| \\
&\quad = \Big| \int_{\mathbb{T}} (f_{(n)}(x))^3 - f_{(n)}(x) f_{(n)}(x+at) f_{(n)}(x+bt)\, dx \Big| \\
&\quad = \Big| \int_{\mathbb{T}} (f_{(n)}(x))^2 (f_{(n)}(x) - f_{(n)}(x+at)) \\
&\qquad + f_{(n)}(x) f_{(n)}(x+at)(f_{(n)}(x) - f_{(n)}(x+bt))\, dx \Big| \\
&\quad \le \|f_{(n)}^2\|_2 \|f_{(n)} - (f_{(n)})_{at}\|_2 + \|f_{(n)}(f_{(n)})_{at}\|_2 \|f_{(n)} - (f_{(n)})_{bt}\|_2 \\
&\quad \le \|f_{(n)}\|_4^2 \|f_{(n)} - (f_{(n)})_{at}\|_2 + \|f_{(n)}\|_4 \|(f_{(n)})_{at}\|_4 \|f_{(n)} - (f_{(n)})_{bt}\|_2 \\
&\quad \le \|f_{(n)} - (f_{(n)})_{at}\|_2 + \|f_{(n)} - (f_{(n)})_{bt}\|_2 \\
&\quad = \|(f - f_{at}) * k_n\|_2 + \|(f - f_{bt}) * k_n\|_2 \\
&\quad \le |||k_n|||(|||f - f_{at}||| + |||f - f_{bt}|||) \le 2K_n\, d(t).
\end{aligned}
$$

As $I_n(0) = \int f_{(n)}(x)^3\, dx \ge \|f_{(n)}\|_1^3 = \|f\|_1^3 \ge \varepsilon^3$ (Jensen's inequality), it follows that $I_n(t) \ge \varepsilon^3 - 2K_n d(t)$ and by (4.8) we get

$$I_{n+1}(t) \ge 2\varepsilon^3/5 - 2K_n\, d(t) \quad \text{for every } t \in \mathbb{T}.$$

So, $I_{n+1}(t) \ge \varepsilon^3/5$ on the set $B(\varepsilon^3/(10K_n))$ whence, by Lemma 4.4 and the choice of δ_{n+1},

$$J(f_{(n+1)}) = \int_{\mathbb{T}} I_n(x)\, dx \ge \frac{\varepsilon^3}{5}\, \lambda\bigg(B\bigg(\frac{\varepsilon^3}{10K_n}\bigg)\bigg) \ge \frac{\varepsilon^3}{5}\, \nu\bigg(\frac{\varepsilon^3}{10K_n}\bigg) \ge 4\delta_{n+1}.$$

Finally, applying Lemma 4.2 and (4.6) we get

$$J(f) \ge J(f_{(n+1)}) - 3\delta_{n+1} \ge \delta_{n+1} \ge \delta_M$$

and the proof is complete. ∎

Disappointingly, we see no way to preserve the uniformity of Theorem 4.1 over all a and b in the following corollaries.

COROLLARY 4.5. *Let $a, b \in \mathbb{N}$. For any $\varepsilon > 0$ there exists $M = M(\varepsilon, a, b) \in \mathbb{N}$ (which may be explicitly identified) such that every set $E \subset \{1, \ldots, M\}$ satisfying $|E| \ge \varepsilon M$ contains a configuration of the form $\{x, x+an, x+bn\}$ for some x and some positive integer n.*

Proof. We may assume that a and b are relatively prime. First we will prove the result substituting the weaker conclusion $n \ne 0$, then show that n can in fact be chosen positive. Let $\delta = \delta(\varepsilon/(4a^2b^2))$ as in Theorem 4.1 and choose N with $N > 1/(4a^2b^2\delta)$. Suppose that $E \subset \{1, \ldots, N\}$ with

$|E| \geq \varepsilon N$. Let

$$A = \frac{1}{2Nab} E + \left(0, \frac{1}{4Na^2b^2}\right).$$

Then $m(A) \geq \varepsilon/(4a^2b^2)$ so by Theorem 4.1 there exists $x, h \in \mathbb{T} \setminus (-\delta/2, \delta/2)$ with $\{x, x + ah, x + bh\} \in A$. This implies that ah and bh both lie in $(-1/(2ab), 1/(2ab)) \subset \mathbb{T}$. In other words, there exist integers k_1 and k_2 such that

$$|ah - k_1| < \frac{1}{2ab} \quad \text{and} \quad |bh - k_2| < \frac{1}{2ab}.$$

In particular, $|bk_1 - ak_2| < 1$, so that $bk_1 = ak_2$, implying $a \mid k_1$ and $b \mid k_2$, which gives us

$$\left|h - \frac{k_1}{a}\right| < \frac{1}{2a^b}, \quad \text{so that} \quad h \in \left(-\frac{1}{2ab}, \frac{1}{2ab}\right).$$

We may write

$$x = \frac{n_1}{2Nab} + \alpha_1, \quad x + ah = \frac{n_2}{2Nab} + \alpha_2, \quad x + bh = \frac{n_3}{2Nab} + \alpha_3,$$

where $\{n_1, n_2, n_3\} \subset E$ and $0 < \alpha_1, \alpha_2, \alpha_3 < 1/(4Na^2b^2)$. Solving for h in the latter two expressions and setting the resulting quantities equal gives

$$(n_2 - n_1)b + 2Nab^2(\alpha_2 - \alpha_1) = (n_3 - n_1)a + 2Na^2b(\alpha_3 - \alpha_1),$$

which implies $(n_2 - n_1)b = (n_3 - n_1)a$ since the quantities $2Nab^2(\alpha_2 - \alpha_1)$ and $2Na^2b(\alpha_3 - \alpha_1)$ are less than $1/2$ in absolute value. In other words, the set $\{n_1, n_2, n_3\}$ has the form $\{x, x + an, x + bn\}$, and $n \neq 0$ since $h > 1/(4Na^2b^2)$.

This proves the result under the relaxed conclusion $n \neq 0$. Let now $M = N(\varepsilon^2/6)$ and suppose that $E \subset \{1, \ldots, N\}$ with $|E| \geq \varepsilon N$. For some l with $1 \leq l \leq 2N$ we have $|(E + l) \cap (2N - E)| > (\varepsilon^2/2)N = (\varepsilon^2/6)(3N)$, hence there exists a configuration $\{x, x + an, x + bn\}$ in $(F + l) \cap (2N - F)$, where $n \neq 0$. If $n > 0$ we have $\{x + l, (x + l) + an, (x - l) + bn\} \subset E$ and we are done. Otherwise, $\{2N - x, (2N - x) + a(-n), (2N - x) + b(-n)\} \subset E$. ■

The derivation of the following from Corollary 4.5 is analogous to the proof of the implication (i)⇒(ii) in Theorem 2.1.

Corollary 4.6. *Suppose $a, b \in \mathbb{N}$ and $\varepsilon > 0$. There exist $N = N(a, b, \varepsilon) \in \mathbb{N}$ and $\beta = \beta(a, b, \varepsilon) > 0$ (both of which may be explicitly identified) such that for every measure preserving system $(X, \mathcal{B}, \mu, T)$ with $\mu(X) = 1$ and every $A \in \mathcal{B}$ with $\mu(A) > \varepsilon$ there exists a positive integer $n \leq N$ such that*

$$\mu(A \cap T^{-an} A \cap T^{-bn} A) \geq \beta.$$

As we mentioned earlier, we are unable to carry over the uniformity with respect to a and b from Theorem 4.1 to Corollaries 4.5 and 4.6. In a sense, this is unsurprising, as Forrest has shown in [Fo] that any estimates that are good for arbitrary powers of the same transformation T are also good for two commuting transformations T and S, and we certainly do not feel that these methods are sufficient to handle such cases without substantial modification.

Gowers mentions in [G2] that he suspects his methods can be adapted to the case of commuting transformations, thus providing explicit bounds for those cases. However, nothing along these lines has been published yet, so for the moment the question of identifying uniform (over all a and b) bounds in these two corollaries appears to be open.

5. Counterexamples and miscellanies. Recall from the introduction that if R is a set of recurrence in $\mathbb{N}$ then we have uniformity: for every $\varepsilon > 0$ there exists a finite subset $R' \subset R$ such that for every probability measure preserving system $(X, \mathcal{B}, \mu, T)$ and every $A \in \mathcal{B}$ with $\mu(A) \geq \varepsilon$ there exists $n \in R'$ such that $\mu(A \cap T^{-n}A) > 0$.

We now seek to show that this result does not carry over to sets of recurrence in $\mathbb{R}$. Namely, we show that there exist sets of recurrence $R \subset \mathbb{R}$ such that for any $\varepsilon > 0$, any finite subset $R' \subset R$, and for any aperiodic measurable measure preserving flow $(X, \mathcal{B}, \mu, \{T_t\}_{t\in\mathbb{R}})$ with $\mu(X) = 1$, there exists a set $A \in \mathcal{B}$ with $\mu(A) > 1/2 - \varepsilon$ such that $\mu(A \cap T_t A) = 0$ for all $t \in R'$. The following lemma, whose proof is modelled after the proof of [BBB], Theorem D, is the key.

LEMMA 5.1. *Given $\nu > 0$, a rationally independent set $\{\lambda_1, \ldots, \lambda_N\} \subset \mathbb{R}$, and any aperiodic measurable flow $(X, \mathcal{B}, \mu, \{T_t\}_{t\in\mathbb{R}})$ with $\mu(X) = 1$, there exists a set $A \in \mathcal{B}$ with $\mu(A) > 1/2 - \nu$ such that $\mu(A \cap T_{\lambda_i} A) = 0$, $1 \leq i \leq N$.*

Proof. Let $\varepsilon > 0$ be so small that $(1-\varepsilon)(1/2-\varepsilon) > 1/2 - \nu$. Scaling if necessary by a constant we may assume that $\{1, \lambda_1, \ldots, \lambda_N\}$ is rationally independent. In this case it is well known that the set

$$\{(n\lambda_1, \ldots, n\lambda_N) : n \in \mathbb{N}\}$$

(multiplication modulo 1) is dense on the N-torus $\mathbb{T}^N$. Therefore there exists $n \in \mathbb{N}$ such that $n\lambda_i \in (1/2-\varepsilon, 1/2+\varepsilon) \pmod 1$, $1 \leq i \leq N$. By [L], Theorem 1, there exists some $L \in \mathbb{N}$, some number a with $1 - \varepsilon \leq a \leq 1$, and a set $B \in \mathcal{B}$ with the following properties:

(i) for all t, s with $0 \leq t < s \leq L$, $T_t B \cap T_s B = \emptyset$,

(ii) for any Lebesgue measurable set $D \subset [0, L]$, $\widetilde{D} = \bigcup_{t\in D} T_t B \in \mathcal{B}$, and moreover $\mu(\widetilde{D}) = am(D)/L$, where m is Lebesgue measure.

Let $S = \{j/n+\beta : j \in \{0, 1, \ldots, Ln-1\}, 0 \leq \beta \leq (1/n)(1/2-\varepsilon)\} \subset [0, L]$. Then $m(S) = L(1/2 - \varepsilon)$ and $S \cap (S + \lambda_i) = \emptyset$, $1 \leq i \leq N$. Let $A = \widetilde{S}$ (as in (ii) above). Then $\mu(A) = a(1/2-\varepsilon) > 1/2-\nu$ and $\mu(A\cap T_{\lambda_i}A) = 0$, $1\le i\le N$. ∎

We are now ready to give the example of the set of recurrence $R\subset\mathbb{R}$ with the specified non-uniformity properties. The set R will be of the form $R=\{n^\alpha : n\in\mathbb{N}\}$. Such sets are sets of recurrence for continuous actions in $\mathbb{R}$ for $\alpha\ne 0$. (For a proof of this fact, see [BBB, Section 3].) For many α, we have the sought-after non-uniformity as well:

THEOREM 5.2. *For all but countably many* $\alpha\in\mathbb{R}$, *the set* $R_\alpha=\{n^\alpha : n\in\mathbb{N}\}$ *has the property that for any* $\varepsilon>0$, *any finite subset* $R'\subset R_\alpha$, *and any aperiodic measurable flow* $(X,\mathcal{B},\mu,\{T_t\}_{t\in\mathbb{R}})$, *where* $\mu(X)=1$, *there exists a set* $A\in\mathcal{B}$ *with* $\mu(A)>1/2-\varepsilon$ *such that* $\mu(A\cap T_tA)=0$ *for all* $t\in R'$.

Proof. By [BBB], Lemma 2.9, R_α is rationally independent for all but countably many $\alpha\in\mathbb{R}$. The result follows from Lemma 5.1. ∎

We now give another negative result. In [F], what is actually proved is that for any $l\in\mathbb{N}$, any probability measure preserving system $(X,\mathcal{B},\mu,T)$, and any $A\in\mathcal{B}$ with $\mu(A)>0$ one has

$$\liminf_{N-M\to\infty}\frac{1}{N-M}\sum_{n=M+1}^{N}\mu(A\cap T^{-n}A\cap\ldots\cap T^{-(l-1)n}A)>0.$$

We now show that it is impossible to extend this result in the manner of Theorem 2.1(iii). Indeed, we have the following:

THEOREM 5.3. *For any invertible ergodic aperiodic probability measure preserving system* $(X,\mathcal{B},\mu,T)$, *any* $\varepsilon>0$, *and any* $N\in\mathbb{N}$ *there exists a set* $A\in\mathcal{B}$ *with* $\mu(A)>1/2-\varepsilon$ *and some* $M\in\mathbb{N}$ *having the property that for all* n *with* $M+1\le n\le M+N$, *one has* $\mu(A\cap T^{-n}A)=0$.

Proof. By Rokhlin's Theorem there exists $k\in\mathbb{N}$ large enough that $k>N/\varepsilon$ and

$$\frac{1-\varepsilon}{2}\cdot\frac{k-N}{k+1}>\frac{1}{2}-\varepsilon,$$

and a set $B\in\mathcal{B}$ such that $T^iB\cap T^jB=\emptyset$, $0\le i<j\le k$, and $\mu(\bigcup_{i=0}^{k}T^iB) > 1-\varepsilon$ (which implies $\mu(B)>(1-\varepsilon)/(k+1)$). Let $M=[(k-N)/2]$ and set $A=\bigcup_{i=0}^{M}T^iB$. Then

$$\mu(A)\ge\frac{1-\varepsilon}{k+1}\cdot\frac{k-N}{2}>\frac{1}{2}-\varepsilon.$$

One may also check that $A\cap T^nA=\emptyset$, $M+1\le n\le k-M$, and since $M\le(k-N)/2$ we have $k-M\ge M+N$. ∎

6. Appendix

Proof of Theorem 3.2. Suppose not. Then there exist $\varepsilon > 0$, $k, l \in \mathbb{N}$, and polynomials $p_{i,j}(n) \in \mathbb{Z}[n]$ with $p_{i,j}(0) = 0$, $1 \le i \le k$, $1 \le j \le l$ having the property that for every $N \in \mathbb{N}$ there exists a set $S_N \subset \{1, \ldots, N\}^l$ with $|S_N| \ge \varepsilon N^l$ such that, if we set $p_{0,j}(n) = 0$, $1 \le j \le l$, then S_N contains no configuration of the form

$$\{(u_1 + p_{i,1}(n), \ldots, u_l + p_{i,l}(n)) : 0 \le i \le k\}, \tag{3.1}$$

for any n, $1 \le n \le N$. We may assume that $p_{1,j}(n) = n$, $1 \le j \le l$, and that $k \ge 2$. For convenience, we will use the following notation: $\mathbf{p}_i(n) = (p_{i,1}(n), \ldots, p_{i,l}(n))$, $0 \le i \le k$. We may assume as well that $\mathbf{p}_i \ne \mathbf{p}_j$ whenever $0 \le i \ne j \le k$.

We will now construct a set $E \subset \mathbb{Z}^l$ with $d^*(E) \ge \varepsilon$ such that E contains no configuration of the form (3.1), thus obtaining a contradiction to Theorem 3.1. The set E will be a countable union of larger and larger finite sets each of which contains no configuration of the type in question and each of which is of density at least ε in some l-dimensional cube. The cubes containing these sets, on the other hand, are shifted (by l-tuples $\mathbf{u}_i$ determined presently) so as to be so sparsely distributed that no applicable configuration can lie in their union without being wholly contained in a single one of the cubes (this is the content of (i) and (ii) below).

Let then $\mathbf{u}_1$ be arbitrary. Having chosen $\mathbf{u}_1, \ldots, \mathbf{u}_{N-1}$, let

$$F_{N-1} = \bigcup_{r=1}^{N-1} (\mathbf{u}_r + \{1, \ldots, r\}^l).$$

We seek to choose $\mathbf{u}_N$ having the following properties:

(i) if there exist $\mathbf{u} \in \mathbb{Z}^l$, $n \in \mathbb{Z}$, and $0 \le i \ne j \le k$ such that $\{\mathbf{u} + \mathbf{p}_i(n), \mathbf{u} + \mathbf{p}_j(n)\} \subset F_{N-1}$ then $\mathbf{u} + \mathbf{p}_m(n) \notin \mathbf{u}_N + \{1, \ldots, N\}^l$ for all m, $0 \le m \le k$,

(ii) if there exist $\mathbf{u} \in \mathbb{Z}^l$, $n \in \mathbb{Z}$, and $0 \le i \ne j \le k$ such that $\{\mathbf{u} + \mathbf{p}_i(n), \mathbf{u} + \mathbf{p}_j(n)\} \subset \mathbf{u}_N + \{1, \ldots, N\}^l$ then $\mathbf{u} + \mathbf{p}_m(n) \notin F_{N-1}$ for all m, $0 \le m \le k$.

Let us now see how to ensure that (i) is satisfied. Let

$$T = \{n : \mathbf{p}_i(n) - \mathbf{p}_j(n) \in F_{N-1} - F_{N-1} \text{ for some } 0 \le i \ne j \le k\}.$$

Then T is a finite set. Suppose that

$$\mathbf{u}_N \notin \bigcup_{0 \le i,m \le k} (F_{N-1} - \{1, \ldots, N\}^l + (\mathbf{p}_m - \mathbf{p}_i)(T)). \tag{3.2}$$

Then suppose that $\mathbf{u}+\mathbf{p}_i(n) \in F_{N-1}$ and $\mathbf{u}+\mathbf{p}_j(n) \in F_{N-1}$ for some $\mathbf{u} \in \mathbb{Z}^l$, $n \in \mathbb{Z}$, and $0 \le i \ne j \le k$. Then $\mathbf{p}_i(n) - \mathbf{p}_j(n) \in F_{N-1} - F_{N-1}$, so $n \in T$. If now $\mathbf{u} + \mathbf{p}_m(n) \in \mathbf{u}_N + \{1, \ldots, N\}^l$ for some m, $0 \leq m \leq k$, then

$$(\mathbf{u} + \mathbf{p}_m(n)) - (\mathbf{u} + \mathbf{p}_i(n)) \in \mathbf{u}_N + \{1, \ldots, N\}^l - F_{N-1},$$

which implies that

$$\mathbf{u}_N \in F_{N-1} - \{1, \ldots, N\}^l + (\mathbf{p}_m - \mathbf{p}_i)(T),$$

contradicting (3.2). Hence if (3.2) is satisfied then (i) is satisfied.

Now we see how to ensure that (ii) is satisfied. Let

$$U = \{n : \mathbf{p}_i(n) - \mathbf{p}_j(n) \in \{-N, -N+1, \ldots, N\}^l \text{ for some } 0 \leq i \neq j \leq k\}.$$

Then U is a finite set. Suppose that

$$\mathbf{u}_N \notin \bigcup_{0 \leq i, m \leq k} (F_{N-1} - \{1, \ldots, N\}^l + (\mathbf{p}_i - \mathbf{p}_m)(U)). \tag{3.3}$$

Then suppose that $\mathbf{u} + \mathbf{p}_i(n) \in \mathbf{u}_N + \{1, \ldots, N\}^l$ and $\mathbf{u} + \mathbf{p}_j(n) \in \mathbf{u}_N + \{1, \ldots, N\}^l$ for some $\mathbf{u} \in \mathbb{Z}^l$, $n \in \mathbb{Z}$, and $0 \leq i \neq j \leq k$. Suppose that $\mathbf{u} + \mathbf{p}_m(n) \in F_{N-1}$ for some m, $0 \leq m \leq k$. Then

$$\mathbf{p}_i(n) - \mathbf{p}_j(n) = (\mathbf{u} + \mathbf{p}_i(n)) - (\mathbf{u} + \mathbf{p}_j(n)) \in \{-N, -N+1, \ldots, N\}^l,$$

so $n \in U$. Furthermore,

$$(\mathbf{u} + \mathbf{p}_i(n)) - (\mathbf{u} + \mathbf{p}_m(n)) \in \mathbf{u}_N + \{1, \ldots, N\}^l - F_{N-1},$$

which implies that

$$\mathbf{u}_N \in F_{N-1} - \{1, \ldots, N\}^l + (\mathbf{p}_i - \mathbf{p}_m)(U),$$

contradicting (3.3). Hence if (3.3) is satisfied, (ii) is satisfied.

Hence, we need only choose $\mathbf{u}_N$ such that (3.2) and (3.3) are satisfied, which is obviously possible since the sets appearing in these displays are finite. This establishes that we may find a sequence $\{\mathbf{u}_1, \mathbf{u}_2, \ldots\}$ having properties (i) and (ii).

Now let

$$E = \bigcup_{N=1}^{\infty} (\mathbf{u}_N + S_N).$$

Since

$$\frac{|(\mathbf{u}_N + S_N) \cap (\mathbf{u}_N + \{1, \ldots, N\}^l)|}{N^l} = \frac{|S_N \cap \{1, , \ldots, N\}^l|}{N^l} \geq \varepsilon$$

for all $N \in \mathbb{N}$, we have $d^*(E) \geq \varepsilon$. Therefore, by Theorem 3.1 there exist $(u_1, \ldots, u_l) \in \mathbb{Z}^l$ and $n \in \mathbb{N}$ such that

$$(u_1 + p_{i,1}(n), \ldots, u_l + p_{i,l}(n)) \in E, \quad 0 \leq i \leq k.$$

Let $N \in \mathbb{N}$ be minimal with respect to the property that

$$(u_1 + p_{i,1}(n), \ldots, u_l + p_{i,l}(n)) \in \bigcup_{i=1}^{N} (\mathbf{u}_i + S_i), \quad 0 \leq i \leq k.$$

Write $\mathbf{u} = (u_1, \ldots, u_l)$. By minimality of N, $\mathbf{u} + \mathbf{p}_m(n) \in \mathbf{u}_N + S_N \subset \mathbf{u}_N + \{1, \ldots, N\}^l$ for some m, $0 \le m \le k$. By (i) and the fact that $S_i \subset \{1, \ldots, i\}^l$, $1 \le i \le N-1$, we have

$$\mathbf{u} + \mathbf{p}_j(n) \in \bigcup_{i=1}^{N-1} (\mathbf{u}_i + S_i) \subset F_{N-1}$$

for at most one j, $0 \le j \le k$. It follows that, since $k \ge 2$, $\mathbf{u} + \mathbf{p}_j(n) \in \mathbf{u}_N + S_N \subset \mathbf{u}_N + \{1, \ldots, N\}^l$ for some $j \ne m$, $0 \le j \le k$. Hence by (ii),

$$\mathbf{u} + \mathbf{p}_r(n) \notin \bigcup_{i=1}^{N-1} (\mathbf{u}_i + S_i) \subset F_{N-1}, \quad 0 \le r \le k.$$

Therefore $\mathbf{u} + \mathbf{p}_i(n) \in \mathbf{u}_N + S_N$, $0 \le i \le k$, which implies that

$$\{(\mathbf{u} - \mathbf{u}_N) + \mathbf{p}_i(n) : 0 \le i \le k\} \subset S_N.$$

Furthermore, since $S_N \subset \{1, \ldots, N\}^l$, $p_{0,1}(n) = 0$, and $p_{1,1}(n) = n$, we have $1 \le n \le N$, contradicting the fact that S_N contains no such configuration and completing the proof. ■

Proof of Theorem 3.7. Let $\nu > 0$ be so small that for any $g \in L^\infty(\mathbb{R}^l)$ with $0 \le g \le 1$ and $\int_{[0,1]^l} g\, d\lambda \ge \varepsilon$ we have

$$\lambda\left(\left\{\mathbf{u} \in \left[\frac{\nu}{4l}, 1 - \frac{\nu}{4l}\right]^l : g(\mathbf{u}) \ge \nu\right\}\right) \ge \nu.$$

Let $a = klt$ and denote by $\mathbf{e}_j$ the jth coordinate unit vector in $\mathbb{R}^a$, $1 \le j \le a$. By Theorem 3.2 there exists $L \in \mathbb{N}$ having the property that if $E \subset \{1, \ldots, L\}^a$ with $|E| \ge \nu L^a/4$ then E contains a configuration of the form

$$\{\mathbf{s} + \mathbf{u}_i(n) : 1 \le i \le k\},$$

where $1 \le n \le L$ and

$$\mathbf{u}_i(n) = \sum_{j=1}^{l} \sum_{b=1}^{t} n^b \mathbf{e}_{b+(j-1)t+(i-1)tl}.$$

Let $\alpha_{i,j,b}$ be the coefficient of x^b in $p_{i,j}(x)$, $1 \le b \le t$, $1 \le i \le k$, $1 \le j \le l$, and put

$$V = 2 \max_{1 \le j \le l} \sum_{i=1}^{k} \sum_{b=1}^{t} |\alpha_{i,j,b}|.$$

Set $\delta = \nu^{k+2}/(8lVL^{a+2})$.

Suppose now that N is any real number with $N \ge 1$. Let M be any number which is not the root of any polynomial having coefficients in

$$\mathbb{Q}[\{N^{b/t}, \alpha_{i,j,b} : 1 \le i \le k,\ 1 \le j \le l,\ 1 \le b \le t\}]$$

and for which $1/M \le \nu/(2lV)$.

Unlike in Theorems 3.4 and 3.6 our intention is not to fix M, with some specific lower bound on $1/M$; rather we wish to point out that the steps in the argument below may be carried out for all but countably many $M \geq 2lV/\nu$. Each of these M's will supply an h, much as in the previous proofs. (However, note that the role of h below is somewhat different than in the previous arguments; actually nh serves the function that h did before.) The proof is completed by noting the the combined measure of these h's, which come from the various M's, is large enough to give the desired conclusion.

Consider the map $\gamma : \{1, \ldots, L\}^a \to \mathbb{R}^l$ given by

$$\gamma(s_1, \ldots, s_a) = \sum_{i=1}^{k} \sum_{j=1}^{l} \sum_{b=1}^{t} s_{b+(j-1)t+(i-1)tl} \left(\frac{\alpha_{i,j,b}}{L^b M^b N^{1-b/t}} \right) \mathbf{c}_j,$$

where $\mathbf{c}_j$ is the jth coordinate unit vector in $\mathbb{R}^l$, $1 \leq j \leq l$. Let

$$A = \left(\frac{\nu}{4l}, \ldots, \frac{\nu}{4l} \right) + \gamma(\{1, 2, \ldots, L\}^a) \subset \mathbf{R}^l.$$

(Compare with the map f of Theorem 3.6.) One easily checks that the range of γ is contained in $[-\nu/(4l), \nu/(4l)]^l$. In particular, $A \subset [0, \nu/(2l)]^l$. Furthermore, since M is not a root of any polynomial having coefficients in

$$\mathbb{Q}[\{N^{b/t}, \alpha_{i,j,b} : 1 \leq i \leq k,\ 1 \leq j \leq l,\ 1 \leq b \leq t\}],$$

we have $|A| = L^a$.

Assume now that $f \in L^\infty(\mathbb{R}^l)$ with $\int_{[0,N]^l} f \, d\lambda \geq \varepsilon N^l$. Let $\widetilde{f}(\mathbf{u}) = f(N\mathbf{u})$, $\mathbf{u} \in \mathbb{R}^l$. Then $\int_{[0,1]^l} \widetilde{f} \, d\lambda \geq \varepsilon$. Recall the property whereby ν was chosen. Namely, if we now let

$$S = \left\{ \mathbf{u} \in \left[\frac{\nu}{4l}, 1 - \frac{\nu}{4l} \right]^l : \widetilde{f}(\mathbf{u}) \geq \nu \right\},$$

then $\lambda(S) \geq \nu$. By Lemma 3.5, if we let

$$U_M = \{x \in [0,1]^l : |(x + A) \cap S|/|A| \geq \nu/4\},$$

we have $\lambda(U_M) \geq \nu/4$. (Notice as well that $U_M \subset [\nu/(4l), 1 - \nu/(4l)]^l$.) For $x \in U_M$, let

$$E_x = \{(s_1, \ldots, s_a) \in \{1, \ldots, L\}^a : x + \gamma(s_1, \ldots, s_a) \in S\}.$$

Then $|E_x| \geq (\nu/4)L^a$, so E_x contains a configuration of the form

$$\{\mathbf{s} + \mathbf{u}_i(n) : 1 \leq i \leq k\},$$

where $1 \leq n \leq L$ and

$$\mathbf{u}_i(n) = \sum_{j=1}^{l} \sum_{b=1}^{t} n^b \mathbf{e}_{b+(j-1)t+(i-1)tl} g.$$

This in turn implies that

$$\{x+\gamma(\mathbf{s}+\mathbf{u}_i(n)) : 1\le i\le k\}\subset S.$$

Note that γ is a linear function. In particular,

$$\begin{aligned}
&x+\gamma(\mathbf{s}+\mathbf{u}_i(n))\\
&\quad = x+\gamma(\mathbf{s})+\gamma(\mathbf{u}_i(n))\\
&\quad = x+\gamma(\mathbf{s})+\sum_{j=1}^{l}\sum_{b=1}^{t}\gamma(n^b\mathbf{e}_{b+(j-1)t+(i-1)tl})\\
&\quad = x+\gamma(\mathbf{s})+\sum_{j=1}^{l}\bigg(\sum_{b=1}^{t}n^b\Big(\frac{\alpha_{i,j,b}}{L^bM^bN^{1-b/t}}\Big)\bigg)\mathbf{c}_j\in S,\qquad 1\le i\le k.
\end{aligned}$$

Let $h=N^{1/t}/(LM)$. Then

$$\begin{aligned}
&(x+\gamma(\mathbf{s}))+\frac{1}{N}\mathbf{p}_i(nh)\\
&\quad = (x+\gamma(\mathbf{s}))+\frac{1}{N}\sum_{j=1}^{l}p_{i,j}(nh)\mathbf{c}_j\\
&\quad = (x+\gamma(\mathbf{s}))+\frac{1}{N}\sum_{j=1}^{l}\Big(\sum_{b=1}^{t}\alpha_{i,j,b}(nh)^b\Big)\mathbf{c}_j\\
&\quad = x+\gamma(\mathbf{s})+\sum_{j=1}^{l}\bigg(\sum_{b=1}^{t}n^b\Big(\frac{\alpha_{i,j,b}}{L^bM^bN^{1-b/t}}\Big)\bigg)\mathbf{c}_j\in S,\qquad 1\le i\le k.
\end{aligned}$$

Therefore,

$$\prod_{i=1}^{k}\widetilde{f}\Big(x+\gamma(\mathbf{s})+\frac{1}{N}\mathbf{p}_i(nh)\Big)\ge\nu^k.$$

At this point, n and $\mathbf{s}$ depend on x. However, recalling that $\mathbf{s}\in\{1,\dots,L\}^a$ and $n\in\{1,\dots,L\}$, we have, for all $x\in U_M$,

$$\sum_{\mathbf{s}\in\{1,\dots,L\}^a}\sum_{n=1}^{L}\prod_{i=1}^{k}\widetilde{f}\Big(x+\gamma(\mathbf{s})+\frac{1}{N}\mathbf{p}_i(nh)\Big)\ge\nu^k,$$

which implies that (recall $U_M\subset[\nu/(4l),1-\nu/(4l)]^l$)

$$\sum_{\mathbf{s}\in\{1,\dots,L\}^a}\sum_{n=1}^{L}\int_{[\nu/(4l),1-\nu/(4l)]^l}\prod_{i=1}^{k}\widetilde{f}\Big(x+\gamma(\mathbf{s})+\frac{1}{N}\mathbf{p}_i(nh)\Big)\,d\lambda(x)\ge\nu^k\lambda(U_M)\ge\frac{\nu^{k+1}}{4}. \tag{3.4}$$

Recall that $h = N^{1/t}/(LM)$, and M can be, with the exception of some countable set, any number for which $1/M \le \nu/(2lV)$. In other words, (3.4) is true for all h, excepting some countable set, for which

$$0 \le h \le \frac{\nu N^{1/t}}{2lVL}.$$

Therefore,

$$\sum_{\mathbf{s}\in\{1,\dots,L\}^a} \sum_{n=1}^{L} \int\limits_{[\nu/(4l),1-\nu/(4l)]^l} \int\limits_0^{\nu N^{1/t}/(2lVL)} \prod_{i=1}^{k} \widetilde{f}\Big(x+\gamma(\mathbf{s})+\frac{1}{N}\mathbf{p}_i(nh)\Big)\,dh\,d\lambda(x) \ge \frac{\nu^{k+2}N^{1/t}}{8lVL}.$$

This in turn implies that for some fixed $\mathbf{s} \in \{1,\dots,L\}^a$ and $n \in \{1,\dots,L\}$,

$$\int\limits_{[\nu/(4l),1-\nu/(4l)]^l} \int\limits_0^{\nu N^{1/t}/(2lVL)} \prod_{i=1}^{k} \widetilde{f}\Big(x+\gamma(\mathbf{s})+\frac{1}{N}\mathbf{p}_i(nh)\Big)\,dh\,d\lambda(x) \ge \frac{\nu^{k+2}N^{1/t}}{8lVL^{a+2}},$$

which is the same as

$$\int\limits_{[\nu/(4l),1-\nu/(4l)]^l} \int\limits_0^{\nu N^{1/t}/(2lVL)} \prod_{i=1}^{k} f(N(x+\gamma(\mathbf{s}))+\mathbf{p}_i(nh))\,dh\,d\lambda(x) \ge \frac{\nu^{k+2}N^{1/t}}{8lVL^{a+2}}.$$

Making the substitution $\mathbf{u} = N(x+\gamma(\mathbf{s}))$ (recall that f is non-negative and $\gamma(\mathbf{s}) \subset [-\nu/(4l),\nu/(4l)]^l$) we have

$$\int\limits_{[0,N]^l} \int\limits_0^{\nu N^{1/t}/(2lVL)} \prod_{i=1}^{k} f(\mathbf{u}+\mathbf{p}_i(nh))\,dh\,d\lambda(\mathbf{u}) \ge \frac{\nu^{k+2}N^{l+1/t}}{8lVL^{a+2}}.$$

Making the substitution $y = nh$, we get

$$\int\limits_{[0,N]^l} \int\limits_0^{\nu n N^{1/t}/(2lVL)} \prod_{i=1}^{k} f(\mathbf{u}+\mathbf{p}_i(y))\,dy\,d\lambda(\mathbf{u}) \ge \frac{n\nu^{k+2}N^{l+1/t}}{8lVL^{a+2}} \ge \frac{\nu^{k+2}N^{l+1/t}}{8lVL^{a+2}},$$

which gives

$$\int\limits_{[0,N]^l} \int\limits_0^{N^{1/t}} f(\mathbf{u})f(\mathbf{u}+\mathbf{p}_1(y)))\dots f(\mathbf{u}+\mathbf{p}_k(y))\,dy\,d\lambda(\mathbf{u}) \ge \delta N^{l+1/t}. \quad \blacksquare$$

REFERENCES

[B] V. Bergelson, *Sets of recurrence of $\mathbb{Z}^m$-actions and properties of sets of differences in $\mathbb{Z}^m$*, J. London Math. Soc. (2) 31 (1985), 295–304.

[BBB] V. Bergelson, M. Boshernitzan, and J. Bourgain, *Some results on non-linear recurrence*, J. Anal. Math. 62 (1994), 29–46.

[BH] V. Bergelson and I. Håland, *Sets of recurrence and generalized polynomials*, in: Convergence in Ergodic Theory and Probability, June 1993, Ohio State University Math. Research Institute Publications, de Gruyter, 1996, 91–110.

[BL] V. Bergelson and A. Leibman, *Polynomial extensions of van der Waerden's and Szemerédi's theorem*, J. Amer. Math. Soc. 9 (1996), 725–753.

[Bo1] J. Bourgain, *A Szemerédi type theorem for sets of positive density in* $\mathbb{R}^k$, Israel J. Math. 54 (1986), 307–316.

[Bo2] —, *A non-linear version of Roth's theorem for sets of positive density in the real line*, J. Anal. Math. 50 (1988), 169–181.

[Fo] A. Forrest, *Recurrence in dynamical systems: a combinatorial approach*, Ph.D. Thesis, Ohio State University, 1990.

[F] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, ibid. 31 (1977), 204–256.

[FK] H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for commuting transformations*, ibid. 34 (1978), 275–291.

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

[G1] W. T. Gowers, *A new proof of Szemerédi's theorem for arithmetic progressions of length four*, Geom. Funct. Anal. 8 (1998), 529–551.

[G2] —, *Fourier analysis and Szemerédi's theorem*, in: Proc. Internat. Congress Math., Vol. I, Doc. Math. (Berlin, 1998), 617–629.

[G3] —, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal., to appear.

[H-B] D. R. Heath-Brown, *Integer sets containing no arithmetic progressions*, J. London Math. Soc. (2) 35 (1987), 385–394.

[La] Y. Lacroix, private communication.

[L] D. Lind, *Locally compact measure preserving flows*, Adv. Math. 15 (1975), 175–193.

[R] K. Roth, *Sur quelques ensembles d'entiers*, C. R. Acad. Sci. Paris 234 (1952), 388–390.

[S1] E. Szemerédi, *On sets of integers containing no* k *elements in arithmetic progression*, Acta Arith. 27 (1975), 199–245.

[S2] —, *Integer sets containing no arithmetic progressions*, Acta Math. Hungar. 56 (1990), 155–158.

Department of Mathematics
The Ohio State University
231 W. 18th Avenue
Columbus, OH 43210, U.S.A.
E-mail: vitaly@math.ohio-state.edu

Équipe d'Analyse et de Mathématiques Appliquées
Université de Marne-la-Vallée
5 Boulevard Descartes, Champs sur Marne
77454 Marne-la-Vallée, Cedex 2, France
E-mail: host@math.univ-mlv.fr

Department of Mathematics
University of Maryland
College Park, MD 20742, U.S.A.
E-mail: randall@math.umd.edu

LAGA, UMR CNRS 7539
Université de Paris 13
93430 Villetaneuse, France
E-mail: parreau@math.univ-paris13.fr

Received 8 November 1999;
revised 10 February 2000 (3849)