Sets of transfer times with small densities

Citation for the original published paper (version of record):

Björklund, M., Fish, A., Shkredov, I. (2021). Sets of transfer times with small densities. Journal de l'Ecole Polytechnique - Mathematiques, 8: 311-329. http://dx.doi.org/10.5802/JEP.147

Journal de l'École polytechnique
Mathématiques

Michael Björklund, Alexander Fish, & Ilya D. Shkredov

Sets of transfer times with small densities

Tome 8 (2021), p. 311-329.

<http://jep.centre-mersenne.org/item/JEP_2021__8__311_0>

Journal de l'École polytechnique
Mathématiques
Tome 8, 2021, p. 311–329 DOI: 10.5802/jep.147

SETS OF TRANSFER TIMES WITH SMALL DENSITIES

by Michael Björklund, Alexander Fish & Ilya D. Shkredov

Abstract. — In this paper we introduce and discuss various notions of doubling for measure-preserving actions of a countable abelian group G. Our main result characterizes 2-doubling actions, and can be viewed as an ergodic-theoretical extension of some classical density theorems for sumsets by Kneser. All of our results are completely sharp and they are new already in the case when $G = (\mathbb{Z}, +)$.

Résumé (Ensembles de temps de transfert avec petites densités). — Dans cet article, nous introduisons et discutons plusieurs notions de doublement pour des actions préservant la mesure sur un groupe abélien dénombrable G. Notre résultat principal caractérise les actions 2-doublantes et peut être vu comme une extension de nature ergodique de certains théorèmes de densité classiques pour les sommes d'ensembles par Kneser. Tous nos résultats sont optimaux et sont nouveaux déjà pour le cas où $G = (\mathbb{Z}, +)$.

Contents

1. Introduction

Throughout this paper, we shall assume that

– G is a countable and discrete abelian group.

– (X, μ) is a standard probability measure space, endowed with an *ergodic* probability measure-preserving action of G. In other words, (X, μ) is an ergodic *Borel G-space.*

Mathematical subject classification (2020). — 37A44, 28D05, 11B13.
Keywords. — Return times, inverse theorems, sumsets.

M.B. was supported by the Swedish Research Council (VR-)grant VR M. Björklund 20-23 (Project number: 11253320), A.F. was supported by the Australian Research Council grant DP210100162, and I.S. was supported by the Russian Science Foundation under grant 19-11-00001.

– (F_n) is a sequence of finite subsets of G with the property that for every bounded measurable function φ on X, there exists a μ-conull subset $X_\varphi \subset X$ such that

$$\lim_n \frac{1}{|F_n|} \sum_{g \in F_n} \varphi(gx) = \int_X \varphi \, d\mu, \quad \text{for all } x \in X_\varphi. \tag{1.1}$$

Definition 1.1. — Let C be a subset of G. The *lower asymptotic density* $\underline{d}(C)$ of the set C with respect to the sequence (F_n) is defined by

$$\underline{d}(C) = \varliminf_n \frac{|C \cap F_n|}{|F_n|}. \tag{1.2}$$

Definition 1.2. — Let A and B be measurable subsets of X with positive μ-measures. The *set of transfer times* $\mathscr{R}_{A,B}$ is defined by

$$\mathscr{R}_{A,B} = \big\{ g \in G \mid \mu(A \cap g^{-1}B) > 0 \big\}. \tag{1.3}$$

We set $\mathscr{R}_A = \mathscr{R}_{A,A}$, and refer to $\mathscr{R}_A$ as the *set of return times to the set* A. If we wish to emphasize the dependence on the measure μ, we write $\mathscr{R}^\mu_{A,B}$ and $\mathscr{R}^\mu_A$ respectively.

The aim of this paper is to establish various lower bounds on $\underline{d}(\mathscr{R}_{A,B})$, and to discuss when these lower bounds are attained. As we shall see in the proofs below, these questions are closely related to direct and inverse theorems for product sets with small doubling, which is an active line of research in additive combinatorics.

Before we state our main results, we make a few preliminary remarks. Firstly, our assumption (1.1) on the sequence (F_n) readily implies that the lower asymptotic density of $\mathscr{R}_{A,B}$ is strictly positive for all measurable subsets A and B of X with positive measures. Secondly,

(i) If $\mu(A) + \mu(B) > 1$, then $\mu(A \cap g^{-1}B) > 0$ for all g, whence $\mathscr{R}_{A,B} = G$.

(ii) If $\mu(A) + \mu(B) = 1$, then *either* $\mathscr{R}_{A,B} = G$, or $\mu(A \cap g_o^{-1}B) = 0$ for some g_o. In the latter case, $B = g_o^{-1}A^c$ modulo μ-null sets, so if denote by H the μ-essential stabilizer of A, then $\mathscr{R}_{A,B} = G \smallsetminus g_o^{-1}H$.

In order to arrive at non-trivial results about sets of transfer times, we shall for the rest of the paper, always assume that the measurable subsets A and B in X satisfy

$$\mu(A) + \mu(B) < 1. \tag{1.4}$$

In particular, if $A = B$, we shall assume that $\mu(A) < 1/2$.

1.1. Main results.

— Our first theorem roughly asserts that if the lower asymptotic density of $\mathscr{R}_A$ is small enough, then the set of transfer times $\mathscr{R}_A$ is in fact a subgroup of G.

Theorem 1.3. — *Let A and B be measurable subsets of X with positive μ-measures.*

(i) $\underline{d}(\mathscr{R}_{A,B}) \geqslant \max(\mu(A), \mu(B))$.

(ii) *If $\underline{d}(\mathscr{R}_A) < \frac{3}{2}\mu(A)$, then there exists a subgroup $G_o < G$ with $[G : G_o] \leqslant 1/\mu(A)$, such that $\mathscr{R}_A = G_o$.*

Our second theorem in particular tells us that a measurable subset $A \subset X$ with positive μ-measure for which (ii) in Theorem 1.3 holds must be rather special. To be able to state the full result, we need the following notation of *control*.

Definition 1.4. — Let (Y, ν) be an ergodic Borel G-space and let $\pi : (X, \mu) \to (Y, \nu)$ be a G-factor map. If (A, B) and (C, D) are pairs of measurable subsets with positive measures in X and Y respectively, then we say that (C, D) *controls* (A, B) if

$$A \subset \pi^{-1}(C) \quad \text{and} \quad B \subset \pi^{-1}(D), \quad \text{modulo } \mu\text{-null sets},$$

and $\mathscr{R}^{\mu}_{A,B} = \mathscr{R}^{\nu}_{C,D}$. If we wish to emphasize the dependence on π, we say that (C, D) π-controls (A, B).

Theorem 1.5. — *Let A and B be measurable subsets of X with positive μ-measures and suppose that $\underline{d}(\mathscr{R}_{A,B}) < \mu(A) + \mu(B)$. Then there exist*

(i) *a proper finite-index subgroup $G_o < G$ and a homomorphism η from G onto the quotient group G/G_o,*

(ii) *a non-trivial G-factor $\sigma : (X, \mu) \to (G/G_o, m_{G/G_o})$, where m_{G/G_o} denote the normalized counting measure on G/G_o and G acts on G/G_o via η,*

(iii) *a finite subset $M \subset G/G_o$,*

such that $\mathscr{R}_{A,B} = \eta^{-1}(M)$. Furthermore, there are finite subsets $I_o, J_o \subset G/G_o$ such that the pair (I_o, J_o) σ-controls (A, B).

Remark 1.6. — Theorem 1.5 can be viewed as an ergodic-theoretical extension of Kneser's celebrated inverse theorem [4] for the lower asymptotic density of sumsets in $(\mathbb{N}, +)$, which roughly asserts that if $\underline{d}(A + B) < \underline{d}(A) + \underline{d}(B)$ for two subsets A and B of $\mathbb{N}$, where $\underline{d}$ denotes the lower asymptotic density with respect to the sequence $([1, n])$, then $A + B$ is periodic (modulo a finite set). The connection between sumsets and sets of transfer times will be discussed in more details below.

Theorem 1.5 also tells us that an ergodic Borel G-space which admits a pair (A, B) of measurable subsets with positive measures with a *small* set of transfer times (in the sense that the inequality $\underline{d}(\mathscr{R}_{A,B}) < \mu(A) + \mu(B)$ holds) must have a non-trivial periodic G-factor. We recall that a Borel G-space is called *totally ergodic* if every finite-index subgroup $G_o < G$ acts ergodically. Such Borel G-spaces cannot have non-trivial periodic G-factors, and thus we conclude the following corollary from Theorem 1.5.

Corollary 1.7. — *Suppose that $G \curvearrowright (X, \mu)$ is totally ergodic. Then, for all measurable subsets $A, B \subset X$ with positive μ-measures,*

$$\underline{d}(\mathscr{R}_{A,B}) \geqslant \min(1, \mu(A) + \mu(B)).$$

Our third theorem characterizes exactly when the lower bound in Corollary 1.7 is attained (assuming that the action is totally ergodic).

Theorem 1.8. — *Suppose that the action $G \curvearrowright (X,\mu)$ is totally ergodic. If*

$$\underline{d}(\mathscr{R}_{A,B}) = \mu(A) + \mu(B) < 1,$$

then there exist

(i) *a homomorphism $\eta : G \to \mathbb{T}$ with dense image.*

(ii) *a G-factor $\sigma : (X,\mu) \to (\mathbb{T}, m_{\mathbb{T}})$, where $m_{\mathbb{T}}$ denotes the normalized Haar measure on $\mathbb{T}$ and G acts on $\mathbb{T}$ via η.*

(iii) *closed intervals I_o and J_o of $\mathbb{T}$ with $m_{\mathbb{T}}(I_o) = \mu(A)$ and $m_{\mathbb{T}}(J_o) = \mu(B)$ such that (I_o, J_o) σ-controls (A,B) and*

$$\mathscr{R}_{A,B} = \eta^{-1}(J_o I_o^{-1}),$$

modulo at most two cosets of the subgroup $\ker \eta$.

Remark 1.9. — Conversely, it is not hard to show that if a pair (A,B) of measurable subsets in X is σ-controlled by a pair (I_o, J_o) as in the theorem above, then

$$\underline{d}(\mathscr{R}_{A,B}) = \min(1, \mu(A) + \mu(B)),$$

and thus Theorem 1.8 really provides a complete characterization of when the lower bound in Corollary 1.7 is attained (assuming total ergodicity).

1.2. Concerning the novelty and sharpness of our results so far

Example 1.1 (The constant $\frac{3}{2}\mu(A)$ in Theorem 1.3 is optimal). — Let $N \geqslant 4$ and consider the action of $G = (\mathbb{Z},+)$ on $X = \mathbb{Z}/N\mathbb{Z}$ by translations modulo N. The normalized counting measure μ on X is clearly invariant and ergodic. Let $A = \{0,1\} \subset X$ and note that

$$\mu(A) = \frac{2}{N} \quad \text{and} \quad \mathscr{R}_A = N\mathbb{Z} \cup \big(N\mathbb{Z}+1\big) \cup (N\mathbb{Z}-1) \subsetneq \mathbb{Z}.$$

It is not hard to check that

$$\underline{d}(\mathscr{R}_A) = \frac{3}{N} = \frac{3}{2}\,\mu(A),$$

but $\mathscr{R}_A$ is *not* a subgroup of $\mathbb{Z}$.

Example 1.2 (Ergodicity of the action is needed in Theorem 1.3). — Given positive real numbers δ and ε, we shall construct a *non-ergodic* probability measure μ for the shift action by $G = (\mathbb{Z},+)$ on the space $2^{\mathbb{Z}}$ of all subsets of $\mathbb{Z}$, endowed with the product topology, such that

$$\mu(A) < \delta \quad \text{and} \quad \underline{d}(\mathscr{R}_A) \leqslant (1+\varepsilon)\mu(A),$$

where $A = \{C \in 2^{\mathbb{Z}} \mid 0 \in C\}$, and such that the set of return times $\mathscr{R}_A$ projects onto every finite quotient of $\mathbb{Z}$. In particular, $\mathscr{R}_A$ cannot be a subgroup of $\mathbb{Z}$, nor can it be contained in a subgroup of $\mathbb{Z}$. Here, the exact choice of the sequence (F_n) in $\mathbb{Z}$ is not so important; for simplicity, we can assume that $F_n = [1,n]$ for all $n \geqslant 1$.

The construction of μ goes along the following lines. Given positive real numbers δ and ε, we choose $0<\eta<1$ such that $1+\varepsilon=(1+\eta)/(1-\eta)$, and we pick a strictly increasing sequence (p_k) of prime numbers such that

$$\frac{1}{p_1}<\delta \quad \text{and} \quad \sum_{k\geqslant 2}\frac{1}{p_k}\leqslant\frac{\eta}{p_1}. \tag{1.5}$$

For every $k\geqslant 1$, we denote by μ_k the uniform probability measure on the $\mathbb{Z}$-orbit of the subgroup $p_k\mathbb{Z}$ in $2^{\mathbb{Z}}$ and we note that $\mu_k(A)=1/p_k$. We now define

$$\mu=(1-\eta)\mu_1+\eta\sum_{k\geqslant 2}\frac{\mu_k}{2^{k-1}},$$

which is clearly a $\mathbb{Z}$-invariant *non-ergodic* Borel probability measure on $2^{\mathbb{Z}}$. One readily checks that

$$\mu(A)=\frac{1-\eta}{p_1}+\eta\sum_{k\geqslant 2}\frac{1}{p_k 2^{k-1}}, \quad \text{and} \quad \mathscr{R}_A=\bigcup_{k\geqslant 1}p_k\mathbb{Z}\subsetneq\mathbb{Z},$$

whence,

$$\frac{1-\eta}{p_1}\leqslant\mu(A)<\delta$$

and, thus, by (1.5) and the choice of η,

$$\underline{d}(\mathscr{R}_A)\leqslant\sum_{k\geqslant 1}\frac{1}{p_k}\leqslant\frac{1+\eta}{p_1}\leqslant\Big(\frac{1+\eta}{1-\eta}\Big)\mu(A)=(1+\varepsilon)\mu(A).$$

Clearly, $\mathscr{R}_A$ projects onto every finite quotient of $\mathbb{Z}$, which finishes our construction.

Example 1.3 (Non-Følner sequences). — All of the results in this paper are new already in the case when $G=(\mathbb{Z},+)$ and $F_n=\{1,\dots,n\}$ (this sequence satisfies (1.1) by Birkhoff's Ergodic Theorem). We stress however that we do *not* need to assume that the sequence (F_n) is Følner (asymptotically invariant) in G. For instance, in the case of $(\mathbb{Z},+)$, our results above also apply to the *sparse* sequence

$$F_n=\big\{\lfloor k\sqrt{2}+k^{5/2}\rfloor \mid k=1,\dots,n\big\}, \quad n\geqslant 1,$$

which is far from being a Følner sequence in $\mathbb{Z}$, but nevertheless satisfies (1.1) by [2].

Example 1.4 ("Non-conventional" lower asymptotic density). — If $\mathbb{Z}\curvearrowright(X,\mu)$ is totally ergodic, then the sequence (F_n) of squares

$$F_n=\{k^2 \mid k=1,\dots,n\}, \quad \text{for } n\geqslant 1,$$

satisfies (1.1) (the almost sure convergence follows from the work of Bourgain [3], while the identification of the limit – for totally ergodic actions – follows from the equidistribution (modulo 1) of the sequence $(n^2\alpha)$, for irrational α). In particular, by Corollary 1.7 we have

$$\varliminf_{n\to\infty}\frac{|\mathscr{R}_{A,B}\cap\{1,4,\dots,n^2\}|}{n}\geqslant\min(1,\mu(A)+\mu(B)),$$

for all measurable subsets $A,B\subset X$ with positive μ-measures.

1.3. On ergodic actions with small doubling

Definition 1.10 (C-doubling actions). — Let $C \geqslant 1$. We say that $G \curvearrowright (X,\mu)$ is *a C-doubling action* if for every $\delta > 0$, there exists a measurable subset $A \subset X$ with $0 < \mu(A) < \delta$ such that $\underline{d}(\mathscr{R}_A) \leqslant C\mu(A)$.

Remark 1.11. — We note that if the action is C-doubling, then it must also be C'-doubling for every $C' \geqslant C$.

It seems natural to ask about the structure of C-doubling actions. The following theorem provides a complete characterization of such actions.

Theorem 1.12. — *Let $C \geqslant 1$. An ergodic action $G \curvearrowright (X,\mu)$ is C-doubling if and only if there exist*

(i) *an infinite compact metrizable group K and a homomorphism $\eta : G \to K$ with dense image.*

(ii) *a G-factor $\sigma : (X,\mu) \to (K, m_K)$, where m_K denotes the normalized Haar measure on K and G acts on K via η.*

Furthermore,

– *if the identity component K^o of K has infinite index, then the action is* 1*-doubling.*

– *if the identity component K^o of K has finite index, then the action is* 2*-doubling.*

Remark 1.13. — Theorem 1.12 in particular asserts that an ergodic action is C-doubling for some $C \geqslant 1$ if and only if it has an infinite Kronecker factor (see e.g. [1] for definitions).

The same line of argument as the one leading up to Theorem 1.12 also proves the following result, whose proof we leave to the reader. We recall that $G \curvearrowright (X,\mu)$ is *weakly mixing* if the diagonal action $G \curvearrowright (X \times X, \mu \otimes \mu)$ is ergodic.

Scholium 1.14. — *There exist measurable subsets A and B of X with $\mu(A), \mu(B) > 0$ such that $\underline{d}(\mathscr{R}_{A,B}) < 1$ if and only if $G \curvearrowright (X,\mu)$ is* not *weakly mixing.*

1.4. A brief outline of the proofs. — Our first observation is that for any two measurable subsets A and B of X with positive μ-measures, there is a measurable μ-conull subset X_1 of X such that

$$\mathscr{R}_{A,B} = B_x A_x^{-1}, \quad \text{for all } x \in X_1,$$

where A_x and B_x are the return times of the point x to the sets A and B (see Section 2.1 below for notation). We then observe in Lemma 2.2 that for some measurable μ-conull subset $X_2 \subset X$,

$$\underline{d}(B_x A_x^{-1}) \geqslant \mu(A_x^{-1} B), \quad \text{for all } x \in X_2,$$

which puts us in the framework of our earlier paper [1]. We combine some of the key points of this paper in Lemma 2.4 below, the outcome of which is that there exist

– a measurable G-invariant μ-conull subset $X_3 \subset X_1 \cap X_2$,

– a compact and metrizable abelian group K with Haar probability measure m_K and a homomorphism $\tau : G \to K$ with dense image,

– a G-equivariant measurable map $\pi : X_3 \to K$ such that $\pi_*(\mu|_{X_3}) = m_K$, where G acts on K via τ,

– two measurable subsets I and J of K,

such that

$$\mu(A_x^{-1}B) = m_K(JI^{-1})$$

and

$$A \cap X_3 \subset \pi^{-1}(I) \quad \text{and} \quad B \cap X_3 \subset \pi^{-1}(J).$$

In particular,

$$\mu(A) \leqslant m_K(I) \quad \text{and} \quad \mu(B) \leqslant m_K(J),$$

and if $A = B$, then we can take $I = J$. In the settings of Theorem 1.3, Theorem 1.5 and Theorem 1.8, we see that

$$m_K(II^{-1}) < \frac{3}{2}\, m_K(I) \quad \text{and} \quad m_K(JI^{-1}) < m_K(I) + m_K(J)$$

and

$$m_K(JI^{-1}) = \min(1, m_K(I) + m_K(J))$$

respectively. At this point, we use some classical results [5] of Kneser for sumsets in compact abelian groups, to conclude that the pair (I, J) is "reduced" to a nicer pair (I_o, J_o) in a much "smaller" quotient group Q of K (see Definition 2.6 for details). The point of all this is that the transfer times $\mathscr{R}_{A,B}$ is *contained in* the transfer times between I_o and J_o, which is equal to the set $\eta^{-1}(J_o I_o^{-1})$. Here $\eta : G \to Q$ is the composition of τ with the quotient map from K to Q. To prove that the sets actually coincide, we shall use the *overshoot relation*

$$\mu(A) + \mu(B) \leqslant m_Q(I_o) + m_Q(J_o) - m_Q(I_o \cap \eta(g)^{-1} J_o), \tag{1.6}$$

for all $g \in \eta^{-1}(J_o I_o^{-1}) \smallsetminus \mathscr{R}_{A,B}$. This inequality is proved in Proposition 2.7. It turns out that in the settings of the theorems above, the sets I_o and J_o have the property that the m_Q-measure of the intersection $I_o \cap \eta(g)^{-1} J_o$, for g in $\eta^{-1}(J_o I_o^{-1}) \smallsetminus \mathscr{R}_{A,B}$, is large enough to contradict (1.6), whence we can conclude that $\mathscr{R}_{A,B} = \eta^{-1}(J_o I_o^{-1})$.

1.5. Ergodic actions of semi-groups. — Our definition of transfer times between two sets makes sense also for actions by non-invertible maps. Suppose that S is a countable abelian semigroup, sitting inside a countable abelian group G. If S acts ergodically by measure-preserving maps on a standard probability measure space (X, μ), then, under some technical assumptions (see e.g. [6] for more details in the general setting), one can construct a so called *natural extension* $(\widetilde{X}, \widetilde{\mu})$ of the S-action, which is a measure-preserving G-action, together with a measurable S-equivariant map $\rho : \widetilde{X} \to X$, mapping $\widetilde{\mu}$ to μ. It is not hard to see that if we set

$$\widetilde{A} = \rho^{-1}(A) \quad \text{and} \quad \widetilde{B} = \rho^{-1}(B),$$

then

$$\mathscr{R}_{\widetilde{A},\widetilde{B}} \cap S = \{s \in S \mid \mu(A \cap s^{-1}B) > 0\},$$

where the transfer times $\mathscr{R}_{\widetilde{A},\widetilde{B}}$ are measured with respect to $\widetilde{\mu}$. We can now apply our results above to the G-action on the natural extension $(\widetilde{X}, \widetilde{\mu})$ (which is ergodic if and only if the semi-group action $S \curvearrowright (X, \mu)$ is), and conclude the same results for the S-action. We leave the details to the interested reader.

Acknowledgements. — The authors thank the anonymous referees for their comments. I.S. is grateful to SMRI and the School of Mathematics and Statistics at Sydney University for funding his visit and for their hospitality. M.B. and A.F. wish to thank the organizers of the MFO workshop "Groups, dynamics and approximation", during which parts of this paper were written, for the invitation.

2. Preliminaries

2.1. Transfer times and action sets. — Given a subset D of X and $x \in X$, we define the *set of return time of x to D* by

$$D_x = \{g \in G \mid gx \in D\} \subset G,$$

and we note that $(gD)_x = gD_x$ and $D_x\, g^{-1} = D_{gx}$ for all $g \in G$. If F is a subset of G, then we define the *action set $FD \subset X$* by

$$FD = \bigcup_{f \in F} fD,$$

and we note that $(FD)_x = FD_x$. If E is another subset of X, then

$$(D \cap E)_x = D_x \cap E_x \quad \text{and} \quad (D \cup E)_x = D_x \cup E_x.$$

In particular,

$$D_x \cap g^{-1}E_x = (D \cap g^{-1}E)_x \quad \text{and} \quad D_x \cup g^{-1}E_x = (D \cup g^{-1}E)_x, \quad \text{for all } g \in G. \tag{2.1}$$

2.2. Transfer times as difference sets. — Let D be a measurable subset of X, and define

$$D^e = \{x \in X \mid D_x = \varnothing\} \quad \text{and} \quad D^{ne} = \{x \in X \mid D_x \neq \varnothing\}.$$

We note that $D^e = \bigcap_{g \in G} gD^c$ and $D^{ne} = GD$. In particular, D^e and D^{ne} are both measurable and G-invariant. Since μ is assumed to be ergodic, we conclude that

$$\mu(D^e) = 1 \quad \text{if } \mu(D) = 0 \tag{2.2}$$

and

$$\mu(D^{ne}) = 1 \quad \text{if } \mu(D) > 0. \tag{2.3}$$

Lemma 2.1. — *Let A and B be two measurable subsets of X with positive μ-measures, and define*

$$X_1 = \Big(\bigcap_{g \in \mathscr{R}_{A,B}} \{x \in X \mid A_x \cap g^{-1}B_x \neq \varnothing\}\Big) \cap \Big(\bigcap_{g \notin \mathscr{R}_{A,B}} \{x \in X \mid A_x \cap g^{-1}B_x = \varnothing\}\Big).$$

Then X_1 is a G-invariant measurable μ-conull subset of X and

$$\mathscr{R}_{A,B} = B_x A_x^{-1}, \quad \textit{for all } x \in X_1.$$

Proof. — Measurability and G-invariance of X_1 is clear, and μ-conullity of X_1 readily follows from applying (2.2) and (2.3) to the sets

$$D(g) := A \cap g^{-1} B, \quad \text{for } g \in G.$$

Indeed, $\mu(D(g)) > 0$ if and only if $g \in \mathscr{R}_{A,B}$, and by (2.1), we have

$$D(g)^e = \big\{x \in X \mid A_x \cap g^{-1} B_x = \varnothing\big\} \quad \text{and} \quad D(g)^{ne} = \big\{x \in X \mid A_x \cap g^{-1} B_x \neq \varnothing\big\}.$$

Note that for every $x \in X$,

$$\begin{aligned} B_x A_x^{-1} &= \big\{g \in G \mid A_x \cap g^{-1} B_x \neq \varnothing\big\} \\ &= \big\{g \in \mathscr{R}_{A,B} \mid A_x \cap g^{-1} B_x \neq \varnothing\big\} \sqcup \big\{g \notin \mathscr{R}_{A,B} \mid A_x \cap g^{-1} B_x \neq \varnothing\big\}. \end{aligned}$$

If $x \in X_1$, then

$$\big\{g \in \mathscr{R}_{A,B} \mid A_x \cap g^{-1} B_x \neq \varnothing\big\} = \mathscr{R}_{A,B} \quad \text{and} \quad \big\{g \notin \mathscr{R}_{A,B} \mid A_x \cap g^{-1} B_x \neq \varnothing\big\} = \varnothing,$$

which finishes the proof. □

2.3. Generic points. — We recall our assumptions on the sequence (F_n) of finite subsets of G: For every bounded measurable function φ on X, there exists a μ-conull subset $X_\varphi \subset X$ such that

$$\lim_n \frac{1}{|F_n|} \sum_{g \in F_n} \varphi(gx) = \int_X \varphi \, d\mu, \quad \text{for all } x \in X_\varphi.$$

The points in X_φ are said to be *generic* with respect to μ, φ and the sequence (F_n).

Lemma 2.2. — *Let A and B be two measurable subsets of X with positive μ-measures. Then there exists a measurable μ-conull subset $X_2 \subseteq X$ such that*

$$\mu(A_x^{-1} B) \leqslant \underline{d}(\mathscr{R}_{A,B}), \quad \textit{for all } x \in X_2.$$

Furthermore, for every finite subset L of G,

$$\underline{d}(L^{-1} A_x) = \mu(L^{-1} A) \quad \text{and} \quad \underline{d}(L^{-1} B_x) = \mu(L^{-1} B),$$

and for every $g \notin \mathscr{R}_{A,B}$,

$$\underline{d}(A_x \cup g^{-1} B_x) = \mu(A) + \mu(B),$$

for all $x \in X_2$.

Proof. — Given a subset $L \subset G$, we define

$$\varphi_L = \chi_{L^{-1}A} \quad \text{and} \quad \psi_L = \chi_{L^{-1}B} \quad \text{and} \quad X_L = X_{\varphi_L} \cap X_{\psi_L}.$$

We note X_L is a measurable μ-conull subset of X and for every $x \in X_L$,

$$\underline{d}(L^{-1} A_x) = \lim_n \frac{1}{|F_n|} \sum_{g \in F_n} \chi_{L^{-1}A}(gx) = \mu(L^{-1} A) \tag{2.4}$$

and

$$\underline{d}(L^{-1}B_x) = \lim_n \frac{1}{|F_n|} \sum_{g \in F_n} \chi_{L^{-1}B}(gx) = \mu(L^{-1}B). \tag{2.5}$$

We now set $X_2' = \bigcap_L X_L$, where the intersection is taken over the countable set of all *finite* subsets of G. Then X_2' is a measurable μ-conull subset of X, and for every $x \in X_2'$ and for every finite subset L of A_x, we have

$$\underline{d}(A_x^{-1}B_x) \geqslant \underline{d}(L^{-1}B_x) = \mu(L^{-1}B).$$

Since μ is σ-additive and $L \subset A_x$ is an arbitrary finite set, we can now conclude that

$$\underline{d}(A_x^{-1}B_x) \geqslant \mu(A_x^{-1}B) \quad \text{for all } x \in X_2'.$$

By Lemma 2.1, there exists a measurable μ-conull subset $X_1 \subseteq X$ such that $\mathscr{R}_{A,B} = B_x A_x^{-1}$ for all $x \in X_1$, and thus, since G is abelian,

$$\underline{d}(\mathscr{R}_{A,B}) = \underline{d}(A_x^{-1}B_x) \geqslant \mu(A_x^{-1}B), \quad \text{for all } x \in X_1 \cap X_2'.$$

Let $X_2 = X_1 \cap X_2'$ and pick $x \in X_2$. We note that if $g \notin \mathscr{R}_{A,B} = B_x A_x^{-1}$, then $A_x \cap g^{-1}B_x = \varnothing$, whence

$$\begin{aligned}\underline{d}(A_x \cup g^{-1}B_x) &= \varliminf_n \Big(\frac{|A_x \cap F_n|}{|F_n|} + \frac{|(g^{-1}B)_x \cap F_n|}{|F_n|} \Big) \\ &= \mu(A) + \mu(B) = \mu(A \cup g^{-1}B),\end{aligned}$$

by (2.4) and (2.5) (applied to the sets $L = \{e\}$ and $L = \{g\}$ respectively), since the limits of each term exist (the last identity follows from the fact that $\mu(A \cap g^{-1}B) = 0$ if $g \notin \mathscr{R}_{A,B}$). Since $x \in X_2$ is arbitrary, this finishes the proof. $\square$

Corollary 2.3. — *For all measurable subsets A and B of X, we have*

$$\underline{d}(\mathscr{R}_{A,B}) \geqslant \max(\mu(A), \mu(B)).$$

Proof. — By Lemma 2.2, there is a measurable μ-conull subset X_2 of X such that

$$\underline{d}(\mathscr{R}_{A,B}) \geqslant \mu(A_x^{-1}B) \geqslant \mu(B), \quad \text{for all } x \in X_2.$$

Since the roles of A and B are completely symmetric, this proves the corollary. $\square$

2.4. A correspondence principle for action sets. — The key ingredient in the proofs of Theorem 1.3, Theorem 1.5 and Theorem 1.8 is the following merger of a series of observations made by the first two authors in [1]. We outline the anatomy of this merger in the proof below. The rough idea is that action sets in an arbitrary ergodic G-action can be controlled by sets in an isometric factor (that is to say, a compact group, on which G acts by translations via a homomorphism from G into the compact group with dense image).

Lemma 2.4. — *Let A and B be measurable subsets of X with positive μ-measures. Then there exist*

- *a G-invariant measurable μ-conull subset $X_3 \subseteq X$,*

– *a compact and metrizable abelian group* K *with Haar probability measure* m_K*, a homomorphism* $\tau : G \to K$ *with dense image, and two measurable subsets* I *and* J *of* K*,*

– *a* G*-equivariant measurable map* $\pi : X_3 \to K$ *such that* $\pi_*(\mu|_{X_3}) = m_K$*, where* G *acts on* K *via* τ*,*

such that

$$A \cap X_3 \subseteq \pi^{-1}(I) \quad \text{and} \quad B \cap X_3 \subseteq \pi^{-1}(J)$$

and

$$\mu(A_x^{-1}B) = m_K(I^{-1}J) \quad \text{and} \quad A_x^{-1}(B \cap X_3) \subseteq \pi^{-1}(\pi(x)I^{-1}J),$$

for all $x \in X_3$*. In the case when* $A = B$*, we can take* $I = J$*. Finally, if* $G \curvearrowright (X, \mu)$ *is totally ergodic, then* K *must be connected.*

Remark 2.5. — If I and J are Borel measurable subsets of K, then their difference set $I^{-1}J$ might fail to be Borel measurable. However, since $I^{-1}J$ is the image of the Borel measurable subset $I \times J$ in $K \times K$ under the continuous map $(k_1, k_2) \mapsto k_1^{-1}k_2$, we see that $I^{-1}J$ is an analytic set, so in particular measurable with respect to the *completion* of the Borel σ-algebra of K with respect to m_K, and thus the expression $m_K(I^{-1}J)$ is well-defined.

Proof. — By [1, Lem. 5.3], there exists a G-invariant measurable μ-conull subset $X_3' \subset X$ such that

$$\mu(A_x^{-1}B) = \mu \otimes \mu(G(A \times B)), \quad \text{for all } x \in X_3'.$$

By [1, Th. 5.1], there exist

– a measurable G-invariant μ-conull subset $X_3'' \subset X$,

– a compact and metrizable abelian group K with Haar probability measure m_K and a homomorphism $\tau : G \to K$ with dense image,

– a G-equivariant measurable map $\pi : X_3'' \to K$ such that $\pi_*(\mu|_{X_3''}) = m_K$, where G acts on K via τ,

– two measurable subsets I and J of K,

such that

$$\mu \otimes \mu(G(A \times B)) = m_K(I^{-1}J)$$

and

$$A \cap X_3'' \subset \pi^{-1}(I) \quad \text{and} \quad B \cap X_3'' \subset \pi^{-1}(J).$$

It follows from the proof of [1, Th. 5.1] that if $A = B$, then we can take $I = J$. Since the set X_3'' is G-invariant, we see that

$$A_x \subset \pi^{-1}(I)_x = \tau^{-1}(I\pi(x)^{-1}), \quad \text{for all } x \in X_3'',$$

whence

$$A_x^{-1}(B \cap X_3'') \subset A_x^{-1}\pi^{-1}(J) = \pi^{-1}(\tau(A_x)^{-1}J) \subset \pi^{-1}(\pi(x)I^{-1}J).$$

Let $X_3 := X_3' \cap X_3''$ and note that X_3 is G-invariant and μ-conull. Thus the proof is finished modulo our assertion about total ergodicity. Suppose that K is not connected.

Then there is an open subgroup U of K, and $G_o = \tau^{-1}(U)$ is a finite-index subgroup of G. We note that $C := \pi^{-1}(U)$ is a G_o-invariant measurable subset of X, with positive μ-measure, but which cannot be μ-conull, since it does not map onto K under π (modulo μ-null sets). We conclude that $G \curvearrowright (X, \mu)$ is not totally ergodic. $\square$

2.5. Putting it all together. — Let K and Q be compact groups with Haar probability measures m_K and m_Q respectively and suppose that there is a continuous homomorphism p from K onto Q.

Definition 2.6 (Pair reduction). — Let (I, J) and (I_o, J_o) be two pairs of measurable subsets of K and Q respectively. We say that (I, J) *reduces to* (I_o, J_o) *with respect to* p if

$$I \subset p^{-1}(I_o) \quad \text{and} \quad J \subset p^{-1}(J_o) \quad \text{and} \quad m_K(JI^{-1}) = m_Q(J_o I_o^{-1}).$$

This notion is quite useful when we now summarize our discussion above.

Proposition 2.7 (A correspondence principle for transfer times). — *Let A and B be measurable subsets of X with positive μ-measures. Then there exist*

- *a compact and metrizable abelian group K with Haar probability measure m_K,*
- *a homomorphism $\tau : G \to K$ with dense image,*
- *a pair (I, J) of measurable subsets of K,*

which satisfy

$$\mu(A) \leqslant m_K(I) \quad \text{and} \quad \mu(B) \leqslant m_K(J) \quad \text{and} \quad m_K(JI^{-1}) \leqslant \underline{d}(\mathscr{R}_{A,B}).$$

Furthermore, suppose that Q is a compact group and $p : K \to Q$ is a continuous surjective homomorphism. If (I_o, J_o) is a pair of measurable subsets of Q such that (I, J) reduces to (I_o, J_o) with respect to p, then

$$\mathscr{R}_{A,B} \subseteq \tau_p^{-1}(J_o I_o^{-1}),$$

where $\tau_p = p \circ \tau$, and for all $g \in \tau_p^{-1}(J_o I_o^{-1}) \smallsetminus \mathscr{R}_{A,B}$, we have

$$\mu(A) + \mu(B) \leqslant m_Q(I_o) + m_Q(J_o) - m_Q(I_o \cap \tau_p(g)^{-1} J_o),$$

Moreover, there exists a G-factor map $\sigma : (X, \mu) \to (Q, m_Q)$, where G acts on Q via τ_p, such that

$$A \subseteq \sigma^{-1}(I_o) \quad \text{and} \quad B \subseteq \sigma^{-1}(J_o), \quad \textit{modulo } \mu\textit{-null sets.}$$

In the case when $A = B$, we can take $I = J$.

Proof. — By Lemma 2.4, we can find a G-invariant measurable μ-conull subset $X_3 \subseteq X$, a compact and metrizable abelian group K with Haar probability measure m_K, a homomorphism $\tau : G \to K$ with dense image, and two measurable subsets I and J of K, a G-equivariant measurable map $\pi : X_3 \to K$ such that $\pi_*(\mu|_{X_3}) = m_K$, where G acts on K via τ, such that

$$A \cap X_3 \subseteq \pi^{-1}(I) \quad \text{and} \quad B \cap X_3 \subseteq \pi^{-1}(J) \tag{2.6}$$

and

$$m_K(JI^{-1}) \leqslant \mu(A_x^{-1}B) \quad \text{and} \quad A_x^{-1}(B\cap X_3) \subseteq \pi^{-1}(\pi(x)I^{-1}J),$$

for all $x \in X_3$. Furthermore, by Lemma 2.1 and Lemma 2.2, there exist measurable μ-conull subsets X_1 and X_2 of X such that

$$\mathscr{R}_{A,B} = B_x A_x^{-1} \quad \text{and} \quad \mu(A_x^{-1}B) \leqslant \underline{d}(\mathscr{R}_{A,B})$$

and, for every $g \notin \mathscr{R}_{A,B}$,

$$\underline{d}(A_x \cup g^{-1}B_x) = \mu(A) + \mu(B) = \mu(A \cup g^{-1}B), \tag{2.7}$$

for all $x \in X_1 \cap X_2$. In particular, since $X_1 \cap X_2 \cap X_3$ is a μ-conull subset of X, and thus non-empty, we have

$$\mu(A) \leqslant m_K(I) \quad \text{and} \quad \mu(B) \leqslant m_K(J) \quad \text{and} \quad m_K(JI^{-1}) \leqslant \underline{d}(\mathscr{R}_{A,B}).$$

Let us now assume that Q is a compact group, $p : K \to Q$ is a continuous surjective homomorphism and I_o and J_o are measurable subsets of Q such that (I,J) reduces to (I_o,J_o). We recall that this means that

$$I \subset p^{-1}(I_o) \quad \text{and} \quad J \subset p^{-1}(J_o) \quad \text{and} \quad m_K(JI^{-1}) = m_Q(J_oI_o^{-1}).$$

Hence, $JI^{-1} \subset p^{-1}(J_oI_o^{-1})$, and

$$m_Q(J_oI_o^{-1}) \leqslant \mu(A_x^{-1}B) \quad \text{and} \quad A_x^{-1}(B \cap X_3) \subseteq \pi^{-1}(\pi(x)p^{-1}(J_oI_o^{-1})),$$

for all $x \in X_3$. We note that we can write

$$\pi^{-1}(\pi(x)p^{-1}(J_oI_o^{-1})) = \sigma^{-1}(\sigma(x)J_oI_o^{-1}),$$

for all $x \in X_3$, where $\sigma = p \circ \pi$, and thus

$$A_x^{-1}(B \cap X_3) \subseteq \sigma^{-1}(\sigma(x)J_oI_o^{-1}), \quad \text{for all } x \in X_3. \tag{2.8}$$

The map σ is a G-factor map from (X,μ) to (Q,m_Q), where G acts on Q via $\tau_p = p\circ\tau$, and it follows from (2.6) that

$$A \cap X_3 \subset \sigma^{-1}(I_o) \quad \text{and} \quad B \cap X_3 \subset \sigma^{-1}(J_o). \tag{2.9}$$

It remains to prove that

$$\mathscr{R}_{A,B} \subseteq \tau_p^{-1}(J_oI_o^{-1}),$$

and that for every $g \in \tau_p^{-1}(J_oI_o^{-1}) \smallsetminus \mathscr{R}_{A,B}$, we have

$$\mu(A) + \mu(B) \leqslant m_Q(I_o) + m_Q(J_o) - m_Q(I_o \cap \tau_p(g)^{-1}J_o). \tag{2.10}$$

To prove the inclusion, we first note that since X_3 is G-invariant, we have

$$\begin{aligned}(A_x^{-1}(B\cap X_3))_x = A_x^{-1}B_x &\subset \sigma^{-1}(\sigma(x)J_oI_o^{-1})_x \\ &= \tau_p^{-1}(\sigma(x)J_oI_o^{-1}\sigma(x)^{-1}) = \tau_p^{-1}(J_oI_o^{-1}),\end{aligned}$$

for all $x \in X_3$. To prove (2.10), we recall from (2.7) that if $g \notin \mathscr{R}_{A,B}$, then

$$\underline{d}(A_x \cup g^{-1}B_x) = \mu(A) + \mu(B) = \mu(A \cup g^{-1}B),$$

whence, by (2.9),

$$\begin{aligned}\underline{d}(A_x \cup g^{-1}B_x) &= \mu(A) + \mu(B) = \mu(A \cup g^{-1}B)\\ &\leqslant \mu(\sigma^{-1}(I_o) \cup g^{-1}\sigma^{-1}(J_o)) = m_Q(I_o \cup \tau_p(g)^{-1}J_o)\\ &= m_Q(I_o) + m_Q(J_o) - m_Q(I_o \cap \tau_p(g)^{-1}J_o),\end{aligned}$$

which finishes the proof. □

2.6. Classical product set theorems in compact groups. — We shall use the following two results about product sets in compact groups due to Kneser in his very influential paper [5].

Theorem 2.8 ([5, Satz 1]). — *Let K be a compact and metrizable abelian group with Haar probability measure m_K and suppose that I and J are measurable subsets of K with positive m_K-measures such that*

$$m_K(JI^{-1}) < m_K(I) + m_K(J).$$

Then JI^{-1} is a clopen subset of K, and there exist

- *a finite group Q and a homomorphism p from K onto Q.*
- *a pair (I_o, J_o) of subsets of Q with*

$$m_Q(J_oI_o^{-1}) = m_Q(J_o) + m_Q(I_o) - m_Q(\{e_Q\}),$$

such that (I,J) reduces to (I_o, J_o) with respect to p. If $I = J$, we can take $I_o = J_o$.

Corollary 2.9. — *Let K be a compact and metrizable abelian group with Haar probability measure m_K and assume that I is a measurable subset of K with positive m_K-measure such that*

$$m_K(II^{-1}) < \frac{3}{2}\, m_K(I).$$

Then there exist a finite group Q, a surjective homomorphism $p : K \to Q$ and a point $q \in Q$ such that (I,I) reduces to $(\{q\},\{q\})$ with respect to p. In particular, II^{-1} is an open subgroup of K.

Proof. — By Theorem 2.8, there exist a finite group Q, a homomorphism p from K onto Q and a subset I_o of Q, such that

$$I \subset p^{-1}(I_o) \quad \text{and} \quad II^{-1} = p^{-1}(I_oI_o^{-1}) \quad \text{and} \quad m_Q(I_oI_o^{-1}) = 2m_Q(I_o) - m_Q(\{e_Q\}).$$

Since $m_K(II^{-1}) < \frac{3}{2}m_K(I)$, we conclude that

$$m_Q(I_oI_o^{-1}) = 2m_Q(I_o) - m_Q(\{e_Q\}) < \frac{3}{2}\, m_Q(I_o), \tag{2.11}$$

whence $m_Q(I_oI_o^{-1}) < \frac{3}{2}m_Q(\{e_Q\})$. Since I_o is non-empty, we conclude that $I_oI_o^{-1}$ must be a point. Hence $I_o = \{q\}$ for some $q \in Q$. □

If K is connected and non-trivial, then there are no proper clopen subsets of K, whence the assumed upper bound in Theorem 2.8 can never occur.

Corollary 2.10. — *Let K be a compact, metrizable and connected abelian group with Haar probability measure m_K. Then, for all measurable subsets I and J of K,*

$$m_K(JI^{-1}) \geqslant \min\bigl(1, m_K(I) + m_K(J)\bigr).$$

In the connected case, Kneser further characterized the pairs of measurable subsets of the group for which the lower bound in Corollary (2.10) is attained. We denote by $\mathbb{T}$ the group $\mathbb{R}/\mathbb{Z}$ endowed with the quotient topology.

Theorem 2.11 ([5, Satz 2]). — *Let K be a compact, metrizable and connected abelian group with Haar probability measure m_K. If I and J are measurable subsets of K such that*

$$m_K(JI^{-1}) = m_K(I) + m_K(J) \leqslant 1,$$

then there exist

– *a continuous homomorphism p from K onto $\mathbb{T}$,*

– *closed intervals I_o and J_o in $\mathbb{T}$ with*

$$m_{\mathbb{T}}(I_o) = m_K(I) \quad \text{and} \quad m_{\mathbb{T}}(J_o) = m_K(J),$$

such that (I, J) reduces to (I_o, J_o) with respect to p.

3. Proof of Theorem 1.3 and Theorem 1.5

Let A and B be measurable subsets of X with positive μ-measures. The first assertion of Theorem 1.3 is contained in Corollary 2.3. Let us assume that either

$$A = B \quad \text{and} \quad \underline{d}(\mathscr{R}_A) < \frac{3}{2}\,\mu(A) \tag{3.1}$$

or

$$\underline{d}(\mathscr{R}_{A,B}) < \mu(A) + \mu(B). \tag{3.2}$$

By the first part of Proposition 2.7, there exist

– a compact and metrizable abelian group K with Haar probability measure m_K,

– a homomorphism $\tau : G \to K$ with dense image,

– a pair (I, J) of measurable subsets of K,

which satisfy

$$\mu(A) \leqslant m_K(I) \quad \text{and} \quad \mu(B) \leqslant m_K(B) \quad \text{and} \quad m_K(JI^{-1}) \leqslant \underline{d}(\mathscr{R}_{A,B}).$$

In the case (3.1), which corresponds to Theorem 1.3, we can take $I = J$, and thus

$$m_K(II^{-1}) \leqslant \underline{d}(\mathscr{R}_A) < \frac{3}{2}\,m_K(I).$$

and in the case (3.2), which corresponds to Theorem 1.5, we have

$$m_K(JI^{-1}) \leqslant \underline{d}(\mathscr{R}_{A,B}) < \mu(A) + \mu(B) \leqslant m_K(I) + m_K(J).$$

In both cases, Theorem 2.8 tells us that there exist a finite group Q, a continuous surjective homomorphism $p : K \to Q$ and a pair (I_o, J_o) of subsets of Q such that (I, J) reduces to (I_o, J_o) with respect to p. By Proposition 2.7, this implies that

$$\mathscr{R}_{A,B} \subset \tau_p^{-1}(J_o I_o^{-1}),$$

and that for all $g \in \tau_p^{-1}(J_o I_o^{-1}) \smallsetminus \mathscr{R}_{A,B}$,

$$\mu(A) + \mu(B) \leqslant m_Q(I_o) + m_Q(J_o) - m_Q(I_o \cap \tau_p(g)^{-1} J_o).$$

In the case (3.1), Corollary 2.9 further asserts that $I_o = J_o = \{q\}$ for some point $q \in Q$, whence $I_o I_o^{-1} = e_Q$ and thus we can conclude from above that $\mathscr{R}_A \subset G_o := \ker \tau_p$, and

$$m_Q(\{e_Q\}) = m_K(I_o I_o^{-1}) \leqslant \underline{d}(\mathscr{R}_{A,B}) < \frac{3}{2}\,\mu(A). \tag{3.3}$$

Since Q is finite, G_o has finite index in G and for every $g \in G_o \smallsetminus \mathscr{R}_A$, we have

$$m_Q(I_o \cap \tau_p(g)^{-1} I_o) \geqslant m_Q(\{e_Q\}).$$

Hence,

$$2\mu(A) \leqslant 2m_Q(I_o) - m_Q(I_o \cap \tau_p(g)^{-1} I_o) \leqslant m_Q(\{e_Q\}).$$

The last inequality clearly contradicts (3.3), so we conclude that $G_o = \mathscr{R}_A$, which finishes the proof of Theorem 1.3.

In the case of (3.2), Theorem 2.8 asserts that the pair (I_o, J_o) in Q satisfies

$$m_Q(J_o I_o^{-1}) = m_Q(I_o) + m_Q(J_o) - m_Q(\{e_Q\}),$$

whence

$$m_Q(I_o) + m_Q(J_o) - m_Q(\{e_Q\}) \leqslant \underline{d}(\mathscr{R}_{A,B}) < \mu(A) + \mu(B). \tag{3.4}$$

By Proposition 2.7, we have

$$\mathscr{R}_{A,B} \subset \tau_p^{-1}(J_o I_o^{-1})$$

and for all $g \in \tau_p^{-1}(J_o I_o^{-1}) \smallsetminus \mathscr{R}_{A,B}$, we have

$$\mu(A) + \mu(B) \leqslant m_Q(I_o) + m_Q(J_o) - m_Q(I_o \cap \tau_p(g)^{-1} J_o).$$

Since $g \in \tau_p^{-1}(J_o I_o^{-1})$ and Q is finite, we have

$$m_Q(I_o \cap \tau_p(g)^{-1} J_o) \geqslant m_Q(\{e_Q\}),$$

whence

$$\mu(A) + \mu(B) \leqslant m_Q(I_o) + m_Q(J_o) - m_Q(\{e_Q\}),$$

which clearly contradicts (3.4). We conclude that $\tau_p^{-1}(J_o I_o^{-1}) \smallsetminus \mathscr{R}_{A,B}$ is empty, and thus

$$\mathscr{R}_{A,B} = \tau_p^{-1}(J_o I_o^{-1}) = M G_o,$$

where $G_o = \ker \tau_p$, and M is a finite subset of G whose image under τ_p equals $J_o I_o^{-1}$. Since Q is finite, G_o has finite index in G. This proves Theorem 1.5 (with $\eta = \tau_p$).

4. Proof of Theorem 1.8

Suppose that $G \curvearrowright (X,\mu)$ is totally ergodic. Let A and B be measurable subsets of X with positive μ-measures, and assume that

$$\underline{d}(\mathscr{R}_{A,B}) = \mu(A) + \mu(B) < 1.$$

By the first part of Proposition 2.7, we can find

– a compact and metrizable abelian group K with Haar probability measure m_K,

– a homomorphism $\tau : G \to K$ with dense image,

– a pair (I,J) of measurable subsets of K,

which satisfy

$$\mu(A) \leqslant m_K(I) \quad \text{and} \quad \mu(B) \leqslant m_K(B) \quad \text{and} \quad m_K(JI^{-1}) \leqslant \underline{d}(\mathscr{R}_{A,B}).$$

Furthermore, since $G \curvearrowright (X,\mu)$ is totally ergodic, K must be connected. In particular, by Corollary 2.10,

$$\min(1, m_K(I) + m_K(J)) \leqslant m_K(JI^{-1}) \leqslant \underline{d}(\mathscr{R}_{A,B}) \leqslant \mu(A) + \mu(B) \leqslant m_K(I) + m_K(J).$$

If $m_K(I) + m_K(J) \geqslant 1$, then $m_K(JI^{-1}) = 1$, whence $\mu(A) + \mu(B) \geqslant 1$, which we have assumed away. Hence, $m_K(I) + m_K(J) < 1$, and thus $\mu(A) = m_K(I)$ and $\mu(B) = m_K(J)$, and

$$m_K(JI^{-1}) = m_K(I) + m_K(J) < 1.$$

Theorem 2.11 now asserts that there is a continuous surjective homomorphism $p : K \to \mathbb{T}$ and closed intervals I_o and J_o of $\mathbb{T}$ such that

$$\mu(A) = m_K(I) = m_{\mathbb{T}}(I_o) \quad \text{and} \quad \mu(B) = m_K(J) = m_{\mathbb{T}}(J_o),$$

and (I,J) reduces to the pair (I_o, J_o) with respect to p. Hence, by the second part of Proposition 2.7,

$$\mathscr{R}_{A,B} \subseteq \tau_p^{-1}(J_o I_o^{-1})$$

and for all $g \in \tau_p^{-1}(J_o I_o^{-1}) \smallsetminus \mathscr{R}_{A,B}$, we have

$$\mu(A) + \mu(B) \leqslant m_{\mathbb{T}}(I_o) + m_{\mathbb{T}}(J_o) - m_{\mathbb{T}}(I_o \cap \tau_p(g)^{-1} J_o).$$

We conclude that

$$m_{\mathbb{T}}(I_o \cap \tau_p(g)^{-1} J_o) = 0, \quad \text{for all } g \in \tau_p^{-1}(J_o I_o^{-1}) \smallsetminus \mathscr{R}_{A,B}.$$

Note that $J_o I_o^{-1}$ is a closed interval in $\mathbb{T}$. Hence $m_{\mathbb{T}}(I_o \cap \tau_p(g)^{-1} J_o) = 0$ for some $g \in \tau_p^{-1}(J_o I_o^{-1})$ if and only if $\tau_p(g)$ is one of the endpoints of this interval. In other words, $\mathscr{R}_{A,B}$ can only differ from the Sturmian set $\tau_p^{-1}(J_o I_o^{-1})$ by at most two cosets of the subgroup $\ker \tau_p$.

Furthermore, since $\mu(A) = m_Q(I_o)$ and $\mu(B) = m_Q(J_o)$, the last part of Proposition 2.7 asserts that there is a G-factor map $\sigma : (X,\mu) \to (\mathbb{T}, m_{\mathbb{T}})$, where G acts on $\mathbb{T}$ via τ_p, such that

$$A = \sigma^{-1}(I_o) \quad \text{and} \quad B = \sigma^{-1}(J_o),$$

modulo μ-null sets. This finishes the proof of Theorem 1.8 (with $\eta = \tau_p$).

5. Proof of Theorem 1.12

Let us first assume that $G \curvearrowright (X,\mu)$ is C-doubling for some $C \geqslant 1$. Then, for every $n \geqslant 1$, there is a measurable subset $A_n \subset X$ such that

$$0 < \mu(A_n) < \frac{1}{n} \quad \text{and} \quad \underline{d}(\mathscr{R}_{A_n}) \leqslant C\mu(A_n) < \frac{C}{n}.$$

To avoid trivialities, we shall from now on assume that $n > C$. By Lemma 2.4, we can find a (non-trivial) compact metrizable group K_n, a homomorphism $\eta_n : G \to K_n$ with dense image, a G-factor map $\pi_n : (X,\mu) \to (K_n, m_{K_n})$ and a measurable subset $I_n \subset K_n$ such that

$$A_n \subset \pi_n^{-1}(I_n) \quad \text{modulo null sets} \quad \text{and} \quad m_{K_n}(I_n^{-1}I_n) \leqslant \frac{C}{n}, \quad \text{for all } n \geqslant 1.$$

In particular, $m_{K_n}(I_n) \leqslant m_{K_n}(I_n^{-1}I_n) \leqslant \frac{C}{n}$. Let K denote the closure in $\prod_n K_n$ of the diagonally embedded subgroup $\{(\eta_n(g)) \mid g \in G\}$. We note that $\pi = (\pi_n) : (X,\mu) \to (K, m_K)$ is a G-factor map, where G acts on K via $\eta = (\eta_n)$. Since the pull-backs to K of the sets I_n provide measurable subsets of K with arbitrarily small m_K-measures, we see that K must be infinite.

Let us now assume that there exist

(i) an infinite compact metrizable group K and a homomorphism $\eta : G \to K$ with dense image.

(ii) a G-factor $\sigma : (X,\mu) \to (K, m_K)$, where m_K denotes the normalized Haar measure on K and G acts on K via η.

We wish to prove that (X,μ) is C-doubling for some $C \geqslant 1$. Since $G \curvearrowright (K, m_K)$ is a G-factor of (X,μ), it is clearly enough to prove that $G \curvearrowright (K, m_K)$ is C-doubling. If K^o has infinite index in K, then K/K^o is an infinite totally disconnected group, and thus we can find a decreasing sequence (U_n) of open subgroups of K with $m_K(U_n) = 1/[K:U_n] < 1/n$ for all n. Since $\mathscr{R}_{U_n} = \eta^{-1}(U_n)$, we have

$$\underline{d}(\mathscr{R}_{U_n}) = \frac{1}{[K:U_n]} = m_K(U_n), \quad \text{for all } n \geqslant 1,$$

which shows that $G \curvearrowright (X,\mu)$ is 1-doubling (we are using here that the sequence (F_n) also satisfy (1.1) for all bounded measurable functions on K). If K^o has finite index in K, then K^o is an open subgroup, and thus has positive m_K-measure. Fix a non-trivial continuous character $\chi : K^o \to \mathbb{T}$, and note that by connectedness, χ is onto. Set

$$I_n = \chi^{-1}\Big(\Big[-\frac{1}{2n}, \frac{1}{2n}\Big]\Big) \subset K^o \subset K, \quad \text{for } n \geqslant 1.$$

Then, $m_K(I_n) = m_K(K^o)/n$ for all n, and it is not hard to show that

$$\underline{d}(\mathscr{R}_{I_n}) = \underline{d}(\eta^{-1}(I_n - I_n)) \leqslant 2m_K(I_n), \quad \text{for all } n,$$

whence $G \curvearrowright (X,\mu)$ is 2-doubling.

References

[1] M. Björklund & A. Fish – "Approximate invariance for ergodic actions of amenable groups", *Discrete Anal.* (2019), article no. 6 (56 pages).

[2] M. Boshernitzan & M. Wierdl – "Ergodic theorems along sequences and Hardy fields", *Proc. Nat. Acad. Sci. U.S.A.* **93** (1996), no. 16, p. 8205–8207.

[3] J. Bourgain – "On the maximal ergodic theorem for certain subsets of the integers", *Israel J. Math.* **61** (1988), no. 1, p. 39–72.

[4] M. Kneser – "Abschätzung der asymptotischen Dichte von Summenmengen", *Math. Z.* **58** (1953), p. 459–484.

[5] ———, "Summenmengen in lokalkompakten abelschen Gruppen", *Math. Z.* **66** (1956), p. 88–110.

[6] Y. Lacroix – "Natural extensions and mixing for semi-group actions", *Publ. Inst. Rech. Math. Rennes* **2** (1995), article no. 7 (10 pages).

Manuscript received 4th May 2020
accepted 26th January 2021

Michael Björklund, Department of Mathematics, Chalmers University
Chalmers Tvärgata 3, SE-412 96 Gothenburg, Sweden
E-mail : `micbjo@chalmers.se`
Url : `http://www.math.chalmers.se/~micbjo/`

Alexander Fish, The School of Mathematics and Statistics, Faculty of Science, The University of Sydney
NSW 2006, Sydney, Australia
E-mail : `afish@maths.usyd.edu.au`
Url : `https://www.maths.usyd.edu.au/u/afish/`

Ilya D. Shkredov, Steklov Mathematical Institute
ul. Gubkina 8, 119991 Moscow, Russia
E-mail : `ilya.shkredov@gmail.com`
Url : `http://www.mi-ras.ru/~ishkredov/`