Proceedings of the Edinburgh Mathematical Society (2013) **56**, 657–666
DOI:10.1017/S0013091512000399

A REMOVAL LEMMA FOR LINEAR CONFIGURATIONS IN SUBSETS OF THE CIRCLE

PABLO CANDELA[1] AND OLOF SISASK[2]

[1] *Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, UK* (pc308@cam.ac.uk)
[2] *School of Mathematical Sciences, Queen Mary, University of London, Mile End Road, London E1 4NS, UK* (o.sisask@qmul.ac.uk)

(Received 9 August 2011)

Abstract We obtain a removal lemma for systems of linear equations over the circle group, using a similar result for finite fields due to Král′, Serra and Vena, and we discuss some applications.

Keywords: removal lemma; torus; systems of linear equations

2010 *Mathematics subject classification:* Primary 11B30
Secondary 11B75

1. Introduction

If a subset of an abelian group contains very few linear configurations of some given type, then one needs to delete only a few elements from the set in order to remove all such configurations. This is the moral of so-called *arithmetic removal lemmas*. For example, if A is a subset of a cyclic group $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$ containing only δN^2 of its own sums (i.e. solutions to $a_1 + a_2 = a_3$), then one can make A completely sum-free by deleting only $\delta' N$ of its elements, where δ' depends only on δ, and $\delta' \to 0$ as $\delta \to 0$. In [**5**] Green proved a result of this type dealing with the removal of solutions to a single linear equation over an arbitrary finite abelian group. Green raised the question of whether similar results held for systems of equations, noting that the Fourier analytic methods employed in [**5**] did not extend to answering this. Shapira [**12**] and (independently) Král′ *et al*. [**7**] used hypergraph removal results to obtain the following extension, dealing with systems of linear equations over finite fields.

Theorem 1.1. *Let $r \leqslant m$ be positive integers and let $\epsilon > 0$. There exists $\delta > 0$ such that the following holds. Let $\mathbb{F}$ be the finite field of order q, let L be an $r \times m$ matrix with coefficients in $\mathbb{F}$ of rank r over $\mathbb{F}$, and suppose $A_1, \dots, A_m \subset \mathbb{F}$ satisfy*

$\mathbb{E}_{x\in\ker L} 1_{A_1}(x_1)\cdots 1_{A_m}(x_m) \leqslant \delta$. *Then there are sets* $E_1,\dots,E_m \subset \mathbb{F}$ *of cardinality at most* ϵq *such that* $(A_1 \setminus E_1) \times \cdots \times (A_m \setminus E_m) \cap \ker L = \emptyset$.*

Our aim here is to obtain a continuous analogue of Theorem 1.1, replacing finite fields with the circle group $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. Previous extensions of discrete additive-combinatorial results to the latter setting include the analogues of the Cauchy–Davenport inequality obtained by Raikov [**10**] and Macbeath [**9**] (see the excellent notes of Ruzsa [**11**] for a more detailed account of this topic) and Lev's work [**8**] on sum-free sets in $\mathbb{T}$.

To state our main result let us set up some notation. For any compact abelian group G we denote the normalized Haar measure on G by μ_G. We denote the closed subgroup $\{x \in G^m : Lx = 0\}$ of the direct product G^m by $\ker_G L$, and to abbreviate the notation we denote by μ_L the normalized Haar measure on $\ker_G L$. For measurable functions $f_1, f_2, \dots, f_m : G \to \mathbb{C}$ we define

$$S_L(f_1,\dots,f_m) = \int_{\ker_G L} f_1(x_1)\cdots f_m(x_m)\,\mathrm{d}\mu_L(x). \tag{1.1}$$

(Throughout the paper 'measurable' refers to Borel measurability.) If each f_i is the indicator 1_{A_i} of a measurable set $A_i \subset \mathbb{T}$, then (1.1) becomes simply $S_L(A_1,\dots,A_m) = \mu_L(A_1 \times \cdots \times A_m \cap \ker_G L)$. We refer to the latter quantity as the *solution measure* of the sets A_i. When $A_i = A$ for all $i \in [m] = \{1,2,\dots,m\}$, we write $S_L(A)$ for the solution measure. If the group G has to be specified to avoid confusion, we shall write $\mu_{L,G}$, $S_{L,G}$ instead of μ_L, S_L. The main result, then, is the following.

Theorem 1.2. *Let* L *be an* $r\times m$ *matrix of integers, of full rank* r. *For any* $\epsilon > 0$, *there exists* $\delta = \delta(L,\epsilon) > 0$ *such that the following holds. If* $A_1,\dots,A_m$ *are measurable subsets of* $\mathbb{T}$ *such that* $S_L(A_1,\dots,A_m) \leqslant \delta$, *then there are measurable sets* $E_1,\dots,E_m \subset \mathbb{T}$ *with* $\mu_{\mathbb{T}}(E_i) \leqslant \epsilon$ *for all* $i \in [m]$, *such that* $(A_1 \setminus E_1) \times \cdots \times (A_m \setminus E_m) \cap \ker_{\mathbb{T}} L = \emptyset$.

For completeness we also prove the following variant concerning sets with zero solution measure, which has a much simpler proof.

Proposition 1.3. *Let* L *be an* $r \times m$ *matrix of integers, of full rank* r, *and suppose* $A_1,\dots,A_m$ *are measurable subsets of* $\mathbb{T}$ *such that* $S_L(A_1,\dots,A_m) = 0$. *Then there are null sets* $E_1,\dots,E_m \subset \mathbb{T}$ *such that* $(A_1 \setminus E_1) \times \cdots \times (A_m \setminus E_m) \cap \ker_{\mathbb{T}} L = \emptyset$. *We can take* $A_i \setminus E_i$ *to be the set of Lebesgue density points of* A_i.

We now discuss briefly some consequences of these results. We say an integer matrix L is *invariant* if its columns sum to zero for the constant vector $\mathbf{1} = (1,1,\dots,1)$. In this case the system $Lx = 0$ is translation invariant in the sense that given any abelian group G, for any $x = (x_1,\dots,x_m) \in G^m$ and $t \in G$, we have $Lx = 0$ if and only if $L(x_1+t,\dots,x_m+t) = 0$. In particular, for any $t \in G$ the element $x = (t,\dots,t)$ is a solution of the system. Therefore, Proposition 1.3 implies that if L is invariant then any set $A \subset \mathbb{T}$ of positive measure has $S_L(A) > 0$. However, the latter positive quantity may depend on the set A. By contrast, Theorem 1.2 implies the following analogue of Szemerédi's Theorem [**13**, Theorem 11.1] for translation-invariant systems on $\mathbb{T}$.

* One can phrase this equivalently with ϵ being a function of δ that vanishes as $\delta \to 0$, as in the initial example.

Theorem 1.4. *Let L be an invariant $r \times m$ integer matrix of full rank r. Then for any $\alpha > 0$, there exists $c = c(\alpha, L) > 0$ such that for any measurable set $A \subset \mathbb{T}$, of measure at least α, we have $S_L(A) \geqslant c$.*

For instance, since arithmetic progressions of arbitrary fixed length are translation invariant, any subset of the circle of positive measure α contains a positive measure c of such progressions, where c depends on α but not on the particular subset.

At the end of the paper we discuss another application of Theorem 1.2, related to the role that groups such as the circle can play as limit objects for certain additive-combinatorial problems.

The paper is structured as follows. Our proof of Theorem 1.2 reduces the problem to the discrete case, where one can appeal to Theorem 1.1. This involves first approximating each set A_i by a simpler set that can be viewed as a subset A_i' of a cyclic group $\mathbb{Z}_p$ for p a prime. This is done in § 2. The relationship between the solution measure of the approximating sets and the solution counts on $\mathbb{Z}_p$ of the sets A_i' is captured in Lemma 2.5. This relationship is somewhat subtle, in that expressing the solution measure in terms of the latter discrete solution counts requires many different shifts of the set $A_1' \times \cdots \times A_m'$, each shift having a corresponding weight. We then require some control on these weights, which is obtained in § 3 using a simple geometric characterization of $\ker_{\mathbb{T}} L$ and its measure μ_L. The proof of Theorem 1.2 is then completed in § 4, where we also deduce Theorem 1.4 and prove Proposition 1.3. Finally, we close with the aforementioned application and some further remarks in § 5.

2. A discrete decomposition of the solution measure

2.1. Approximating measurable sets

For any positive integer N, we refer to the partition $\mathbb{T} = \bigsqcup_{x \in [N]} [(x-1)/N, x/N)$ as the *N-partition of* $\mathbb{T}$, and we say $A \subset \mathbb{T}$ is *N-measurable* if A is a union of intervals from the N-partition. The aim in this subsection is to show that, for the proof of Theorem 1.2, the sets A_i can be assumed to be p-measurable for some large prime p.

Lemma 2.1. *Let L be an $r \times m$ matrix of integers, of full rank r, such that any $r \times (m-1)$ submatrix of L also has rank r. Let $\delta > 0$ and let $C_1, \dots, C_m$ be measurable subsets of $\mathbb{T}$. Then for any large $p \in \mathbb{N}$, there exist p-measurable sets $A_i \subset \mathbb{T}$ such that $\mu_{\mathbb{T}}(C_i \Delta A_i) \leqslant \delta/m$ for all $i \in [m]$, and $|S_L(C_1, \dots, C_m) - S_L(A_1, \dots, A_m)| \leqslant \delta$.*

The submatrix condition in this lemma can be assumed without loss of generality when proving Theorem 1.2. Indeed, suppose that deleting column j from L yields a matrix L' of rank $r-1$. Then for some non-zero vector $v \in \mathbb{Z}^r$, we have $v^{\mathrm{T}} L' = 0$. Since L has rank r, the jth entry of $v^{\mathrm{T}} L$ must be a non-zero integer ℓ. Now if $x \in A_1 \times \cdots \times A_m$ satisfies $Lx = 0$, then in $\mathbb{T}$ we have $\ell x_j = (v^{\mathrm{T}} L) \cdot x = v^{\mathrm{T}} \cdot (Lx) = 0$. Therefore, we can delete all such solutions x by removing the finite set $\{a \in A_j : \ell a = 0\}$ from A_j, so Theorem 1.2 is clearly true for this system.

To prove Lemma 2.1 we use the following basic result, which will also be used later.

Lemma 2.2. *Let G be a locally compact abelian group with a Haar measure μ and let H be a closed subgroup of G^m with a Haar measure μ_H such that, for some $i \in [m]$, the projection $\pi\colon H \to G$, $x \mapsto x_i$ is surjective. Then there is a constant $c > 0$ such that, for any functions $f_1, \dots, f_m\colon G \to \mathbb{C}$ with $\|f_j\|_{L_\infty} \leqslant 1$ for all j, we have*

$$\left| \int_H f_1(x_1) f_2(x_2) \cdots f_m(x_m) \, \mathrm{d}\mu_H(x) \right| \leqslant c \|f_i\|_{L_1}.$$

If G, H are compact abelian groups and μ_G, μ_H are their respective unique probability Haar measures, then we can take $c = 1$.

Proof. The left-hand side above is at most

$$\int_H |f_i(x_i)| \, \mathrm{d}\mu_H(x) = \int_H |f_i \circ \pi(x)| \, \mathrm{d}\mu_H(x).$$

The map π is a continuous surjective homomorphism from H to G, whence the measure $\mu_H \circ \pi^{-1}$ is a Haar measure on G, so by uniqueness there exists $c > 0$ such that $\mu_H \circ \pi^{-1} = c\mu_G$, and $c = 1$ if μ_G, μ_H are both probability measures. It follows that

$$\int_H |f_i \circ \pi(x)| \, \mathrm{d}\mu_H(x) = c \int_G |f_i(y)| \, \mathrm{d}\mu_G(y) = c \|f_i\|_{L_1}.$$

$\square$

Proof of Lemma 2.1. First, it follows from measure theory that for sufficiently large p there exist p-measurable sets, A_i, such that $\mu_{\mathbb{T}}(C_i \Delta A_i) \leqslant \delta/m$ for all i (say, by first approximating by unions of dyadic intervals and then approximating these by p-intervals). Now, by the multilinearity of S_L we have

$$\begin{aligned} &|S_L(C_1, \dots, C_m) - S_L(A_1, \dots, A_m)| \\ &\qquad\qquad \leqslant \sum_{i \in [m]} |S_L(1_{A_1}, \dots, 1_{A_{i-1}}, 1_{C_i} - 1_{A_i}, 1_{C_{i+1}}, \dots, 1_{C_m})|, \end{aligned}$$

and the assumption that every $r \times (m-1)$ submatrix of L has rank r is easily seen to imply that each projection $\ker_{\mathbb{T}} L \to \mathbb{T}$, $x \mapsto x_i$ is surjective, whence by Lemma 2.2 the ith summand above is at most $\|1_{C_i} - 1_{A_i}\|_{L_1(\mathbb{T})} \leqslant \mu_{\mathbb{T}}(C_i \Delta A_i) \leqslant \delta/m$. $\square$

2.2. The main formula

From now on, given a p-measurable set $A \subset \mathbb{T}$, we denote by A' the subset of $\mathbb{Z}_p$ defined by $1_{A'}(x) = 1_A(x/p)$. In order to apply Theorem 1.1, we express $S_L(A)$ in terms of solution measures in $\mathbb{Z}_p$ involving A'. This is done in Lemma 2.5.

For any positive integer p, let $\Lambda = \Lambda(p)$ denote the discrete torus $\mathbb{Z}_p^m/p \leqslant \mathbb{T}^m$, with elements denoted $j/p = (j(1)/p, \dots, j(m)/p)$, $j \in \mathbb{Z}_p^m$.

Definition 2.3. For any $r \times m$ integer matrix L and any positive integer p, we define

$$J = J(L, p) = \{j/p \in \Lambda \colon \mu_L((j/p + [0, 1/p)^m) \cap \ker_{\mathbb{T}} L) > 0\}.$$

The fact that J consists of $O_L(1)$ shifts of $\Lambda \cap \ker_{\mathbb{T}} L$ is central to the whole argument.

Lemma 2.4. *For some $K_L > 0$ depending only on L, for any p there exist elements $j_1/p, j_2/p, \dots, j_K/p \in J(L,p)$, with $K \leqslant K_L$, such that $J = \bigsqcup_{k\in[K]}(j_k/p + (\Lambda \cap \ker_{\mathbb{T}} L))$.*

Proof. If $j/p \in J$, then $L(j/p)$ lies in $-(L([0,1)^m) \cap \mathbb{Z}^r)/p \mod 1$. The cardinality of the latter set is bounded above in terms of L alone. Choosing $j_1/p, \dots, j_K/p \in J$ such that L is a bijection from $\{j_1/p, \dots, j_K/p\}$ to $L(J)$, we then have $K = O_L(1)$, and the result follows since $J \subset L^{-1}(L(J))$ and $L^{-1}(L(j_k/p)) \cap \Lambda = j_k/p + (\Lambda \cap \ker_{\mathbb{T}} L)$. □

We can now prove the main formula.

Lemma 2.5. *Let L be an $r \times m$ matrix of integers of full rank r, let p be a large prime, and let $A_1, \dots, A_m$ be p-measurable subsets of $\mathbb{T}$. Then there exist $j_1/p, \dots, j_K/p \in J$, with $K \leqslant K_L$, such that*

$$S_{L,\mathbb{T}}(A_1, \dots, A_m) = \sum_{k\in[K]} \lambda_k S_{L,\mathbb{Z}_p}(A_1' - j_k(1), \dots, A_m' - j_k(m)), \tag{2.1}$$

where $\lambda_k = p^{m-r}\mu_{L,\mathbb{T}}((j_k/p + [0,1/p)^m) \cap \ker_{\mathbb{T}} L)$.

Proof. We have $S_{L,\mathbb{T}}(A_1, \dots, A_m)$ equal to

$$\mu_L((A_1 \times \cdots \times A_m) \cap \ker_{\mathbb{T}} L) = \sum_{j \in A_1' \times \cdots \times A_m'} \mu_L((j/p + [0,1/p)^m) \cap \ker_{\mathbb{T}} L).$$

By the definition of the set J, this sum can be restricted directly to the shifts $j_1/p + \ker_{\mathbb{T}} L, \dots, j_K/p + \ker_{\mathbb{T}} L$ occurring in Lemma 2.4; since the subgroup $\Lambda \cap \ker_{\mathbb{T}} L$ of $\mathbb{T}^m$ is clearly isomorphic to the subgroup $\ker_{\mathbb{Z}_p} L$ of $\mathbb{Z}_p^m$, we see that $S_{L,\mathbb{T}}(A_1, \dots, A_m)$ equals

$$\begin{aligned}\sum_{k\in[K]} \sum_{j \in A_1' \times \cdots \times A_m' \cap (j_k + \ker_{\mathbb{Z}_p} L)} & \mu_L((j/p + [0,1/p)^m) \cap \ker_{\mathbb{T}} L) \\ &= \sum_{k\in[K]} \sum_{j \in (A_1' \times \cdots \times A_m' - j_k) \cap \ker_{\mathbb{Z}_p} L} \mu_L(((j + j_k)/p + [0,1/p)^m) \cap \ker_{\mathbb{T}} L).\end{aligned}$$

By invariance of μ_L under translation by $j/p \in \ker_{\mathbb{T}} L$, this equals

$$\sum_{k\in[K]} \sum_{j \in (A_1' \times \cdots \times A_m' - j_k) \cap \ker_{\mathbb{Z}_p} L} \mu_L((j_k/p + [0,1/p)^m) \cap \ker_{\mathbb{T}} L),$$

and (2.1) follows. □

3. A positive lower bound for the weights λ_k

For each $j \in \mathbb{Z}_p^m$, let $\lambda(j) = p^{m-r}\mu_L((j/p+[0,1/p)^m)\cap\ker_{\mathbb{T}} L)$. In order to use Lemma 2.5, we require that the weights λ_k be bounded away from 0, uniformly over p. Such a bound is guaranteed by the following result.

https://doi.org/10.1017/S0013091512000399 Published online by Cambridge University Press

Lemma 3.1. *Let L be an $r \times m$ matrix of integers of full rank r. Then there exists $\lambda^* > 0$ depending only on L such that for any large positive integer p, for any $j/p \in J(L,p)$, we have $\lambda(j) \geqslant \lambda^*$.*

The proof relies on a compactness argument coupled with the geometric characterization of μ_L given in Lemma 3.3. In what follows we always consider $\mathbb{T}^m$ as the set $[0,1)^m \subset \mathbb{R}^m$ with coordinatewise addition modulo 1 (and with the quotient topology on $\mathbb{R}^m/\mathbb{Z}^m$). Then $\ker_{\mathbb{T}} L$ is the closed subgroup $\{x \in [0,1)^m \colon Lx \in \mathbb{Z}^r\} \leqslant \mathbb{T}^m$. This subgroup is described more precisely by the following simple result.

Lemma 3.2. *Let $x_1, \dots, x_M$ be a choice of points in $[0,1)^m$ such that the linear map L over $\mathbb{R}$ gives a bijection $\{x_i \colon i \in [M]\} \to L([0,1)^m) \cap \mathbb{Z}^r$. Then we have the partition*

$$\ker_{\mathbb{T}} L = \bigsqcup_{i \in [M]} ((x_i +_{\mathbb{R}} \ker_{\mathbb{R}} L) \cap [0,1)^m). \tag{3.1}$$

Here we use $+_{\mathbb{R}}$ to denote addition in $\mathbb{R}^m$ (or more generally addition over $\mathbb{R}$), to distinguish it from addition in $\mathbb{T}^m$, which we may denote by $+_{\mathbb{T}}$. We now use (3.1) to relate the Haar measure μ_L to the $(m-r)$-dimensional Lebesgue measure on $\ker_{\mathbb{R}} L$, which we denote $\mu_{L,\mathbb{R}}$.

Lemma 3.3. *For any Borel set $A \subset \ker_{\mathbb{T}} L$, let $A^{(i)} := A \cap (x_i +_{\mathbb{R}} \ker_{\mathbb{R}} L)$ for each $i \in [M]$. Then there is a constant $c_L > 0$ such that $\mu_{L,\mathbb{T}}(A) = c_L \sum_i \mu_{L,\mathbb{R}}(A^{(i)} -_{\mathbb{R}} x_i)$.*

Proof. Let G denote the group $\{x \in \mathbb{R}^m \colon Lx \in \mathbb{Z}^r\}$. This is a closed subgroup of $\mathbb{R}^m$, and $H := \mathbb{Z}^m \leqslant G$. Clearly, we may identify $\ker_{\mathbb{T}} L$ with G/H. Thus, in the notation from (3.1), we have $G = (\bigsqcup_{i \in [M]} x_i + \ker_{\mathbb{R}} L) + \mathbb{Z}^m$, so we may write $G = \bigsqcup_{z \in Z}(z + \ker_{\mathbb{R}} L)$ for some collection $Z \subset \bigcup_i \{x_i\} + \mathbb{Z}^m$ containing the x_i. It is then easy to verify that any Haar measure on G must be a multiple of

$$\mu_G(A) := \sum_{z \in Z} \mu_{L,\mathbb{R}}(A^{(z)} - z), \tag{3.2}$$

where $A^{(z)} = A \cap (z + \ker_{\mathbb{R}} L)$, as may be seen by restricting to $\ker_{\mathbb{R}} L$. Endowing H with counting measure, by the quotient integral formula [**2**, Theorem 1.5.2] there is an invariant Radon measure $\mu_{G/H} \neq 0$ on G/H such that

$$\int_G f \, \mathrm{d}\mu_G = \int_{G/H} \sum_{n \in \mathbb{Z}^m} f(x +_{\mathbb{R}} n) \, \mathrm{d}\mu_{G/H}(x) \tag{3.3}$$

for any $f \in L_1(G)$. By the uniqueness of Haar measure we have $\mu_{L,\mathbb{T}} = c_L \mu_{G/H}$ for some constant $c_L > 0$. Now, given a Borel subset A of $\ker_{\mathbb{T}} L$, the function $f = 1_A$ on G is integrable, and the function $\sum_{n \in \mathbb{Z}^m} f(x+n)$ on G/H is simply 1_A, whence by (3.3) we have $\mu_{G/H}(A) = \int_G 1_A(x) \, \mathrm{d}\mu_G(x)$ and by (3.2) this is $\sum_{i \in [M]} \mu_{L,\mathbb{R}}(A^{(i)} - x_i)$. $\square$

Lemma 3.1 follows immediately from the following result.

Lemma 3.4. *There exists a finite set Λ^* of positive real numbers, depending only on L, such that for all large positive integers p we have $\{\lambda(j) \colon j \in J(L,p)\} \subset \Lambda^*$.*

Proof. First we show that there is a finite set $U \subset \ker_{\mathbb{T}} L$, depending only on L, such that for any large p and $j/p \in J$ there exist $v \in \mathbb{Z}^m$ and $u \in U$ such that

$$\lambda(j) = c_L \mu_{L,\mathbb{R}}((u +_{\mathbb{R}} v +_{\mathbb{R}} [0,1)^m) \cap \ker_{\mathbb{R}} L) > 0, \tag{3.4}$$

where c_L is the constant from Lemma 3.3. For p large enough depending only on L, by (3.1) the set $(j/p + [0,1/p)^m) \cap \ker_{\mathbb{T}} L$ lies entirely in $x_i +_{\mathbb{R}} \ker_{\mathbb{R}} L$ for some $i \in [M]$, so

$$\begin{aligned}\lambda(j) &= p^{m-r} \mu_L((j/p + [0,1/p)^m) \cap \ker_{\mathbb{T}} L)\\ &= c_L p^{m-r} \mu_{L,\mathbb{R}}((j/p -_{\mathbb{R}} x_i +_{\mathbb{R}} [0,1/p)^m) \cap \ker_{\mathbb{R}} L)\\ &= c_L \mu_{L,\mathbb{R}}((j -_{\mathbb{R}} px_i +_{\mathbb{R}} [0,1)^m) \cap \ker_{\mathbb{R}} L),\end{aligned}$$

where $j \in \mathbb{Z}^m$. Now let $U = \bigcup_{i \in [M]} \{-px_i \bmod 1 \colon p \in \mathbb{N}\} \subset \ker_{\mathbb{T}} L$. This is a finite subset of $\ker_{\mathbb{T}} L$ if we take the x_i to have rational coordinates (as we do). For any $j/p \in J$, we then have $j -_{\mathbb{R}} px_i = u +_{\mathbb{R}} v$ for some $u \in U$ and $v \in \mathbb{Z}^m$, whence (3.4) follows.

Now, by the translation invariance of $\mu_{L,\mathbb{R}}$ by elements of $\ker_{\mathbb{R}} L$, the measure in (3.4) depends only on $L(u+v)$. But if this measure is positive, then $L(u+v)$ is contained in the finite set $L(\bigcup_{w \in U} w + \mathbb{Z}^m) \cap -L([0,1)^m)$. Hence there are only finitely many possible values for the left-hand side of (3.4). □

4. Proofs of the main results

Recall that whenever A is a p-measurable subset of $\mathbb{T}$ we denote by A' the corresponding subset of $\mathbb{Z}_p$ defined by $1_{A'}(x) = 1_A(x/p)$.

Proof of Theorem 1.2. Given the matrix L, fix $\epsilon > 0$, let $\lambda^* > 0$ be the lower bound given by Lemma 3.1, and let $K_L > 0$ be as defined in Lemma 2.4. Let $\delta' > 0$ be such that Theorem 1.1 holds with initial parameter $\epsilon/2K_L$, and let $\delta = \min(\delta'\lambda^*, \epsilon)/2$. Now let $A_i \subset \mathbb{T}$, $i \in [m]$, be any Borel sets satisfying $S_L(A_1, \dots, A_m) \leqslant \delta$. Applying Lemma 2.1, we can assume that the given sets A_i are p-measurable for some large prime p, up to an error of measure $\delta/m \leqslant \epsilon/2$ for each set, and such that $S_L(A_1, \dots, A_m) \leqslant 2\delta \leqslant \delta'\lambda^*$. It follows from (2.1) and the lower bound $\lambda_k \geqslant \lambda^*$ that for some $K \leqslant K_L$ and each $k \in [K]$ we have $S_{L,\mathbb{Z}_p}(A'_1 - j_k(1), \dots, A'_m - j_k(m)) \leqslant \delta'$, and so Theorem 1.1 gives us subsets $E_{k,1}, \dots, E_{k,m}$ of $\mathbb{Z}_p$ of cardinality at most $\epsilon p/2K_L$, such that

$$((A'_1 \setminus E_{k,1}) - j_k(1)) \times \dots \times ((A'_m \setminus E_{k,m}) - j_k(m)) \cap \ker_{\mathbb{Z}_p} L = \emptyset. \tag{4.1}$$

Now for each $i \in [m]$, define the p-measurable set $E_i = \bigcup_{k \in [K]} (E_{k,i}/p + [0,1/p))$, and note that $\mu_{\mathbb{T}}(E_i) \leqslant \epsilon/2$. Finally, let $\varDelta$ be the null set $\mathbb{Z}_p/p$ in $\mathbb{T}$. We now claim that

$$\prod_{i \in [m]} A_i \setminus (E_i \cup \varDelta) \cap \ker_{\mathbb{T}} L = \emptyset.$$

Suppose for a contradiction that this set is non-empty, containing some point x. Then by the p-measurability of the sets $A_i \setminus E_i$ and the definition of $\varDelta$, letting j denote the

https://doi.org/10.1017/S0013091512000399 Published online by Cambridge University Press

point $(\lfloor px_1 \rfloor, \ldots, \lfloor px_m \rfloor) \in \mathbb{Z}_p^m$, we have

$$\prod_{i\in[m]} A_i \setminus (E_i \cup \Delta) \supset j/p + (0, 1/p)^m \ni x.$$

But then $(j/p+(0,1/p)^m)\cap \ker_{\mathbb{T}} L$ is a non-empty open subset of $\ker_{\mathbb{T}} L$, so this set must have positive μ_L-measure, and so $j/p \in J$. Then, by the covering of J in Lemma 2.4, there exists $k \in [K]$ such that $j \in j_k + \ker_{\mathbb{Z}_p} L$, and so $j - j_k$ belongs to $\prod_i((A_i' \setminus E_i') - j_k(i)) \cap \ker_{\mathbb{Z}_p} L$, contradicting (4.1). □

We can now quickly deduce Theorem 1.4. We say $A \subset \mathbb{T}$ is *L-free* if $A^m \cap \ker_{\mathbb{T}} L = \emptyset$.

Proof of Theorem 1.4. Let c be a positive value of δ such that Theorem 1.2 holds with initial parameter $\epsilon = \alpha/2m$. Suppose $S_L(A) \leqslant c$. Then by Theorem 1.2 there exists a measurable set $E \subset A$ such that $A \setminus E$ is L-free and $\mu_{\mathbb{T}}(E) \leqslant \mu_{\mathbb{T}}(E_1)+\cdots+\mu_{\mathbb{T}}(E_m) \leqslant \alpha/2$. Since for any $a \in A \setminus E$ the constant element $(a, \ldots, a) \in \mathbb{T}^m$ is in $\ker_{\mathbb{T}} L$, we must have $A \setminus E = \emptyset$, and therefore $\mu_{\mathbb{T}}(A) = \mu_{\mathbb{T}}(E) < \alpha$. □

While Theorem 1.4 follows very easily from Theorem 1.2, one can in fact simplify the overall argument somewhat if one is only interested in the former theorem; see the first remark in the next section.

Proof of Proposition 1.3. For each $i \in [m]$ let D_i denote the set of Lebesgue density points of A_i. Suppose for a contradiction that there exists some point x in $D_1 \times \cdots \times D_m \cap \ker_{\mathbb{T}} L$, and fix $\epsilon > 0$. By the Lebesgue density theorem, there exists $\delta > 0$ such that, letting Q denote the cube centred on x and of side length δ, we have $\mu_{\mathbb{T}}(D_i \cap \pi_i Q) \geqslant (1-\epsilon)\delta$ for all i (where π_i denotes projection to the ith component on $\mathbb{T}^m$). Now, by Lemma 3.2 and the characterization of $\mu_{L,\mathbb{T}}$, setting $C_i := D_i \cap \pi_i Q$ for each i, there exists a constant $c_L > 0$ such that

$$\mu_L(C_1 \times \cdots \times C_m \cap \ker_{\mathbb{T}} L) \geqslant c_L \delta^{m-r} \mu_{L,\mathbb{R}}(B_1 \times \cdots \times B_m \cap \ker_{\mathbb{R}} L),$$

where $B_i \subset [-\frac{1}{2}, \frac{1}{2})$ is the dilation by δ^{-1} of the set $B_i' - x_i$, when the latter is viewed as a subset of $I := [-\frac{1}{2}, \frac{1}{2}] \subset \mathbb{R}$. We claim that the large density of each B_i inside I implies $\mu_{L,\mathbb{R}}(B_1 \times \cdots \times B_m \cap \ker_{\mathbb{R}} L) > 0$, which gives a contradiction. Indeed, by multilinearity and Lemma 2.2 we have that $|\mu_{L,\mathbb{R}}(I^m \cap \ker_{\mathbb{R}} L) - \mu_{L,\mathbb{R}}(B_1 \times \cdots \times B_m \cap \ker_{\mathbb{R}} L)|$ is at most

$$\begin{aligned}\sum_{i\in[m]} \left| \int_{\ker_{\mathbb{R}} L} 1_{B_1}(x_1) \ldots 1_{B_{i-1}}(x_{i-1})(1_I - 1_{B_i})(x_i) 1_I(x_{i+1}) \ldots 1_I(x_m)\, \mathrm{d}\mu_{L,\mathbb{R}}(x) \right| \\ \leqslant c \sum_{i\in[m]} \|1_I - 1_{B_i}\|_{L_1(\mathbb{R})} \\ \leqslant cm\epsilon.\end{aligned}$$

Setting $\epsilon = \mu_{L,\mathbb{R}}(I^m \cap \ker_{\mathbb{R}} L)/2cm$ yields the claim. Note that the measure here is strictly positive since $I^m \cap \ker_{\mathbb{R}} L$ contains a non-empty open set. (In fact $\mu_{L,\mathbb{R}}(I^m \cap \ker_{\mathbb{R}} L) \geqslant 1$ by Vaaler's Theorem [**14**].) □

5. Remarks

The precision of Lemma 2.5 is not required for a proof of Theorem 1.4 per se; one can make do with a simpler inequality of the form $S_{L,\mathbb{T}}(A_1,\dots,A_m) \gg_L S_{L,\mathbb{Z}_p}(A'_1,\dots,A'_m)$. (If L is invariant, one can also apply Vaaler's Theorem to obtain the more precise inequality $S_{L,\mathbb{T}}(A_1,\dots,A_m) \geqslant S_{L,\mathbb{Z}_p}(A'_1,\dots,A'_m)$ for p-measurable sets A_i.) On the other hand, the non-trivial shifts of $A'_1 \times \cdots \times A'_m$ that contribute to $S_L(A_1,\dots,A_m)$ in Lemma 2.5 need to be taken into account when removing solutions from $A_1 \times \cdots \times A_m$, as in Theorem 1.2.

As mentioned in § 1, Theorem 1.2 can be used when studying $\mathbb{T}$ as a limit object or model for certain finite additive-combinatorial questions. A well-known question of this kind asks for the maximal density $d_L(\mathbb{Z}_p)$ of a subset of $\mathbb{Z}_p$ not containing solutions to a given system $Lx = 0$. In [**1**], the special case of Theorem 1.2 for a single equation was used to show that if L is a linear form with integer coefficients in at least three variables, then $d_L(\mathbb{Z}_p)$ converges to the natural analogue $d_L(\mathbb{T}) := \sup\{\mu_{\mathbb{T}}(A)\colon A \subset \mathbb{T} \text{ is } L\text{-free}\}$ as $p \to \infty$ through the primes. Theorem 1.2 enables us to extend this convergence result to so-called systems of *complexity* 1. A notion of complexity for systems of linear forms on finite abelian groups was introduced in [**4**], to which we refer the reader for more background on this topic. We use the following variant of this notion, specific to groups $\mathbb{Z}_p$ and $\mathbb{T}$.

Definition 5.1. Let L be an $r \times m$ integer matrix. We say the system of equations $Lx = 0$ (alternatively, the matrix L) has *complexity* k if k is the smallest integer such that, for any $\epsilon > 0$, there exists $\delta > 0$ with the following property: let $G = \mathbb{T}$ or $\mathbb{Z}_p$ for any large prime $p > p_0(L)$; then for any $f, g\colon G \to \mathbb{C}$ with $\|f\|_{L_\infty(G)}$, $\|g\|_{L_\infty(G)}$ both at most 1 and $\|f - g\|_{U^{k+1}(G)} \leqslant \delta$, we have $|S_{L,G}(f) - S_{L,G}(g)| \leqslant \epsilon$.

Here the notation $\|f\|_{U^k(G)}$ refers to the kth Gowers uniformity norm, which is defined on $L_\infty(G)$ for any compact abelian group G [**3**]. Using Theorem 1.2, the main convergence result from [**1**] can be extended as follows.

Theorem 5.2. *Let $\mathcal{F}$ be a finite family of full-rank integer matrices of complexity* 1*, and let $d_{\mathcal{F}}(\mathbb{Z}_p)$ denote the maximal density of an $\mathcal{F}$-free subset of $\mathbb{Z}_p$. Then $d_{\mathcal{F}}(\mathbb{Z}_p) \to d_{\mathcal{F}}(\mathbb{T})$ as $p \to \infty$ over primes.*

Here $d_{\mathcal{F}}(\mathbb{T}) := \sup\{\mu_{\mathbb{T}}(A)\colon A \subset \mathbb{T} \text{ is } \mathcal{F}\text{-free}\}$, where we say a measurable set $A \subset \mathbb{T}$ is $\mathcal{F}$-free if A is L-free for every $L \in \mathcal{F}$. Generalizing the argument in [**1**] to obtain Theorem 5.2 is not hard; we omit the details in this paper.

Let us close with remarks regarding further generalizations of removal lemmas. Recently, Král′ *et al*. extended Theorem 1.1 to all finite abelian groups [**6**], and upon inspection Green's proof [**5**] for single equations can be seen to hold over arbitrary compact abelian groups. Can Theorem 1.2 be generalized to all compact abelian groups? The desired generalization should hold with a function $\delta(L, \epsilon)$ independent of the group, so in particular δ should not depend on the group's topological dimension. The argument in this paper, when applied with $\mathbb{T}^n$ instead of $\mathbb{T}$, gives a parameter δ which decays to 0 as n grows, so additional ideas are required.

Acknowledgements. The authors thank Tim Austin for helpful conversations. Parts of this work were carried out while the authors attended the Discrete Analysis programme at the Isaac Newton Institute, whose support is gratefully acknowledged. Both authors are EPSRC postdoctoral fellows and gratefully acknowledge the support of the EPSRC.

References

1. P. Candela and O. Sisask, On the asymptotic maximal density of a set avoiding solutions to linear equations modulo a prime, *Acta Math. Hungar.* **132**(3) (2011), 223–243.
2. A. Deitmar and S. Echterhoff, *Principles of harmonic analysis* (Springer, 2009).
3. T. Eisner and T. Tao, Large values of the Gowers–Host–Kra seminorms, *J. Analyse Math.* **117** (2012), 133–186.
4. W. T. Gowers and J. Wolf, The true complexity of a system of linear equations, *Proc. Lond. Math. Soc.* **100** (2010), 155–176.
5. B. J. Green, A Szemerédi-type regularity lemma in abelian groups, with applications, *Geom. Funct. Analysis* **15**(2) (2005), 340–376.
6. D. Král′, O. Serra and L. Vena, On the removal lemma for linear systems over abelian groups, *Eur. J Combin.* **34** (2013), 248–259.
7. D. Král′, O. Serra and L. Vena, A removal lemma for systems of linear equations over finite fields, *Israel J. Math.* **187** (2012), 193–207.
8. V. F. Lev, On sum-free subsets of the torus group, *Funct. Approx. Comment. Math.* **37**(2) (2007), 277–283.
9. A. M. Macbeath, On measure of sum sets, II, *Proc. Camb. Phil. Soc.* **49** (1953), 40–43.
10. D. A. Raikov, On the addition of point-sets in the sense of Schnirelmann, *Mat. Sb.* **5**(47) (1939), 425–440.
11. I. Z. Ruzsa, Sumsets and structure, in *Combinatorial number theory and additive group theory*, pp. 87–210, Advanced Courses in Mathematics—CRM Barcelona (Springer, 2009).
12. A. Shapira, A proof of Green's conjecture regarding the removal properties of sets of linear equations, *J. Lond. Math. Soc.* **81**(2) (2010), 355–373.
13. T. Tao and V. Vu, *Additive combinatorics* (Cambridge University Press, 2006).
14. J. D. Vaaler, A geometric inequality with applications to linear forms, *Pac. J. Math.* **83**(2) (1979), 543–553.