ISRAEL JOURNAL OF MATHEMATICS, Vol. 60, No. 3, 1987

CONSTRUCTION OF SETS OF POSITIVE MEASURE NOT CONTAINING AN AFFINE IMAGE OF A GIVEN INFINITE STRUCTURE

BY

J. BOURGAIN

Institut des Hautes Etudes Scientifiques, 35, route de Chartres, 91440 Bures-sur-Yvette, France

ABSTRACT

This paper deals with the problem of existence of infinite structures in euclidean space such that every set of positive measure contains an affine image of it. We contribute to P. Erdös' question about sequences in the real line, by showing that no triple sum of infinite sets has this property.

1. Introduction

P. Erdös [1] asked the question whether there exists an infinite sequence S of real numbers such that every measurable subset of $\mathbf{R}$ of positive measure contains an affine image of S. Presumably his problem was motivated by the results of Szemeredi [4], Furstenberg [2] and Furstenberg–Katznelson [3] relative to finite structures. It will be convenient to say, more generally, that a subset S of $\mathbf{R}^d$ ($d = 1, 2, \ldots$) has property (E) provided any $A \subset \mathbf{R}^d$ of positive measure contains an affine image of S. It seems reasonable to conjecture that such a set S has to be finite. Let $d = 1$ and $\{\varepsilon_k\}$ a sequence of positive numbers, decreasing to 0. It is known and easy to show (considering appropriate Cantor sets) that the condition $(\varepsilon_{k-1} - \varepsilon_k)/\varepsilon_k \to 0$ implies that $\{\varepsilon_k\}$ fails (E). The problem seems open for $\varepsilon_k = 2^{-k}$, for instance.

Our purpose here is to disprove an infinite version of the 3-dimensional Szemeredi phenomenon.

PROPOSITION 1. *If S is an infinite set of reals, then the set $S \times S \times S$ in $\mathbf{R}^3$ fails property* (E).

Received May 14, 1987

It should be pointed out that the previous statement remains valid if we weaken the definition of property (E) allowing different scalings according to the dimensions 1, 2, 3. An elaboration of the argument also permits to prove the following 1-variable result.

PROPOSITION 2. *If* S_1, S_2, S_3 *are infinite subsets of* $\mathbf{R}$, *then* $S_1 + S_2 + S_3$ *fails* (E).

We denote $A + B = \{a + b \mid a \in A, b \in B\}$. There are variants of this result and for certain sequences S, for instance $\{2^{-k}\}$, the method yields that already $S + S$ fails (E).

The author is grateful to B. Weiss for some discussions on these problems.

2. A reformulation of Property (E)

(E) is equivalent to a statment of a more "finite" nature and which may be formulated as an inequality.

LEMMA 1. *The following conditions are equivalent for a bounded subset* S *of* $\mathbf{R}^d$:

(i) *There is a constant* C *such that*

$$\int \inf_{1<t<2} \sup_{y \in S_0} |f(x+ty)|\,dx \leqq C \int |f(x)|\,dx \tag{1}$$

whenever S_0 *is a finite subset of* S *and* f *a continuous function on* $\mathbf{R}^d$ *with compact support.*

(ii) S *has property* (E).

Notice the scale invariance of (1), in the sense that (1) is equivalent to

$$\int \inf_{\delta<t<2\delta} \sup_{y \in S_0} |g(x+ty)|\,dx \leqq C \int |g(x)|\,dx \tag{2}$$

where

$$g(x) = \frac{1}{\delta^d}\, f\left(\frac{x}{\delta}\right); \qquad \delta > 0 \text{ arbitrarily chosen.}$$

Implication (ii)$\Rightarrow$(i) is the relevant one in proving the proposition.

PROOF OF (i)$\Rightarrow$(ii). Denote μ the d-dimensional Lebesgue measure and let

$S \subset B(0, D)$. Let K be a compact subset of $\mathbf{R}^d$, $\mu(K) > 0$. Considering a Lebesgue density point, we may find a cube Q satisfying

$$\mu(Q \cap K) > \left(1 - \frac{1}{4^d C}\right) \mu(Q) \tag{3}$$

where C is the constant appearing in (1).

Fix $\tau > 0$ and consider a continuous function $0 \leqq \varphi \leqq 1$ satisfying

$$\varphi = 0 \text{ on } K, \quad \varphi = 1 \text{ on } Q \setminus K_\tau \quad (K_\tau = \tau\text{-neighborhood of } K), \tag{4}$$

$$\int_{\mathbf{R}^d} \varphi(x) dx \leqq 2\mu(Q \setminus K). \tag{5}$$

(Clearly we may assume $\mu(Q \setminus K) > 0$.)

Let x_0 be the center of Q and define

$$f(x) = \varphi(x_0 + \varepsilon x), \qquad \varepsilon = \frac{\text{size } Q}{10D}.$$

Then, by (3), (5), $\hat{Q} \equiv$ cube with center x_0 and half the side-length of Q,

$$\int f(x) dx = \varepsilon^{-d} \int \varphi(x) dx < 2\varepsilon^{-d} \frac{1}{4^d C} \mu(Q) < \frac{1}{2C} \int \chi_{\hat{Q}}(x_0 + \varepsilon x) dx.$$

Hence, by (1), for any finite subset S_0 of S

$$\int \inf_{1<t<2} \sup_{y \in S_0} f(x + ty) dx < \frac{1}{2} \int \chi_{\hat{Q}}(x_0 + \varepsilon x) dx$$

from which the existence of a point x and $1 \leqq t \leqq 2$ (both dependent of S_0) such that for $y \in S_0$

$$\varphi(x_0 + \varepsilon x + \varepsilon t y) < \tfrac{1}{2} \chi_{\hat{Q}}(x_0 + \varepsilon x). \tag{6}$$

(6) implies that

$$x_0 + \varepsilon x \in \tilde{Q}, \tag{7}$$

$$x_0 + \varepsilon x + \varepsilon t y \notin Q \setminus K_\tau \qquad \text{for each } y \in S_0. \tag{8}$$

Notice that since

$$|z| < 2\varepsilon D < \frac{\text{size } Q}{5} \qquad \text{for } z \in \varepsilon t S_0$$

from (7)

$$x_0 + \varepsilon x + \varepsilon t S_0 \subset Q$$

and thus, from (8),

$$x_0 + \varepsilon x + \varepsilon t S_0 \subset K_\tau.$$

Since εt is bounded below independently of S_0 and τ, a compactness argument implies that K contains a translate of $t'S$, for some $t' > 0$.

PROOF OF (ii)$\Rightarrow$(i). Assume (i) fails. Thus for any $0 < \delta < 1$ and $0 < M < \infty$ there is a finite subset S_0 of S and a continuous positive function f such that

$$\int \inf_{\delta < t < 2\delta} \sup_{y \in S_0} f(x + ty)dx > M \int f(x)dx. \tag{9}$$

Denoting F the function

$$F(x) = \inf_{\delta < t < 2\delta} \sup_{y \in S_0} f(x + ty),$$

(9) gives

$$\int_0^\infty \mu[F \geqq \lambda]d\lambda > M \int_0^\infty \mu[f \geqq \lambda]d\lambda$$

and hence for some $\lambda > 0$

$$\mu[F \geqq \lambda] > M\mu[f \geqq \lambda].$$

Considering $A = [f \geqq \lambda]$, it follows thus that

$$\mu(A_1) > M\mu(A) \qquad \text{where } A_1 = \bigcap_{\delta < t < 2\delta} \left[\bigcup_{y \in S_0} (A - ty) \right] \tag{10}$$

(A, A_1 are compact sets).

Our next purpose is to replace A_1 by a set of small complementary measure and A by a set of small measure (in an appropriate quotient space). Assume again $S \subset B(0, D)$ and S_0 containing the origin 0. Considering the lattice $D\mathbf{Z}^d$, split $\mathbf{R}^d$ in cubes of size D and consider such a D-cube I satisfying

$$\mu(A_1 \cap I) \geqq \frac{M}{2^d} \mu(A \cap \tilde{I}) \tag{11}$$

where $\tilde{I} = I + [0, D]^d$. This is possible by (10). It also follows from (10) that if

$x \in A_1 \cap I$, then for any $\delta < t < 2\delta$, $x + ty \in A$ for some $y \in S_0$, hence $x + ty \in A \cap (I + tS) \subset A \cap \tilde{I}$.

Let G stand for the torus $\mathbf{R}^d/2D\mathbf{Z}^d$ and ν the normalized quotient measure on G. Let B (resp. B_1) be the images of $A \cap I$ (resp. $A_1 \cap I$) under the quotient map $\mathbf{R}^d \to G$. Thus (11) gives

$$\nu(B) < \tau\nu(B_1) \tag{12}$$

where $\tau = 2^d/M$ can be made arbitrarily small. The advantage of dealing with subsets of the torus G is that by random translations, it is now possible to construct sets $\tilde{B}$, $\tilde{B}_1$ where

$$\nu(G \setminus \tilde{B}_1) < \tau \quad \text{and} \quad \nu(\tilde{B}) < \tau \log \frac{1}{\tau}. \tag{12}$$

Since $\tilde{B}$, $\tilde{B}_1$ are the union of translates of B, B_1 by the same sequence of points, it is clear that given $x \in \tilde{B}_1$ and $\delta < t < 2\delta$, the subset of G

$$\{x + ty \mid y \in S_0\}$$

will intersect $\tilde{B}$. By construction, this property is indeed true for B, B_1. Define $\Lambda = \tilde{B}_1 \setminus \tilde{B}$ for which by (12)

$$\nu(\Lambda) > 1 - 2\tau \log \frac{1}{\tau}$$

and which does not contain any translate of tS_0 for $\delta < t < 2\delta$. Take δ of the form 2^{-k} ($k = 0, 1, 2, \ldots$) and consider subsets Λ_k of G each satisfying

$$\nu(\Lambda_k) > 1 - \tfrac{1}{10} 2^{-k}. \tag{13}$$

Λ_k does not contain a translate of tS_k, $2^{-k} \leqq t < 2^{-k+1}$ where S_k is an appropriate finite subset of S.

By (13), the set $\Lambda = \bigcap \Lambda_k$ has measure $\nu(\Lambda) > \frac{1}{2}$. Moreover, it does not contain a translate of tS, for any $0 < t < 2$. Lifting Λ back to a subset of $\mathbf{R}^d$, property (E) for the set S is clearly disproved.

3. Construction of random functions

Consider an infinite subset S of $\mathbf{R}$. Our purpose is to disprove a uniform inequality (1) (see Lemma 1) relative to the set $S \times S \times S$ on $\mathbf{R}^3$. We will generate functions using a random method.

By passing to a subset and translation, we may identify S with a decreasing sequence $\{\varepsilon_k\}$ of positive numbers in $[0, 1]$.

Fix integers K, J, N (J, N will depend polynomially on K). Consider numbers

$$\xi = \xi(s) = \frac{s_1}{N\varepsilon_1} + \cdots + \frac{s_K}{N\varepsilon_K} \qquad \text{where } s_k \in \{0, 1, \ldots, N^2\}. \tag{14}$$

Assuming ε_k rapidly enough decreasing, we can ensure that

$$\xi\varepsilon_k \approx \frac{s_k}{N} + \frac{\varepsilon_k}{\varepsilon_{k+1}}\frac{s_{k+1}}{N} + \cdots + \frac{\varepsilon_k}{\varepsilon_K}\frac{s_K}{N}. \tag{15}$$

Let $\Omega = \{0, 1, 2, \ldots, N^2\}^K$ be endowed with normalized counting measure and for $1 < t < 2$, consider the map

$$\alpha_t : \Omega \to \Pi^K : s \to (t\xi\varepsilon_1, \ldots, t\xi\varepsilon_K)$$

where $\Pi = \mathbf{R}/\mathbf{Z}$ and the relation between ξ and s given by (14). Using (15), it is easily seen that if N is taken large enough, then for each $1 < t < 2$ the image measure of α_t becomes an arbitrary fine discretization of the normalized Haar measure of Π^k.

Let φ be a bumpfunction, $0 \leqq \varphi \leqq 1$, $\varphi = 1$ on $[0, 3]^3$ and $\varphi = 0$ outside $[-1, 4]^3$. Let h be a function on Π to be specified later, $h \in L^2(\Pi)$, $\| h \|_2 \leqq 10$. Define

$$f(x, y, z) = \varphi(x, y, z) \sum_{j=1}^{J} \theta_j h(\xi_j x) h(\xi_j y) h(\xi_j z) \tag{17}$$

where θ_j are ± 1 random signs and the ξ_j's as in (14). Thus

$$\int |f_\theta(x, y, z)|\, dx dy dz d\theta < c\sqrt{J}\, \| h \|_2^3 \sim \sqrt{J} \tag{18}$$

where the θ-integration is performed on $\{1, -1\}^J$ (assuming ξ_j not all 0 and ε_k's small enough).

We are considering translates by elements of the set

$$S_0 = \{(\varepsilon_k, \varepsilon_l, \varepsilon_m) \mid 1 \leqq k, l, m \leqq K\}.$$

Clearly

$$\int \inf_{1<t<2} \max_{k,l,m} |f_\theta(x+\varepsilon_k t, y+\varepsilon_l t, z+\varepsilon_m t)|\, dxdydzd\theta$$

$$\geqq \inf_{\substack{1<t<2\\ x,y,z,\theta}} \max_{k,l,m} \left| \sum_{j=1}^{J} \theta_j h(\xi_j x + \alpha_t(\xi_j)_k) h(\xi_j y + \alpha_t(\xi_j)_l) h(\xi_j z + \alpha_t(\xi_j)_m) \right|.$$

Our purpose is to prove the existence of ξ_j's, hence elements $s^1, \ldots, s^J$ in Ω, such that for all $1<t<2$; $x_j, y_j, z_j \in \Pi$, $\theta_j = \pm 1$ $(1 \leqq j \leqq J)$ we have

$$\max_{k,l,m} \left| \sum_{j=1}^{J} \theta_j h(x_j + \alpha_t(\xi_j)_k) h(y_j + \alpha_t(\xi_j)_l) h(z_j + \alpha_t(\xi_j)_m) \right| \gg \sqrt{J}. \tag{20}$$

This would give a lower bound on (19) and from (18), (19) we conclude the existence of some $\theta \in \{1, -1\}^J$ such that the corresponding f_θ and the set S_0 disprove (i) of Lemma 1.

This existence of a system $s^1, \ldots, s^J$ in Ω satisfying (20) will follow from statistical considerations.

Assume a derivative estimate on h (which is polynomial in K). We first reduce the number of relevant values of t's and elements x_j, y_j, z_j. Observe, from the definition of $\alpha_t(\xi)$, that $\alpha_t(\xi_j)$ is well defined by the system

$$\{e^{2\pi i (t/N)(\varepsilon_k/\varepsilon_l)}\}_{1 \leqq k \leqq l \leqq K}$$

defining an element of $\Pi^{(K^2)}$. Hence we are led to consider a τ-net $(\log(1/\tau) \sim \log K)$ in Π^{K^2+3J} and the number of relevant values $x_j, y_j, z_j \in \Pi$, $1<t<2$ is reduced to

$$\exp(cK^2 \log K + cJ \log K). \tag{21}$$

For θ, there are 2^J possibilities.

Thus fixing x_j, y_j, z_j, t, θ, we are interested in the probability that $s^1, \ldots, s^J$ in Ω^J will satisfy the event

$$\max_{k,l,m} \left| \sum_{j=1}^{J} \theta_j h(x_j + \alpha_t(\xi_j)_k) h(y_j + \alpha_t(\xi_j)_l) h(z_j + \alpha_t(\xi_j)_m) \right| > M\sqrt{J} \tag{22}$$

where M is a fixed large constant. By (21), we will be done by showing that this probability always exceeds

$$1 - \exp[-c(K^2+J)\log K]. \tag{23}$$

The map

$$(s^1, \dots, s^J) \to \{\alpha_t(\xi_j)_k\}_{\substack{1 \leq j \leq J \\ 1 \leq k \leq K}}$$

induces a discrete measure on Π^{KJ} and for N large enough we get arbitrary fine expectations of it approximating the invariant measure. By the gradient estimate on h, $\log N \sim \log K$ already will permit to identify the image measures on $\mathbf{R}^{JK}$ of the respective maps

$$\Omega^{JK} \to \mathbf{R}^{JK} : (s^1, \dots, s^J) \mapsto \{h(x_j + \alpha_t(\xi_j)_k)\}$$

and

$$\Pi^{JK} \to \mathbf{R}^{JK} : (u_{j,k}) \mapsto \{h(u_{j,k})\}.$$

Thus we need to evaluate

$$\text{(24)} \qquad \mathbf{P}_{\Pi^{JK}}\left[\max_{1<k,l,m \leq K} \left| \sum_{j=1}^{J} \theta_j h(u_{j,k}) h(u_{j,l}) h(u_{j,m}) \right| < M\sqrt{J}\right].$$

It remains to specify h. The ideal should be to ensure that $h : \Pi \to \mathbf{R}$ has Gaussian image measure. However, such h would not satisfy a gradient estimate and therefore some care is needed. A substitute for the Gaussian is given by

LEMMA 2. *Given $0 < \beta$ there is a map $h_\beta : \Pi \to \mathbf{R}$ with image measure ν_β satisfying*

$$\text{(25)} \qquad |H'_\beta| < c\beta^{-1} \quad \textit{and} \quad \nu_\beta < (1+\beta)\nu$$

where ν is the Gaussian measure.

Take $\beta = 1/2KJ$ and let $\mu_\beta = \nu_\beta \otimes \cdots \otimes \nu_\beta$ be the JK-fold product measure on $\mathbf{R}^{JK}$. Thus (25) implies

$$\text{(26)} \qquad \mu_\beta \leqq 2\mu$$

where μ is the Gaussian measure on $\mathbf{R}^{JK}$.

Taking $h = h_\beta$ in (24), we get thus by (26) the condition

$$\mu_\beta\left[\max_{k,l,m}\left|\sum_{j=1}^{J}\theta_j x_{jk}x_{jl}x_{jm}\right|<M\sqrt{J}\right]\leqq 2\mu\left[\max_{k,l,m}\left|\sum_{j=1}^{J}\theta_j x_{j,k}\,x_{j,l}x_{j,m}\right|<M\sqrt{J}\right]$$

$$<\exp[-c(K^2+J)\log K]$$

where $(x_{j,k})$ is a random variable in $\mathbf{R}^{JK}$.

The proof will thus be completed by the following fact:

LEMMA 3. *For $K^2<cJ$ and $\{g_{jk}\}_{1\leqq j\leqq J,\,1\leqq k\leqq K}$ independent Gaussians*

$$\mathbf{P}\left[\max_{1\leqq k,l,m\leqq K}\left|\sum_{j=1}^{J}g_{jk}g_{jl}g_{jm}\right|<M\sqrt{J}\right]<\exp(-c(M)K^3). \tag{27}$$

It is likely that a version of Lemma 3 remains valid replacing Gaussians by Steinhaus functions or independent signs. The interest of Gaussians in the proof is the rotational invariance.

4. Proof of Lemmas

PROOF OF LEMMA 2. Realize an (L^1)-normalized symmetric Gaussian g as a symmetric map $g:[-\frac12,\frac12]\to\mathbf{R}$ defined by

$$g(\tfrac12-t)=-g(t)\quad\text{and}\quad \nu[x>g(t)]=1-2t\qquad\text{for }\tfrac14\leqq t\leqq\tfrac12.$$

Thus $g(\frac14)=g(-\frac14)=0$ and g behaves like

$$\begin{cases}-\left(\log\dfrac{1}{|t|}\right)^{1/2} & \text{if } |t|\to 0\\[2ex] \left(\log\dfrac{1}{\frac12-|t|}\right)^{1/2} & \text{if } |t|\to\frac12.\end{cases}$$

Fixing $\beta>0$ consider the affine map $c_\beta:[0,\frac12]\to[\beta,\frac12-\beta]$ defined by $c_\beta(t)=\frac14+(1-4\beta)(t-\frac14)$. Define $h_\beta(t)=g(c_\beta(t))$ on $[0,\frac12]$ and $h_\beta(t)=h_\beta(-t)$. Thus

$$|h'_\beta|\leqq-\frac{d}{dt}\left(\log\frac1t\right)^{1/2}\Big|_{t=\beta}\leqq\beta^{-1}$$

while the image measure ν_β of h_β fulfils ($m=[0,\frac12],dx$)

$$\tfrac{1}{2}\nu_\beta(A) = m[g(c_\beta(t)) \in A] = m[c_\beta^{-1}(g^{-1}(A))] < \frac{1}{1-4\beta}\, m[g^{-1}(A)]$$

$$= \frac{1}{1-4\beta}\,\frac{\nu(A)}{2}$$

for A a subset of $\mathbf{R}$.

PROOF OF LEMMA 3. (27) is equivalent to

$$\mathbf{P}\left[\max_{k,l,m}\left|\sum_{j=1}^{J} g_{jk}g'_{jl}g''_{jm}\right| < M\sqrt{J}\right] < \exp(-c(M)K^3) \tag{28}$$

where now the g_{jk}, g'_{jk}, g''_{jk} are independent. The proof of (28) is based on two observations.

Assume first $(\xi^r)_{1\leq r\leq R}$ a sequence of vectors in $\mathbf{R}^J$ such that

$$\operatorname{dist}(\xi^r, \operatorname{span}(\xi^1, \ldots, \xi^{r-1})) > \gamma. \tag{29}$$

Then

$$\mathbf{P}\left[\max_{1\leqq r\leqq R}\left|\sum_{j=1}^{J} g_j\xi_j^r\right| < M\right] < \exp(-c(M,\gamma)R). \tag{30}$$

To prove (30), write

$$\xi^1 = e_1,$$

$$\xi^r = c_r e_r + f_r, \qquad f_r \in \operatorname{span}(e_1, \ldots, e_{r-1}) \qquad (1 < r \leqq R),$$

where $(e_r)_{1\leqq r\leqq R}$ is an orthonormal system and $|c_r| > \gamma$ for each r. Denoting $G = (g_1, \ldots, g_J)$, the system $(\langle G, e_r\rangle)_{1\leqq r\leqq R}$ is a system of independent Gaussian variables. Thus we have to majorize

$$\mathbf{P}\left[\max_{1\leqq r\leqq R}\left|g_r - \sum_{s=1}^{r-1}\alpha_{r,s}g_s\right| < \frac{M}{\gamma}\right] \tag{31}$$

where the $(\alpha_{r,s})_{s<r}$ are real numbers. But, by the independence of the (g_r), (31) is clearly less than $[1-\exp(-2M^2/\gamma^2)]^R$, hence (30).

Say that a system of K^2 vectors $(y^s)_{1\leqq s\leqq K^2}$ fulfils $P_{R,\gamma}$, provided there is a subsequence $(\xi^r)_{1\leqq r\leqq R}$ for which (29) holds. It is clear from (30) that the left member of (28) can be estimated in the following way:

$$\text{(32)} \qquad \mathbf{P}\left[\left\{\frac{1}{\sqrt{J}}(g_{jk}g'_{jl})_{1\leqq j\leqq J}\right\}_{1\leqq k,l\leqq R} \text{ fails } P_{R,\gamma}\right] + \exp(-c(M,\gamma)RK).$$

Thus it remains to evaluate the first term in (32), taking $R \sim K^2$. For x a vector in $\mathbf{R}^J$, define

$$|x|_0 = \frac{1}{J}\ \#\left\{1 \leqq j \leqq J;\ |x_j| > \frac{1}{100}\right\}.$$

It is straightforward, using elementary entropy considerations, to show that if E is any $J/100$-dimensional subspace of $\mathbf{R}^J$, then

$$\text{(33)} \qquad \mathbf{P}\left[\inf_{x\in X} |G-x|_0 \leqq \frac{1}{100}\right] < e^{-J}.$$

Using this observation, it follows that for ω fixed outside an exceptional set of measure at most

$$\text{(34)} \qquad \exp\left(-\frac{K}{3}J\right)$$

the system $\{(g_{jk}(\omega))_{1\leqq j\leqq J}\}_{1\leqq k\leqq K}$ has a subsystem $\{\eta^r\}_{r=1}^{K/2}$ satisfying the condition

$$\text{(35)} \qquad \inf_{x\in[\eta^1,\dots,\eta^{r-1}]} |\eta^r - x|_0 > \frac{1}{100} \qquad \left(1 \leqq r \leqq \frac{K}{2}\right).$$

We now use the ω'-variable. If $\{(\eta^r_j g'_{jl})\}_{1\leqq r\leqq K/2,\,1\leqq l\leqq K}$ fails $P_{R,\gamma}$ with $R = K^2/4$, there has to be at least $K/2$ values of l for which

$$\min_{1\leqq r\leqq K/2} \operatorname{dist}[(\eta^r_j g'_{jl}), \operatorname{span}[(\eta^s_j g'_{j,l_1}), (\eta^{r_1}_j g'_{jl}) \mid l_1 < l,\ 1 \leqq s \leqq K/2,\ r_1 < r]]$$

is bounded by $\gamma\sqrt{J}$.

A probabilistic estimate for this event is given by

$$\text{(36)} \qquad \binom{K}{K/2}\left(\frac{K}{2}\alpha\right)^{K/2}$$

where α is an upper bound for the probability of the event

$$\text{(37)} \qquad \sqrt{J}\gamma > \operatorname{dist}[(\eta^r_j g_j), E + \operatorname{span}\{(\eta^{r_1}_j g_j) \mid r_1 < \eta\}]$$

where E is any $K^2/2$-dimensional subspace of $\mathbf{R}^J$.

If $(g_j(\omega))$ satisfies (37), then

$$\sum_j |c_j g_j(\omega) - x_j|^2 < \gamma^2 J$$

for some $x \in E$ and where, by (35), $c_j = \eta_j^r - \Sigma_{s<r} \alpha_s \eta_j^s$ satisfies $|(c_j)|_0 > \frac{1}{100}$. Hence there is a coordinate set I, $|I| = J/100$ such that

$$\sum_{j \in I} |g_j(\omega) - y_j|^2 < 10^4 \gamma^2 J \tag{38}$$

for some $y \in P_I(E)$, P_I = coordinate projection on $[e_j \mid j \in I]$. Again, entropy considerations show that this probability is bounded by

$$\alpha < \binom{J}{J/100} \left(\frac{1}{\gamma}\right)^{K^2} (10^3 \gamma)^{J/200} < e^{-J} \tag{39}$$

for an appropriate constant γ and assuming K^2 small enough with respect to J.

Collecting estimates, the following bound on (28) is obtained from (32), (34), (36), and (39):

$$\exp\left(-\frac{1}{4} c(M,\gamma) M^3\right) + \exp\left(-\frac{K}{3} J\right) + \binom{K}{K/2} \left(\frac{K}{2} e^{-J}\right)^{K/2}$$

provided $K^2 < cJ$.

This proves (28) and (27).

References

1. P. Erdös, Private communication, 1983.

2. H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204–256.

3. H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for commuting transformation*, J. Analyse Math. **34** (1978), 275–291.

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