⌐ GAFA Geometric And Functional Analysis

AN INVERSE THEOREM FOR THE UNIFORMITY SEMINORMS ASSOCIATED WITH THE ACTION OF $\mathbb{F}_p^\infty$

Vitaly Bergelson, Terence Tao and Tamar Ziegler

Abstract. Let $\mathbb{F}$ a finite field. We show that the universal characteristic factor for the Gowers–Host–Kra uniformity seminorm $U^k(\mathrm{X})$ for an ergodic action $(T_g)_{g\in\mathbb{F}^\omega}$ of the infinite abelian group $\mathbb{F}^\omega$ on a probability space $X = (X, \mathcal{B}, \mu)$ is generated by phase polynomials $\phi : X \to S^1$ of degree less than $C(k)$ on X, where $C(k)$ depends only on k. In the case where $k \leqslant \operatorname{char}(\mathbb{F})$ we obtain the sharp result $C(k) = k$. This is a finite field counterpart of an analogous result for $\mathbb{Z}$ by Host and Kra [HK]. In a companion paper [TZ] to this paper, we shall combine this result with a correspondence principle to establish the inverse theorem for the Gowers norm in finite fields in the high characteristic case $k \leqslant \operatorname{char}(\mathbb{F})$, with a partial result in low characteristic.

1 Introduction

1.1 Gowers–Host–Kra seminorms. This paper is concerned with the structural theory of measure preserving actions of abelian groups. We begin with some general definitions.

DEFINITION 1.1 (G-systems). Let $(G, +)$ be a locally compact abelian group. A *G-system* $\mathrm{X} = (X, \mathcal{B}_X, \mu_X, (T_g)_{g\in G})$ is a probability space $X = (X, \mathcal{B}_X, \mu_X)$ which is *separable modulo null sets* (i.e. $\mathcal{B}_X$ is countably generated modulo null sets), together with an action $g \mapsto T_g$ of G on X by measure-preserving transformations $T_g : X \to X$, and such that the map $(g, x) \mapsto T_g x$ is jointly measurable in g and x. (In order to carry out certain measure-theoretical constructions such as the disintegration of measures with respect to a factor, we will tacitly assume that the underlying measure spaces of the G-systems we will be dealing with are *regular*, meaning that in the triple $(X, \mathcal{B}, \mu)$, X is a compact metric space, $\mathcal{B}$ is the (completion of the) σ-algebra of Borel sets, and μ is a Borel measure. Since every separable (modulo null sets) probability measure space is equivalent to a regular space (see for example [Fu2, Prop. 5.3]), this assumption can be made without any loss of generality. Here, equivalence means equivalence of abstract σ-algebras modulo null sets; see [Fu2, Def. 5.2] for a precise definition.

Moreover, when dealing with factor maps between G-systems (see Definition 1.7 below) we will be assuming without specifically mentioning this that this regularity assumption applies simultaneously to a G-system and its factor. (Cf. [Fu2, Th. 5.15].))

We define the L^p spaces $L^p(\mathrm{X}) = L^p(X, \mathcal{B}_X, \mu_X)$ for $1 \leqslant p \leqslant \infty$ in the usual manner (in particular, we identify any two functions in $L^p(\mathrm{X})$ which agree μ-almost everywhere). If X is a point, we write $\mathrm{X} = \mathrm{pt}$. Given any measurable $\phi : X \to \mathbb{C}$ and $h \in G$, we define the shift $T_h \phi := \phi \circ T_h$ and the multiplicative derivative $\boldsymbol{\Delta}_h \phi := \overline{\phi} \cdot T_h \phi$. Similarly, if $(U, +)$ is an abelian group and $\phi : X \to U$ is a measurable function, we define the shift $T_h \phi := \phi \circ T_h$ and the additive derivative $\Delta_h \phi := T_h \phi - \phi$. We observe the commutativity relations $\boldsymbol{\Delta}_h \boldsymbol{\Delta}_k = \boldsymbol{\Delta}_k \boldsymbol{\Delta}_h$ and $\Delta_h \Delta_k = \Delta_k \Delta_h$ for all $h, k \in G$. Observe also that for any $h \in G$, $\boldsymbol{\Delta}_h$ and Δ_h are multiplicative and additive homomorphisms from the group of measurable functions from $X \to \mathbb{C}$ or $X \to U$ endowed with pointwise multiplication or pointwise addition respectively, to itself.

We say that a G-system is *ergodic* if the only functions in $L^2(\mathrm{X})$ which are invariant under the G-action (by the shifts introduced above) are the constants.

REMARK 1.2. Most of our analysis will take place in the setting of ergodic systems, but for various technical reasons we will sometimes have to work with non-ergodic systems. In some (but not all) cases, results on ergodic systems can be extended satisfactorily to the non-ergodic case using the ergodic decomposition. The hypothesis of separability is a technical one (used in particular in Appendix C to obtain a certain measurability property), but can often be removed in applications by restricting the σ-algebra $\mathcal{B}_X$ to the sub-algebra generated by the functions one is interested in studying, together with all of their shifts.

In most of our analysis, the group G will be countable, discrete and abelian and hence has a *Følner sequence*, i.e. a sequence $(\Phi_n)_{n=1}^\infty$ of finite subsets of G satisfying $|(\Phi_n + h)\Delta\Phi_n|/|\Phi_n| \to 0$ as $n \to \infty$ for all $h \in G$. (We use $|E|$ to denote the cardinality of a finite set E, and Δ to denote symmetric difference.) It is well known that one can always choose a Følner sequence to be nested and to satisfy the condition $G = \bigcup_{n=1}^\infty \Phi_n$, and we will assume throughout this paper that the Følner sequences we deal with have this additional property. Model examples include the integers $\mathbb{Z}$ (with $\Phi_n = [-n, n]$), and the (additive group of) countably infinite vector space $\mathbb{F}^\omega := \oplus \mathbb{F} \equiv \bigcup_n \mathbb{F}^n$ over a finite field $\mathbb{F}$ (with $\Phi_n = \mathbb{F}^n$). For our initial discussion we will allow G to be any countable abelian group, but we will eventually restrict our attention to the vector space $\mathbb{F}^\omega$. (We will also need to consider the actions of various compact abelian groups, and in particular closed subgroups of the Pontryagin dual $\widehat{\mathbb{F}^\omega} \equiv \prod \mathbb{F}$ of $\mathbb{F}^\omega$.) However, it may be useful for future applications to note that several of the tools used here are in fact valid for arbitrary countable discrete abelian G. (In fact, it seems likely that the hypothesis that G be discrete could be dropped in much of the theory. It may also be possible to generalize from abelian groups G to nilpotent groups G. We will not pursue these matters here.)

This paper is concerned with the following seminorms for G-systems:

DEFINITION 1.3 (Gowers–Host–Kra uniformity seminorms, cf. [HK]).   Let G be a countable abelian group, let $X = (X, \mathcal{B}_X, \mu_X, (T_g)_{g \in G})$ be a G-system, let $\phi \in L^\infty(X)$, and let $k \geqslant 1$ be an integer. We define the *Gowers–Host–Kra seminorm* $\|\phi\|_{U^k(X)}$ of order k of ϕ recursively by the formula

$$\|\phi\|_{U^1(X)} := \lim_{n \to \infty} \left\| \mathbb{E}_{h \in \Phi_n^1} T_h \phi \right\|_{L^2(X)}$$

for $k = 1$, and

$$\|\phi\|_{U^k(X)} := \lim_{n \to \infty} (\mathbb{E}_{h \in \Phi_n^k} \|\boldsymbol{\Delta}_h \phi\|_{U^{k-1}(X)}^{2^{k-1}})^{1/2^k}$$

for $k \geqslant 1$, where, for each k, $\Phi_1^k \subset \Phi_2^k \subset \dots$ is a Følner sequence, and we use the expectation notation $\mathbb{E}_{h \in H} f(h) := \frac{1}{|H|} \sum_{h \in H} f(h)$ for finite non-empty sets H and functions $f : H \to \mathbb{C}$ throughout this paper.

REMARK 1.4. In the case of ergodic G-systems, we can use the mean ergodic theorem to simplify the U^1 norm as $\|\phi\|_{U^1(X)} = \left| \int_X \phi \, d\mu_X \right|$. In the non-ergodic case, the mean ergodic theorem gives the formula

$$\|\phi\|_{U^1(X)} = \left(\lim_{n \to \infty} \int_X \mathbb{E}_{h \in \Phi_n^k} \boldsymbol{\Delta}_h \phi \, d\mu_X \right)^{1/2}$$

and more generally

$$\|\phi\|_{U^k(X)} = \left(\lim_{n_k \to \infty} \dots \lim_{n_1 \to \infty} \int_X \mathbb{E}_{h_k \in \Phi_{n_k}^k} \dots \mathbb{E}_{h_1 \in \Phi_{n_1}^1} \boldsymbol{\Delta}_{h_k} \dots \boldsymbol{\Delta}_{h_1} \phi \, d\mu_X \right)^{1/2^k}.$$

The existence of the limits (independently of the choice of Følner sequences) as well as similar integral formulae for the higher-order Gowers–Host–Kra norms are given for $\mathbb{Z}$-actions in [HK], and for actions of a general countable abelian group in Lemma A.18. One can also show that the Gowers–Host–Kra seminorms are indeed seminorms on $L^\infty(X)$; see Lemma A.20.

EXAMPLE 1.5.   Let $\phi : X \to S^1$ be an *eigenfunction* for an ergodic G-system, thus ϕ is measurable and $T_h \phi = \lambda(h) \phi$ for all $h \in G$ and some character $\lambda : G \to S^1$, where $S^1 := \{z \in \mathbb{C} : |z| = 1\}$ is the unit circle. Then $\|\phi\|_{U^k(X)} = 1$ for all $k \geqslant 2$. If λ is trivial, then $\|\phi\|_{U^1(X)} = 1$ as well, otherwise $\|\phi\|_{U^1(X)} = 0$.

REMARK 1.6. In the case when G is finite and X is just G with normalized counting measure, the σ-algebra that consists of all subsets of G, and the translation action, the Gowers–Host–Kra seminorms simplify to the *Gowers uniformity norms*

$$\|f\|_{U^k(G)} = \left(\mathbb{E}_{x, h_1, \dots, h_k \in G} \boldsymbol{\Delta}_{h_k} \dots \boldsymbol{\Delta}_{h_1} f(x) \right)^{1/2^k}.$$

These norms (in the special case $G = \mathbb{Z}/N\mathbb{Z}$) were first introduced by Gowers in [G1], where he derives quantitative bounds for Szemerédi's theorem on arithmetic progressions in sets of positive upper density in the integers. The above seminorms in the context of ergodic $\mathbb{Z}$-systems were introduced by Host and Kra in [HK], as a tool in the study of the ergodic averages related to Furstenberg's ergodic theoretic proof [Fu1] of Szemerédi's theorem [Sz]. The Gowers uniformity norms for other finite

abelian groups, such as finite-dimensional vector spaces $\mathbb{F}^n$ over a finite field $\mathbb{F}$, were studied in [GrT2,4,5], [S], [LoMS], [GW].

1.2 Universal characteristic factors. A fundamental concept in the study of the Gowers–Host–Kra uniformity seminorms is that of the *universal characteristic factor* for such norms. To describe this concept we need some notation.

DEFINITION 1.7 (Factors). A *factor* $Y = (Y, \mathcal{B}_Y, \mu_Y, (S_g)_{g \in G}, \pi_Y^X)$ of a G-system $X = (X, \mathcal{B}_X, \mu_X, (T_g)_{g \in G})$ is another G-system $(Y, \mathcal{B}_Y, \mu_Y, (S_g)_{g \in G})$, together with a measurable *factor map* $\pi_Y^X : X \to Y$, such that the push-forward $(\pi_Y^X)_* \mu_X$ of μ_X by π_Y^X is equal to μ_Y, and such that $\pi_Y^X \circ T_g = S_g \circ \pi_Y^X$ μ_X-a.e. for all $g \in G$. We will often write $(Y, (S_g)_{g \in G}, \pi_Y^X)$, (Y, π_Y^X) or just Y, for the factor $(Y, \mathcal{B}_Y, \mu_Y, (S_g)_{g \in G}, \pi_Y^X)$. If U is a measure space and $f : Y \to U$ is a measurable map, we write $(\pi_Y^X)^* f : X \to U$ for the pullback $(\pi_Y^X)^* f := f \circ \pi_Y^X$. Conversely, if $f \in L^2(X)$, we write $(\pi_Y^X)_* f \in L^2(Y)$ for the pushforward of f, and $\mathbb{E}(f|Y) := (\pi_Y^X)^* (\pi_Y^X)_* f \in L^2(X)$ for the conditional expectation of f to Y. We say that $f \in L^2(X)$ is $\mathcal{B}_Y$-*measurable* if $f = \mathbb{E}(f|Y)$, of equivalently if $f = (\pi_Y^X) F$ for some $F \in L^2(Y)$. We refer to (X, π_Y^X) as an *extension* of the G-system $(Y, \mathcal{B}_Y, \mu_Y, (S_g)_{g \in G})$. If Y is a point, $(\pi_{\text{pt}}^X)_* f = \int_X f \, d\mu_X$.

One factor $Y = (Y, \mathcal{B}_Y, \mu_Y, (S_g)_{g \in \mathcal{G}}, \pi_Y^X)$ of X is said to *extend* another factor $Y' = (Y', \mathcal{B}_{Y'}, \mu_{Y'}, (S_g')_{g \in \mathcal{G}}, \pi_{Y'}^X)$ of X (or equivalently, Y' is a *sub-factor* of Y) if every Y'-measurable function is also Y-measurable; in this case, we write $Y \geqslant Y'$. Two factors are said to be *isomorphic* if they extend each other. Note that the notion of extension is a partial order modulo measure equivalence. When we say that a factor is *maximal* (resp. *minimal*) with respect to some property, we mean that there is no extension (resp. sub-factor) of this factor, that obeys that property, which is not already equivalent to that factor.

We say that a factor $Y = (Y, \mathcal{B}_Y, \mu_Y, (S_g)_{g \in G}, \pi_Y^X)$ is *generated* by a collection $\mathcal{F}$ of measurable functions $f : X \to \mathbb{C}$ if the σ-algebra $(\pi_Y^X)^{-1}(\mathcal{B}_Y)$ is generated (modulo μ_X-null sets) by the pre-images of level sets $T_g f^{-1}(V)$ of functions $f \in \mathcal{F}$, where $g \in G$ and V ranges over Borel subsets of $\mathbb{C}$. Equivalently, Y is the minimal factor such that all functions in $\mathcal{F}$ are $\mathcal{B}_Y$-measurable.

REMARK 1.8. Observe that any σ-algebra $\mathcal{B} \subset \mathcal{B}_X$ which is preserved by the G-action induces a factor $Y = (X, \mathcal{B}, \mu_X, (T_g)_{g \in G}, \text{id})$ of $X = (X, \mathcal{B}_X, \mu_X, (T_g)_{g \in G})$ (indeed, up to isomorphism, all factors arise in this manner, and we will often abuse notation by identifying factors with invariant σ-algebras). In this case we see that $(\pi_Y^X)_* f = \mathbb{E}(f|\mathcal{B})$ and $(\pi_Y^X)^* f = f$. Note that if X is separable modulo null sets, then $L^2(X)$ is separable, hence on taking orthogonal projections $L^2(Y)$ is separable, hence Y is also separable modulo null sets.

REMARK 1.9. Observe that any factor of an ergodic G-system is also ergodic. The converse, of course, is not true.

PROPOSITION 1.10 (Universal characteristic factor). *Let G be a countable abelian group, let X be a G-system, and let $k \geqslant 1$. Then there exists a factor $\mathcal{Z}_{<k} = \mathcal{Z}_{<k}(X) = \left(Z_{<k}(X), \mathcal{B}_{\mathcal{Z}_{<k}}, \mu_{\mathcal{Z}_{<k}}, (S_g)_{g \in G}, \pi_{Z_{<k}(X)}^X\right)$ of X with the property that for*

every $f \in L^\infty(\mathrm{X})$, $\|f\|_{U^k(\mathrm{X})} = 0$ if and only if $(\pi^X_{Z_{<k}(\mathrm{X})})_* f = 0$ (or equivalently $\mathbb{E}(f|\mathcal{Z}_{<k}) = 0$). This factor is unique up to isomorphism.

Proof. The uniqueness of $\mathcal{Z}_{<k}$ is clear; the existence follows immediately from Lemma A.32. $\qquad\qquad\square$

REMARK 1.11.   As U^k is a seminorm on $L^\infty(\mathrm{X})$, an equivalent characterization of $\mathcal{Z}_{<k}$ is that it is the maximal factor for which $\|f\|_{U^k(X)} = \|\mathbb{E}(f|\mathcal{Z}_{<k})\|_{U^k(X)}$ for all $f \in L^\infty(\mathrm{X})$. The factor $\mathcal{Z}_{<k}$ is also referred to as $\mathcal{Z}_{k-1}$ in the literature (and in particular in [HK]). From (A.9) we have the monotonicity $\mathcal{Z}_{<j} \leqslant \mathcal{Z}_{<k}$ for $k \geqslant j$.

EXAMPLE 1.12 (Universal characteristic factors for small k).   Let X be a G-system, then from the ergodic theorem we see that $\mathcal{Z}_{<1}(\mathrm{X})$ is generated by the G-invariant functions on X; in particular, for ergodic G-systems $Z_{<1}(X)$ is simply a point. Some spectral theory also reveals (in the ergodic case) that $\mathcal{Z}_{<2}(\mathrm{X})$ is the *Kronecker factor* of X, that is, the factor generated by the eigenfunctions of X (see Example 1.5); this system is isomorphic to $(H, \mathcal{B}, \mu, (T_g)_{g \in G})$ where H is a closed subgroup of the Pontryagin dual $\hat{G}$ of G, $\mathcal{B}$ the Borel σ-algebra, μ the Haar measure, and the action of G being given by a homomorphism from G to H, acting on H by translation.

In view of Proposition 1.10 and the decomposition $f = \mathbb{E}(f|\mathcal{Z}_{<k}) + (f - \mathbb{E}(f|\mathcal{Z}_{<k}))$, we see that any function $f \in L^\infty(\mathrm{X})$ can be decomposed into a $\mathcal{Z}_{<k}$-measurable function, plus a function with vanishing U^k norm. When coupled with an explicit description of $\mathcal{Z}_{<k}$ (as was done for $k = 1, 2$ in Example 1.12), this decomposition leads to some highly non-trivial multiple recurrence and convergence theorems in ergodic theory: see for instance [HK], [Z], [FHK], [FW]. (See also [GrT1,2,3], [GW] for some finitary analogues of this decomposition, and some applications to additive combinatorics and analytic number theory.)

It is thus of interest to describe the universal characteristic factors $\mathcal{Z}_{<k}(\mathrm{X})$ as explicitly as possible. One particular class of functions related to such factors are the *phase polynomials*:

DEFINITION 1.13 (Phase polynomials).   Let G be a countable discrete abelian group, X be a G-system, let $\phi \in L^\infty(\mathrm{X})$, and let $k \geqslant 0$ be an integer. We say that ϕ is a *phase polynomial of degree less than k* if we have $\boldsymbol{\Delta}_{h_1} \dots \boldsymbol{\Delta}_{h_k} \phi = 1$ μ_X-a.e., for all $h_1, \dots, h_k \in G$. (In particular, setting $h_1 = \dots = h_k = 0$, we see that phase polynomials must take values in the unit circle S^1 μ_X-almost everywhere.) We let $\mathcal{P}_{<k}(\mathrm{X})$ denote the set of phase polynomials of degree less than k.

We write $\mathrm{Abr}_{<k}(\mathrm{X})$ for the factor of X generated by $\mathcal{P}_{<k}(\mathrm{X})$, and say that X is an *Abramov system of order $< k$* if X is "generated" by $\mathcal{P}_{<k}(\mathrm{X})$, or equivalently if $\mathcal{P}_{<k}(\mathrm{X})$ spans $L^2(\mathrm{X})$. (It was Abramov who studied (under the name "systems with quasi-discrete spectrum" and for $\mathbb{Z}$-actions, see [A]) systems of this type.)

EXAMPLE 1.14. $\mathcal{P}_{<0}(\mathrm{X})$ consists only of the constant function 1, so the only Abramov system of order < 0 is a trivial (one point) system. $\mathcal{P}_{<1}(\mathrm{X})$ consists of the G-invariant functions from X to S^1 (which, in the ergodic case, are just the constants), and only the Abramov systems of order < 1 are those for which the action of G is trivial. In the ergodic case, $\mathcal{P}_{<2}(\mathrm{X})$ consists of the eigenfunctions

from Example 1.5, and so the ergodic Abramov systems of order < 2 are precisely the Kronecker systems (i.e. systems generated by translations on compact abelian groups). There is an analogous relationship between higher-degree phase polynomials and higher-order eigenfunctions of X. Observe that every phase polynomial $\phi \in \mathcal{P}_{<k}(\mathrm{X})$ takes the form $\phi = e(P)$, where $e : \mathbb{R}/\mathbb{Z} \to S^1$ is the standard character $e(x) := e^{2\pi i x}$, and $P : X \to \mathbb{R}/\mathbb{Z}$ is a *polynomial of degree $< k$* in the sense that $\Delta_{h_1} \dots \Delta_{h_k} P = 0$ for all $h_1, \dots, h_k \in G$.

The following observations are immediate:

LEMMA 1.15 (Trivial facts about phase polynomials). *Let G be a countable discrete abelian group, X be a G-system, and let $k \geqslant 0$.*

  (i) *(Monotonicity) We have $\mathcal{P}_{<k}(\mathrm{X}) \subseteq \mathcal{P}_{<k+1}(\mathrm{X})$. In particular, $\mathrm{Abr}_{<k}(\mathrm{X}) \leqslant \mathrm{Abr}_{<k+1}(\mathrm{X})$, and an Abramov system of order $< k$ is also an Abramov system of order $< k + 1$.*
 (ii) *(Homomorphism) $\mathcal{P}_{<k}(\mathrm{X})$ is an abelian group under pointwise multiplication, and for each $h \in G$, $\boldsymbol{\Delta}_h$ is a homomorphism from $\mathcal{P}_{<k+1}(\mathrm{X})$ to $\mathcal{P}_{<k}(\mathrm{X})$.*
(iii) *(Polynomiality criterion) Conversely, if $\phi : X \to \mathbb{C}$ is a measurable function such that $\boldsymbol{\Delta}_h \phi \in \mathcal{P}_{<k}(\mathrm{X})$ for all $h \in G$, then $\phi \in \mathcal{P}_{<k+1}(\mathrm{X})$.*
 (iv) *(Functoriality) If Y is a factor of X, then the pullback map $(\pi_X^Y)^*$ is a homomorphism from $\mathcal{P}_{<k}(\mathrm{Y})$ to $\mathcal{P}_{<k}(\mathrm{X})$. Conversely, if $f : \mathrm{Y} \to \mathbb{C}$ is such that $(\pi_X^Y)^* f \in \mathcal{P}_{<k}(\mathrm{X})$, then $f \in \mathcal{P}_{<k}(\mathrm{Y})$.*

It is not hard to show that every phase polynomial of degree $< k$ is in fact $\mathcal{Z}_{<k}(\mathrm{X})$-measurable (see Lemma A.35); thus $\mathrm{Abr}_{<k}(\mathrm{X}) \leqslant \mathcal{Z}_{<k}(\mathrm{X})$. However, in the case of $\mathbb{Z}$-systems, the characteristic factor $\mathcal{Z}_{<k}(\mathrm{X})$ also contains some functions which do not arise from phase polynomials, even when one assumes ergodicity; this fact was essentially first observed by Furstenberg and Weiss [FuB]. Indeed, it is not too difficult to show (see [HK]) that any factor $\mathrm{Y} = (N/\Gamma, (x \mapsto a^g x)_{g \in \mathbb{Z}}, \pi_Y^X)$ of X which is a $< k$-*step nilsystem*, thus N is a nilpotent Lie group of step $< k$, Γ is a discrete cocompact subgroup, N/Γ is given Haar measure, and $a \in N$ is a group element, is a sub-factor of $\mathcal{Z}_{<k}(\mathrm{X})$. The converse statement is much deeper, and is due to Host and Kra [HK]:

Theorem 1.16 (Description of $\mathcal{Z}_{<k}(\mathrm{X})$ for ergodic $\mathbb{Z}$-systems [HK, Th. 10.1]). *Let X be an ergodic $\mathbb{Z}$-system. Then $\mathcal{Z}_{<k}(\mathrm{X})$ is the minimal factor that extends all $<k$-step nilsystem factors of X. Indeed, $\mathcal{Z}_{<k}(\mathrm{X})$ is itself the inverse limit of $<k$-step nilsystems.*

REMARK 1.17. In the case of ergodic $\mathbb{Z}$-systems, every Abramov system of order $<k$ is the inverse limit of $<k$-step nilsystems (this is implicit from [HK]). However, the converse is not true; see [FuB] for further discussion. (For instance, the Heisenberg system $H(\mathbb{R})/H(\mathbb{Z})$ discussed in section 1.4 is a 2-step nilsystem which is not an Abramov system of any order, if the underlying shifts α, β and 1 are independent over $\mathbb{Q}$.)

REMARK 1.18. The finitary counterpart to Theorem 1.16, where G and X are $\mathbb{Z}/N\mathbb{Z}$, is known as the *inverse conjecture for the Gowers norm* for cyclic groups

$\mathbb{Z}/N\mathbb{Z}$, and would have a number of applications to additive combinatorics and analytic number theory, see e.g. [GrT3]. It is currently only proven for $k \leqslant 3$ [GrT2].

1.3   Main result.   In view of Theorem 1.16, it is natural to ask what the universal characteristic factors $\mathcal{Z}_{<k}(\mathrm{X})$ are for ergodic G-systems, when G is an abelian group other than the integers $\mathbb{Z}$. (The non-ergodic case can then be recovered, at least in principle, from the ergodic case by the ergodic decomposition, though we will not attempt to do so here.)

The main results of this paper give a sharp answer to this question in the case when G is the (additive group of the) countably infinite vector space $G = \mathbb{F}^\omega$ over a finite field $\mathbb{F}$ of characteristic $\geq k$ which we will refer to as the *high characteristic* case. We also partially answer this question for general k:

Theorem 1.19 (Sharp description of $\mathcal{Z}_{<k}(\mathrm{X})$ for ergodic $\mathbb{F}^\omega$-systems in high char). *Let $\mathbb{F}$ be a finite field, and let X be an ergodic $\mathbb{F}^\omega$-system. Let $k \leqslant \mathrm{char}(\mathbb{F})$. Then for each $k \geqslant 1$, we have $\mathrm{Abr}_{<k}(\mathrm{X}) = \mathcal{Z}_{<k}(\mathrm{X})$.*

Theorem 1.20 (Partial description of $\mathcal{Z}_{<k}(\mathrm{X})$ for ergodic $\mathbb{F}^\omega$-systems in low char). *Let $\mathbb{F}$ be a finite field, and let X be an ergodic $\mathbb{F}^\omega$-system. Then for each $k \geqslant 1$, we have $\mathrm{Abr}_{<k}(\mathrm{X}) \leqslant \mathcal{Z}_{<k}(\mathrm{X}) \leqslant \mathrm{Abr}_{<C(k)}(\mathrm{X})$ for some $C(k)$ depending only on k.*

REMARK 1.21.   We also have a slightly more precise structural description of $\mathcal{Z}_{<k}(\mathrm{X})$, as a tower of abelian extensions by polynomial cocycles; see Theorems 4.8 and 8.7. The quantity $C(k)$ can in principle be computed explicitly from the proof of Theorem 1.20, but we have not sought to obtain the best possible value of $C(k)$. We believe in fact that $\mathcal{Z}_{<k}(\mathrm{X})$ should equal $\mathrm{Abr}_{<k}(\mathrm{X})$ for all k (not just in the high characteristic case $k \leqslant \mathrm{char}(\mathbb{F})$).

From Theorems 1.20, 1.19 and Proposition 1.10 we have the following immediate corollaries:

COROLLARY 1.22 (Ergodic inverse Gowers conjecture – high characteristic).   *Let $\mathbb{F}$ be a finite field, and let X be an ergodic $\mathbb{F}^\omega$-system. Let $1 \leqslant k \leqslant \mathrm{char}(\mathbb{F})$ and $f \in L^\infty(\mathrm{X})$ be such that $\|f\|_{U^k(\mathrm{X})} > 0$. Then there exists $\phi \in \mathcal{P}_{<k}(\mathrm{X})$ such that $\int_X f\overline{\phi}\, d\mu_X \neq 0$.*

COROLLARY 1.23 (Partial ergodic inverse Gowers conjecture).   *Let $\mathbb{F}$ be a finite field, and let X be an ergodic $\mathbb{F}^\omega$-system. Let $k \geqslant 1$ and $f \in L^\infty(\mathrm{X})$ be such that $\|f\|_{U^k(\mathrm{X})} > 0$. Then there exists $\phi \in \mathcal{P}_{<C(k)}(\mathrm{X})$ such that $\int_X f\overline{\phi}\, d\mu_X \neq 0$.*

In a companion paper [TZ], we will combine Corollary 1.22 with a version of the Furstenberg correspondence principle, as well as the equidistribution theory in [GrT5], to obtain a finitary counterpart to this theorem:

Theorem 1.24 (Inverse theorem for the Gowers norm over finite fields in high characteristic [TZ]).   *Let $\mathbb{F}$ be a finite field of characteristic p, let $1 \leqslant k \leqslant \mathrm{char}(\mathbb{F})$ be an integer, and let $\delta > 0$. Then there exists $c = c(p, k, \delta) > 0$ such that for finite-dimensional vector space G over $\mathbb{F}$ and any function $f : G \to \mathbb{C}$ with $\|f\|_{L^\infty(G)} \leqslant 1$ and $\|f\|_{U^k(G)} \geqslant \delta$, one has $|\mathbb{E}_{x \in G} f(x)\overline{\phi(x)}| \geqslant c$ for some $\phi \in \mathcal{P}_{<k}(G)$.*

REMARK 1.25. We conjecture that this result should in fact hold without the restriction on the characteristic depending on k. Once this restriction is removed, it becomes important here that the values of the phase polynomial ϕ be allowed to range freely in the unit circle S^1. If one constrains ϕ to take values in the pth roots of unity C_p, then the claim can fail for small p, as first observed in [LoMS], [GrT5]. However, such examples do not obstruct Theorem 1.24 from holding when ϕ takes values in S^1 (see [TZ] for further discussion). By using Corollary 1.23 instead of Corollary 1.22, one can obtain a partial analogue of Theorem 1.24 in the low characteristic case $k > \operatorname{char}(\mathbb{F})$, in which $\mathcal{P}_{<k}(G)$ is replaced by $\mathcal{P}_{<C(k)}(G)$; see [TZ].

REMARK 1.26. Theorem 1.19 should also allow one (assuming sufficiently high characteristic) to obtain a formula for the limit of multiple ergodic averages of quantities such as $c(g) := \mu(A \cap T^g A \cap \ldots \cap T^{(k-1)g} A)$ (as in [Z]), and to be able to show that $c(g)$ can be approximated by a function of polynomials in g, in the spirit of the results in [BeHK]. We hope to report on these and other applications in a subsequent paper.

1.4 The Heisenberg example.

To illustrate the above results we now pause to describe the model case of a *Heisenberg system*. (The discussion in this section is not directly used in the remainder of the paper.) To simplify the discussion we restrict attention to the $k = 3$ case.

We first review the more familiar case of $\mathbb{Z}$-systems. For any commutative ring R, let $H(R)$ be the Heisenberg group

$$H(R) := \begin{pmatrix} 1 & R & R \\ 0 & 1 & R \\ 0 & 0 & 1 \end{pmatrix}.$$

This is clearly a 2-step nilpotent group. The quotient space $H(\mathbb{R})/H(\mathbb{Z})$ is then a 2-step nilmanifold that has a natural Haar measure. For fixed $\alpha, \beta \in \mathbb{R}$, the function $\psi : \mathbb{Z} \to H(\mathbb{R})$ defined by

$$\psi(n) := \begin{pmatrix} 1 & n\alpha & n\gamma + \binom{n}{2}\alpha\beta \\ 0 & 1 & n\beta \\ 0 & 0 & 1 \end{pmatrix} \tag{1.1}$$

can be easily checked to be a homomorphism $\mathbb{Z} \to H(\mathbb{R})$, and thus defines an action $(T_g)_{g \in \mathbb{Z}}$ on $H(\mathbb{R})/H(\mathbb{Z})$; if $1, \alpha, \beta$ are linearly independent over $\mathbb{Q}$, one can show that this action is ergodic (see e.g. [P]). The function $f : H(\mathbb{R}) \to \mathbb{C}$ defined by

$$f \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} = e(z - \{x\}y), \tag{1.2}$$

where $\{x\} := x - \lfloor x \rfloor$ is the fractional part of x, induces a function $f : H(\mathbb{R})/H(\mathbb{Z}) \to \mathbb{C}$, with

$$f(a^n x_0) = e\left(\binom{n}{2}\alpha\beta + n\gamma - \{n\alpha\}n\beta\right) \tag{1.3}$$

where $x_0 := H(\mathbb{Z}) \in H(\mathbb{R})/H(\mathbb{Z})$ (one can easily check that this function is well defined as a function on $H(\mathbb{R})/H(\mathbb{Z})$).

The argument on the right-hand side of (1.3) is an example of a *generalized polynomial* (see [BeL]). It can be shown that the function f is asymptotically orthogonal to all phase polynomials of degree < 3 (in fact, it is asymptotically orthogonal to phase polynomials of all degrees), but that $\|f\|_{U^3(X)} > 0$. For the system $X = (H(\mathbb{R})/H(\mathbb{Z}), (T_g)_{g \in \mathbb{Z}})$, the algebra generated by the function f and the eigenfunctions of the $\mathbb{Z}$ action (which take the form $e(ax + by + \theta)$ for integers a, b and $\theta \in \mathbb{R}/\mathbb{Z}$) is dense in $L^2(X)$.

One can imitate this construction for an $\mathbb{F}^\omega$-action, for $\mathbb{F}$ with char $\mathbb{F} > 2$. The analogues of the ring of integers $\mathbb{Z}$ and the field of reals $\mathbb{R}$ will be the ring $\mathbb{F}[t] := \left\{ \sum_{i=0}^m a_i t^i : a_i \in \mathbb{F} \right\}$ of polynomials in one variable and the field $\mathbb{F}((t)) := \left\{ \sum_{i=-\infty}^m a_i t^i : a_i \in \mathbb{F} \right\}$ of Laurent polynomials of one variable, respectively. The analogue of the torus $\mathbb{R}/\mathbb{Z}$ is the abelian group $\mathbb{F}((t))/\mathbb{F}[t]$. If $x = \sum_{i=-\infty}^m a_i t^i$ is an element of $\mathbb{F}((t))$, we write $\lfloor x \rfloor := \sum_{i=0}^m a_i t^i \in \mathbb{F}[t]$ for the "integer part", and $\{x\} := x - \lfloor x \rfloor$ for the "fractional part".

Observe that *as an additive group*, $\mathbb{F}[t]$ is isomorphic to $\mathbb{F}^\omega$, and so an ergodic $\mathbb{F}[t]$-system is also an ergodic $\mathbb{F}^\omega$-system. Consider the homogeneous space $X := H(\mathbb{F}((t)))/H(\mathbb{F}[t])$; this is a "manifold" over the base field $\mathbb{F}((t))$ and comes with a natural Haar measure. For fixed $\alpha, \beta \in \mathbb{F}((t))$, the function $\psi : \mathbb{F}[t] \to H(\mathbb{F}((t)))$ defined by (1.1) is still a homomorphism, and thus defines an action $(T_g)_{g \in \mathbb{F}[t]}$ on X as before. If we define $f : H(\mathbb{F}((t))) \to \mathbb{C}$ by (1.2), then this function again induces a well-defined function on X and we have the formula (1.3) as before with $x_0 := \mathbb{F}[t] \in X$.

Thus far, the $\mathbb{F}^\omega$-action case has proceeded in exact analogy with the $\mathbb{Z}$-action case. However, the key difference between the two case is that the phase $P(n) := n^2 \alpha\beta + n\gamma - \{n\alpha\}n\beta$ on the right-hand side of (1.3) is not just a generalized polynomial; it is a genuine polynomial, indeed one easily checks the identity $\Delta_{h_1}\Delta_{h_2}\Delta_{h_3} P = 0$ for all $h_1, h_2, h_3 \in \mathbb{F}[t]$. (This ultimately stems from the fact that the maps $x \mapsto \lfloor x \rfloor$ and $x \mapsto \{x\}$ are genuine homomorphisms in the $\mathbb{F}^\omega$-action case (where there is no "carry" operation), whereas they are only "approximate" homomorphisms in the integer case.) Thus the system $X = (X, (T_g)_{g \in \mathbb{F}[t]})$ is isomorphic to the system $((\mathbb{T}(\mathbb{F}))^2 \times \mathbb{T}(\mathbb{F}), (T_g)_{g \in \mathbb{F}^\omega})$, where $\mathbb{T}(\mathbb{F}) := \mathbb{F}^{\mathbb{Z}} \equiv \mathbb{F}((t))/\mathbb{F}[t]$ is the Pontryagin dual $\widehat{\mathbb{F}^\omega}$ of $\mathbb{F}^\omega$, and the $\mathbb{F}^\omega$-action is given by

$$T_g\big((x, y), z\big) = \left((x, y) + (g\alpha, g\beta), z + g\gamma + \binom{g}{2}\alpha\beta + \lfloor g\alpha \rfloor y - g\beta\{x\} - \{g\alpha\}g\beta \right)$$

for $x, y \in \mathbb{T}(\mathbb{F}) \equiv \mathbb{F}((t))/\mathbb{F}[t]$ and $g \in \mathbb{F}^\omega \equiv \mathbb{F}[t]$. For any character χ of $\mathbb{T}(\mathbb{F})$, $\chi \in \widehat{\mathbb{T}}(\mathbb{F})$, the function $z \mapsto \chi(z)$ lies in $\mathcal{P}_2(\mathbb{T}(\mathbb{F}))$. By Fourier decomposition, any function in $L^2(X)$ can be written as $\sum_{\chi \in \widehat{\mathbb{T}}(\mathbb{F})} f_\chi(x, y)\chi(z)$, where $f_\chi(x, y) \in L^2(\mathbb{T}(\mathbb{F})^2)$. Since for any χ, the function $f_\chi(x, y)$ is defined on a Kronecker system (arising from the eigenfunctions $(x, y) \mapsto e(ax + by)$ for $a, b \in \mathbb{F}[t]$), we have $f_\chi \in \mathcal{Z}_{<2}(X)$. It follows that X is an Abramov system of order < 3 and is thus equal to $\mathcal{Z}_{<3}(X)$ by Theorem 1.20. This latter fact can also be checked by calculating the Gowers–Host–Kra

seminorms directly. The reader may also wish to verify Theorem 4.8 explicitly for this example.

The main point of the example above is that it demonstrates the peculiarities of the polynomial aspect of nilpotency in the case of an $\mathbb{F}^\omega$-action.

$$
\begin{array}{ccccc}
\text{Prop. 6.1} & \longrightarrow & \text{Thm. 5.4} & \longleftarrow & \text{Prop. 7.1} \\
\uparrow & & \downarrow & & \\
\text{Lems. 5.1, 5.3} & \longrightarrow & \text{Thm. 4.5} & \longrightarrow \text{Lem. 4.7} \longrightarrow & \text{Thm. 4.8} \\
& & \downarrow & & \\
\text{Prop. 4.4} & \longrightarrow & \text{Thm. 3.8} & \longrightarrow \text{Thm. 3.3} \longleftarrow & \text{Prop. 3.4} \\
& & & \downarrow & \\
\text{[TZ], [GrT5]} & \longrightarrow & \text{Thm. 1.24} & \longleftarrow & \text{Thm. 1.20}
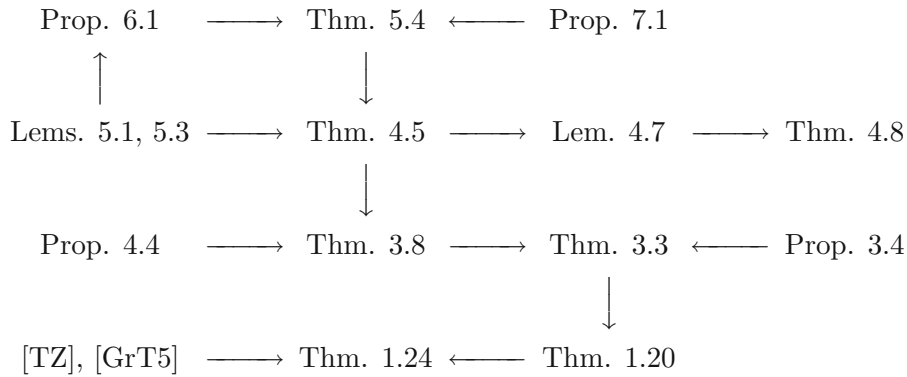\end{array}
$$

Figure 1: Logical dependencies in the proofs of key theorems in the paper in the general characteristic case (ignoring the material in the appendices, which are used throughout the paper). Theorem 4.8 is also used inductively to establish Propositions 6.1, 7.1. The high-characteristic case follows a similar logic.

1.5   Overview of the structure of the paper.   The bulk of the paper will be devoted to the proof of Theorems 1.20, 1.19. We first prove Theorem 1.20, and then specialize to the high characteristic case. The proof will be established via a series of reductions. The main task is to establish the inequality $\mathcal{Z}_{<k}(X) \leqslant \mathrm{Abr}_{<C(k)}(X)$ for some $C(k)$ depending on k, which will imply that $\mathcal{Z}_{<k}(X)$ is generated by polynomials of high degree. Thus we do not need to keep careful track of the degree of the polynomials that arise in our analysis, except to ensure that they are bounded. The next few steps closely follow the approach of Host and Kra [HK]. Namely, by following the methods in [HK], one can reduce matters to understanding a certain type of cocycle $f : \mathbb{F}^\omega \times X \to S^1$ on a structured type of ergodic $\mathbb{F}^\omega$-system X; specifically, one needs to show that any $(\mathbb{F}^\omega, X, S^1)$-cocycle of type $< k$ on an ergodic $\mathbb{F}^\omega$-system of order $< k$ (see Definition 3.1) is cohomologous to a $(\mathbb{F}^\omega, X, S^1)$-phase polynomial; see Theorem 4.5. (All definitions for the terms used here are given later in the paper, when these results are formalized.) As mentioned earlier, the degree of this polynomial is not too important so long as it is bounded.

By an inductive hypothesis one can understand the underlying system X reasonably well; it turns out to be a finite tower of abelian group extensions, each of which is given by a cocycle which is cohomologous to a polynomial.

To proceed further one needs to understand the condition that a cocycle $f : \mathbb{F}^\omega \times X \to S^1$ has bounded type. By definition, this means that a certain "iterated derivative" $d^{[k]}f$ of that cocycle is a coboundary. It turns out that the underlying system X can be expressed as an abelian extension $X = Y \times_\rho U$ of a simpler system Y,

where $U = (U, \cdot)$ is a compact abelian group. The group U then acts freely on X by the action $V_u : (y, v) \mapsto (y, uv)$, with this action commuting with the action of $\mathbb{F}^\omega$. One can differentiate the cocycle f repeatedly in the "vertical" direction by using the operation $\boldsymbol{\Delta}_u f(g, x) := f(g, V_u x)/f(g, x)$. It turns out that each such vertical differentiation reduces the type of the cocycle; iterating this, one concludes that one can find an m such that $\boldsymbol{\Delta}_{t_1} \ldots \boldsymbol{\Delta}_{t_m} f$ is a cocycle of type < 0 (i.e. a coboundary) for every $t_1, \ldots, t_m \in U$. In other words, one has an equation of the form

$$\boldsymbol{\Delta}_{t_1} \ldots \boldsymbol{\Delta}_{t_m} f(g, x) = \boldsymbol{\Delta}_g F_{t_1, \ldots, t_m}(x) := \frac{F_{t_1, \ldots, t_m}(T_g x)}{F_{t_1, \ldots, t_m}(x)} \qquad (1.4)$$

for some functions $F_{t_1, \ldots, t_m}(x)$. A key technical point is that while the function $F_{t_1, \ldots, t_m}(x)$ is *a priori* only measurable in x, it can be made to be measurable in the parameters $t_1, \ldots, t_m$ also (see Lemma C.4). This will be rather important for us as we will be relying quite heavily on the measurability property in our arguments. (For instance, we will need a variant of the classical Steinhaus theorem that asserts that if A is a measurable subset of a compact abelian group U with positive measure, then the difference set $A - A$ contains a neighborhood of the origin (cf. Lemma D.1). Curiously, analogous results are exploited in the additive-combinatorial approach to the Gowers inverse problem (see e.g. [GrT2]), where they go by the name of "Bogolyubov-type lemmas".)

We would like to use the equation (1.4) to show that f itself is a coboundary, up to a polynomial error. The idea is to "integrate" the derivatives $\boldsymbol{\Delta}_{t_1}, \ldots, \boldsymbol{\Delta}_{t_m}$ one at a time. At any intermediate stage of this process, one will have obtained another function $\tilde{f} : \mathbb{F}^\omega \times X \to S^1$ which differs from f (multiplicatively) by a phase polynomial, and which obeys an equation of the form

$$\boldsymbol{\Delta}_{t_1} \ldots \boldsymbol{\Delta}_{t_j} \tilde{f}(g, x) = c_{t_1, \ldots, t_j}(g, x) \boldsymbol{\Delta}_g F_{t_1, \ldots, t_j}(x) \qquad (1.5)$$

for some $0 \leqslant j \leqslant m$, some functions $F_{t_1, \ldots, t_j}(x)$, some phase polynomials $c_{t_1, \ldots, t_j}$, any $t_1, \ldots, t_j \in U$, $g \in G$, and almost all $x \in X$. A technical point here is that whereas the original function f was a cocycle in $\mathbb{F}^\omega$, the new function $\tilde{f}$ need not be; however, it turns out that the $\mathbb{F}^\omega$-cocycle property is not actually needed in the rest of our analysis; indeed, as we shall shortly see, we can rely primarily on the cocycle behavior in t_j instead. (Indeed, there seem to be significant technical difficulties if one attempts to work purely within the world of cocycles, stemming ultimately from the fact that $\mathbb{F}^\omega$ is not finitely generated. For instance, it appears quite difficult to show that the abstract nilpotent group $\mathcal{G}^{[k]}$ defined in [HK] acts transitively in the non-finitely generated case.)

The left-hand side of (1.5) exhibits some "linearity" in t_j, thanks to the cocycle equation $\boldsymbol{\Delta}_{t_j + t_j'} f = (\boldsymbol{\Delta}_{t_j} V_{t_j'} f) \boldsymbol{\Delta}_{t_j'} f$. This induces some approximate linearity properties on $F_{t_1, \ldots, t_j}$. One can conjugate $F_{t_1, \ldots, t_j}$ to upgrade this approximate linearity to genuine linearity (here we crucially exploit the measurability of all our objects with respect to $t_1, \ldots, t_j$), at least for $t_1, \ldots, t_j$ in an open subgroup U' of U (here we rely heavily on the finite characteristic of $\mathbb{F}^\omega$, which forces U to be a torsion group, so that every open neighborhood of U contains an open subgroup). It still remains to handle the behavior in the quotient group U/U', but this is a finite abelian group

and can be worked out explicitly, splitting this group as the product of finite cyclic groups and "straightening" the various actors in (1.5) along each such group. At the end of the day, one is able to "integrate" and remove one of the derivatives in (1.5), at the cost of introducing an additional U-invariant factor on the right-hand side (cf. the classical "$+C$" ambiguity when solving an equation $\frac{d}{dx}f = F$ by integration). However, this additional factor (which basically lives on Y rather than on X) can be shown to be cohomologous to a polynomial by an induction hypothesis, and so can be absorbed into the other terms on the right-hand side. One can then iterate this procedure until the number j of derivatives in (1.5) reaches zero, at which point we obtain the desired characterization of f.

The high-characteristic case $k \leqslant \mathrm{char}(\mathbb{F})$ (i.e. Theorem 1.19) follows a similar logic, but one has to be much more careful in keeping track of the degrees and types of the various objects that arise in the proof, thus requiring "exact" counterparts of many of the lemmas used in the above analysis. The high-characteristic hypothesis is used to ensure that all the polynomials which appear take values in a coset of the cyclic group C_p, and have constant "integral" along various lines (see Lemma D.3 for a more precise statement.)

REMARK 1.27. As mentioned above, our approach largely follows that of Host and Kra [HK], although the abstract nilpotent structure group $\mathcal{G}^{[k]}$, which plays a central role in [HK], is not used explicitly in this paper (although of course we will be exploiting several symmetries of the cubic measure $\mu^{[k]}$, particularly with respect to "vertical rotations", which can be viewed as special elements of this structure group). Furthermore, the heart of the Host-Kra argument (the "lifting" proposition in [HK, Prop. 10.10]) does not have an equivalent in our context due to the fact that the group $\mathbb{F}^\omega$ is not finitely generated. As discussed above, the emphasis is instead on solving various equations of "Conze–Lesigne" type[CL]. In this respect, the arguments here share some features in common with [Z] as well as [HK].

REMARK 1.28. One curious distinction between the finite-characteristic case and the $\mathbb{Z}$-action case is that in the former case (and assuming sufficiently high characteristic), the systems $Z_k(X)$ are toral systems rather than inverse limits of such systems.

1.6 Acknowledgements.

2   Abelian Cohomology

Throughout the paper we will be relying heavily on the language of abelian cohomology of dynamical systems. We record the key definitions here; further discussion of these concepts can be found in Appendix B.

DEFINITION 2.1 (Abelian cohomology).   Let G be a countable discrete abelian group, let $X = (X, \mathcal{B}_X, \mu_X, (T_g)_{g \in G})$ be a G-system, and let $(U, \cdot)$ be a compact abelian group.

- We let $M(X, U)$ denote the set of all measurable functions $\phi : X \to U$, with two functions $\phi, \phi' : X \to U$ identified if they agree μ_X-almost everywhere; this is an abelian group under pointwise multiplication. We refer to elements of $M(X, U)$ as (X, U)-*functions*.
- We let $M(G, X, U)$ denote the set of all measurable functions $\rho : G \times X \to U$, with two functions $\rho, \rho' : G \times X \to U$ identified if $\rho(g, x) = \rho'(g, x)$ for all $g \in G$ and almost every $x \in X$; this is an abelian group under pointwise multiplication. We refer to elements of $M(G, X, U)$ as (G, X, U)-*functions*.
- We let $Z^1(G, X, U)$ denote the subgroup of $M(G, X, U)$ consisting of those (G, X, U)-functions that obey the cocycle equation $\rho(g + g', x) = \rho(g, T_{g'} x) \rho(g', x)$ for all $g, g' \in G$ and almost every $x \in X$. We refer to elements of $Z^1(G, X, U)$ as *abelian* (G, X, U)-*cocycles*, or *cocycles* for short.
- If ρ is an abelian (G, X, U)-cocycle, we define the *abelian extension* $X \times_\rho U$ of X by ρ to be the product space $(X \times U, \mathcal{B}_X \times \mathcal{B}_U, \mu_X \times \mu_U)$ with shift maps $(x, u) \mapsto (T_g x, \rho(g, x) u)$ for $g \in G$, where μ_U is normalized Haar measure on U. Note that this is indeed an extension of X, with the obvious factor map $\pi^{X \times_\rho U} X : (x, u) \mapsto x$.
- If F is a (X, U)-function, we define the *derivative* $\mathbf{\Delta} F$ of F to be the (G, X, U) function $\mathbf{\Delta} F(g, x) := \mathbf{\Delta}_g F(x) = F(T_g x) / F(x)$. We refer to F as an *antiderivative* of $\mathbf{\Delta} F$. We write $B^1(G, X, U) = \mathbf{\Delta} M(X, U)$ denote the space of all derivatives; this is a subgroup of $Z^1(G, X, U)$. We refer to elements of $B^1(G, X, U)$ as (G, X, U)-*coboundaries*, or *coboundaries* for short.
- More generally, if B is a G-invariant subset of X, we say that a (G, X, U)-function ρ is a (G, B, U)-*coboundary* if there exists some measurable $F : B \to U$ such that $\rho(g, x) = \mathbf{\Delta}_g F(x) = F(T_g x) / F(x)$ for all $g \in G$ and μ_X-almost every $x \in B$. We refer to F as an *antiderivative* of the coboundary.
- We say that two (G, X, U)-functions ρ, ρ' are (G, X, U)-*cohomologous* (or *cohomologous* for short), if $\rho / \rho' \in B^1(G, X, U)$. Note that we do not require ρ, ρ' to be cocycles here, though clearly any function cohomologous to a cocycle is again a cocycle.

REMARK 2.2.   Observe that if ρ and $\tilde{\rho}$ are cohomologous, then $X \times_\rho U$ and $X \times_{\tilde{\rho}} U$ are measure-equivalent systems. Thus, from the perspective of measure equivalence, (X, G, U) cocycles ρ are only determined up to their representative $[\rho]_{G, X, U}$ in the cohomology group $H^1(G, X, U) := Z^1(G, X, U) / B^1(G, X, U)$. As usual we have the short exact sequence

$$0 \longrightarrow B^1(G, X, U) \longrightarrow Z^1(G, X, U) \longrightarrow H^1(G, X, U) \longrightarrow 0.$$

REMARK 2.3. We will primarily be working with (G, X, U)-cocycles, but for technical reasons related to the fact that $\mathbb{F}^\omega$ is not countably generated, we will also need to work in the more general setting of (G, X, U)-functions. In practice, however, these functions will be "close" to coboundaries in various senses (for instance, they may differ from a coboundary by a polynomial function).

REMARK 2.4. We caution that abelian extensions of ergodic G-systems are not necessarily ergodic.

We will need several technical results concerning abelian cohomology groups, which we have collected in Appendix B, and which we will refer to as necessary in the main text of the paper.

3   Reduction to Abelian Extensions of Order $< k + 1$

Recall the definition of $\mathcal{Z}_{<k}(\mathrm{X})$ from Proposition 1.10. Before starting the proof of Theorem 1.20, we make a basic definition:

DEFINITION 3.1 (System of order $< k$).   Let $k \geqslant 1$, and let G be a countable discrete abelian group. A G-system X is said to be *of order $< k$* if $\mathcal{Z}_{<k}(\mathrm{X}) = \mathrm{X}$.

EXAMPLE 3.2. Trivial systems are of order < 1; Kronecker systems are of order < 2. From (A.9) we see that for any system X, $\mathcal{Z}_{<k}(\mathrm{X})$ is of order $< k$, and that any system of order $< k$ is automatically of order $< k + 1$.

By Theorem A.35, every Abramov system of order $< k$ is also a G-system of order $< k$. This and (A.9), will allow us to immediately derive Theorem 1.20 from the following claim:

Theorem 3.3 (First reduction).   *Let $\mathbb{F}$ be a finite field, let $k \geqslant 1$, and let X be an ergodic $\mathbb{F}^\omega$-system of order $< k$. Then X is an Abramov system of order $< O_k(1)$. (Here and in the sequel, we use $O_k(1)$ to denote any quantity bounded by $C(k)$ for some constant $C(k)$ depending only on k, and similarly for other choices of subscripts in the $O(\ )$ notation.)*

To prove Theorem 3.3, we will in fact prove a more precise statement. We will use the language of abelian cohomology, in particular the notions of (G, X, U)-functions, (abelian) (G, X, U)-cocycles, (G, X, U)-coboundaries, and extensions $\mathrm{X} \times_\rho U$ of a G-system X by a cocycle $\rho : G \times X \to U$, where U is a compact abelian group; see Definition 2.1 for full details.

The following basic fact was established by Host and Kra [HK]:

PROPOSITION 3.4 (Order $< k + 1$ systems are abelian extensions of order $< k$ systems [HK, Prop. 6.3]).   *Let G be a discrete countable abelian group, let $k \geqslant 1$, and let X be an ergodic G-system of order $< k + 1$. Then X is an abelian extension $\mathrm{X} \equiv \mathcal{Z}_{<k}(\mathrm{X}) \times_\rho U$ of the order $< k$ system $\mathcal{Z}_{<k}(\mathrm{X})$ for some compact abelian group U and some $(G, \mathcal{Z}_{<k}(\mathrm{X}), U)$-cocycle ρ. (All topological groups in this paper are assumed to be Hausdorff.)*

DEFINITION 3.5 (Phase polynomials, II). Let $(G, +)$ be a discrete countable abelian group, let $\mathrm{X} = (X, \mathcal{B}_X, \mu_X, (T_g)_{g \in G})$ be a G-system, let $(U, \cdot)$ be a compact abelian group, and let $k \geqslant 1$.

- We let $\mathcal{P}_{<k}(\mathrm{X}, U)$ denote the set of all (X, U)-functions ϕ satisfying the equation $\mathbf{\Delta}_{h_1} \dots \mathbf{\Delta}_{h_k} \phi(x) = 1$ for μ_X-a.e. x and for all $h_1, \dots, h_k \in G$, where $\mathbf{\Delta}_h \phi(x) := \phi(T_h x)/\phi(x)$, and we identify functions that agree μ_X-almost everywhere. We refer to elements of $\mathcal{P}_{<k}(\mathrm{X}, U)$ as (X, U)-*phase polynomials of degree* $< k$.
- We let $\mathcal{P}_{<k}(G, \mathrm{X}, U)$ denote the set of all (G, X, U)-functions ρ satisfying the equation $\mathbf{\Delta}_{h_1} \dots \mathbf{\Delta}_{h_k} \rho(g, x) = 1$ for μ_X-a.e. x and for all $g, h_1, \dots, h_k \in G$, where $\mathbf{\Delta}_h \rho(g, x) := \rho(g, T_h x)/\rho(g, x)$, and where we identify functions that agree μ_X-almost everywhere for each $g \in G$. We refer to elements of $\mathcal{P}_{<k}(G, \mathrm{X}, U)$ as (G, X, U)-*phase polynomials of degree* $< k$. We refer to elements of the intersection $\mathcal{P}_{<k}(G, \mathrm{X}, U) \cap Z^1(G, \mathrm{X}, U)$ as (G, X, U)-*phase polynomial cocycles of degree* $< k$.

EXAMPLE 3.6. If X is ergodic, then all phase polynomials of degree < 1 are constant in the X variable, and so $\mathcal{P}_{<1}(\mathrm{X}, U) \equiv U$ and $\mathcal{P}_{<1}(G, \mathrm{X}, U) \equiv M(G, U)$ in this case (for the definition of $M(G, U)$ see Definition 2.1).

REMARK 3.7. Observe that $\mathcal{P}_{<k}(\mathrm{X}, U)$ and $\mathcal{P}_{<k}(G, \mathrm{X}, U)$ are subgroups of $M(\mathrm{X}, U)$ and $M(G, \mathrm{X}, U)$ respectively. This definition is of course closely related to Definition 1.13; indeed, we observe that $\mathcal{P}_{<k}(\mathrm{X}) \equiv \mathcal{P}_{<k}(\mathrm{X}, S^1)$. Also observe that a (G, X, U)-function ρ is a phase polynomial coboundary of degree $< k$ if and only if it is a derivative $\rho(g, x) = \mathbf{\Delta}_g F(x)$ of a (X, U)-phase polynomial F of degree $< k + 1$. Indeed, we have the short exact sequence

$$0 \longrightarrow \mathcal{P}_{<1}(\mathrm{X}, U) \longrightarrow \mathcal{P}_{<k+1}(\mathrm{X}, U) \overset{\Delta}{\longrightarrow} \mathcal{P}_{<k}(G, \mathrm{X}, U) \cap B^1(G, \mathrm{X}, U) \longrightarrow 0, \quad (3.1)$$

and similarly

$$0 \longrightarrow \mathcal{P}_{<1}(\mathrm{X}, U) \longrightarrow M(\mathrm{X}, U) \overset{\Delta}{\longrightarrow} B^1(G, \mathrm{X}, U) \longrightarrow 0$$

(for the definitions of $M(\mathrm{X}, U), B^1(G, \mathrm{X}, U)$ see Definition 2.1).

Using Proposition 3.4 and Fourier analysis, we can reduce matters to studying projections of abelian cocycles to the unit circle. More precisely, in future sections we will show the following result.

Theorem 3.8 (Second reduction). *Let $\mathbb{F}$ be a finite field, let $k \geqslant 1$, and let X be an ergodic $\mathbb{F}^\omega$-system of order $< k$, and let $\mathrm{X} \times_\rho U$ be a (possibly non-ergodic) abelian extension of X by a $(\mathbb{F}^\omega, \mathrm{X}, U)$-cocycle which is of order $< k + 1$. Then for every character $\chi \in \hat{U}$ (i.e. every continuous homomorphism $\chi : U \to S^1$), the $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cocycle $\chi \circ \rho$ is cohomologous to a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-phase polynomial of degree $< O_k(1)$, i.e. $\chi \circ \rho \in \mathcal{P}_{<O_k(1)}(\mathbb{F}^\omega, \mathrm{X}, S^1) \cdot B^1(\mathbb{F}^\omega, \mathrm{X}, S^1)$ for all $\chi \in \hat{U}$.*

Proof of Theorem 3.3 assuming Theorem 3.8. We induct on k. The claim is trivial for $k = 1$, so suppose that $k \geqslant 2$ and that Theorem 3.3 has already been proven for $k - 1$. By Proposition 3.4, we may assume that $\mathrm{X} = \mathcal{Z}_{<k-1}(\mathrm{X}) \times_\rho U$ for some compact abelian group U and some $(\mathbb{F}^\omega, Z_{<k-1}(X), U)$-cocycle ρ.

Let $\chi \in \hat{U}$ be any element of the Pontryagin dual $\hat{U}$ of U, thus $\chi : U \to S^1$ is a character (i.e. a continuous homomorphism). Applying Theorem 3.8, we see that there exists a measurable function $F_\chi \in M(\mathcal{Z}_{<k-1}(\mathrm{X}), S^1)$ and a $(\mathbb{F}^\omega, \mathcal{Z}_{<k-1}(\mathrm{X}), S^1)$-phase polynomial cocycle $\tilde{\rho}_\chi$ of degree $< O_k(1)$, for which we have the equation

$$\chi \circ \rho = \tilde{\rho}_\chi \mathbf{\Delta} F_\chi . \tag{3.2}$$

If we define $\phi_\chi \in L^\infty(\mathrm{X})$ to be the function $\phi_\chi(x, u) := \overline{F_\chi}(x)\chi(u)$, then we see from (3.2) that $\mathbf{\Delta}\phi_\chi = \tilde{\rho}_\chi$, and thus (by (3.1)) ϕ_χ is a phase polynomial of degree $< O_k(1)$. On the other hand, $F_\chi(x)$ lies in $L^2(\mathcal{Z}_{<k-1}(\mathrm{X}))$ and can thus be approximated in that space by finite linear combinations of phase polynomials in $\mathcal{P}_{<O_k(1)}(\mathrm{X})$, by the induction hypothesis. Thus we see that the characters $(x, u) \mapsto \chi(u)$ can also be approximated in $L^2(\mathrm{X})$ by finite linear combinations of phase polynomials in $\mathcal{P}_{<O_k(1)}(\mathrm{X})$. Since $L^2(\mathrm{X})$ is generated by $L^2(\mathcal{Z}_{<k-1}(\mathrm{X}))$ and these characters, we see that X is an Abramov system of order $< O_k(1)$ as claimed. $\square$

It remains to prove Theorem 3.8. This will be the objective of the next few sections.

4 Functions of Type $< k$

We make a further reduction, introducing the useful notion of a *function of type $< k$*. (This concept is essentially that of a *cocycle of type k* from [HK], but generalized to non-cocycles and to more general group actions. We have replaced "k" by "$< k$" as such cocycles will have "degree" strictly less than k in some sense.)

DEFINITION 4.1 (Functions of type $< k$). Let $(G, +)$ be a discrete countable abelian group, let $\mathrm{X} = (X, \mathcal{B}_X, \mu_X, (T_g)_{g \in G})$ be a G-system, and let $U = (U, \cdot)$ be a compact abelian group. Let $k \geqslant 0$, and let $\mathrm{X}^{[k]} = \big(X^{[k]}, \mathcal{B}^{[k]}, \mu^{[k]}, (T_g^{[k]})_{g \in G}\big)$ be the G-system defined in Definition A.12.

- For each $f \in M(\mathrm{X}, U)$, we define $d^{[k]}f \in M\big(\mathrm{X}^{[k]}, U\big)$ to be the function

$$d^{[k]}f\big((x_{\mathbf{w}})_{\mathbf{w} \in \mathbf{2}^k}\big) := \prod_{\mathbf{w} \in \mathbf{2}^k} f(x_{\mathbf{w}})^{\mathrm{sgn}(\mathbf{w})}$$

where $\mathrm{sgn}(w_1, \ldots, w_k) := w_1 \ldots w_k \in \{-1, +1\}$.

- Similarly, for each $\rho \in M(G, \mathrm{X}, U)$, we define $d^{[k]}\rho \in M\big(G, \mathrm{X}^{[k]}, U\big)$ to be the function

$$d^{[k]}\rho\big(g, (x_{\mathbf{w}})_{\mathbf{w} \in \mathbf{2}^k}\big) := \prod_{\mathbf{w} \in \mathbf{2}^k} \rho(g, x_{\mathbf{w}})^{\mathrm{sgn}(\mathbf{w})} .$$

- A (G, X, U)-*function of type $< k$* is any function $\rho \in M(G, \mathrm{X}, U)$ such that $d^{[k]}\rho$ is a $\big(G, \mathrm{X}^{[k]}, U\big)$-coboundary. We let $M_{<k}(G, \mathrm{X}, U)$ denote the space of (G, X, U)-functions of type $< k$, and let $Z^1_{<k}(G, \mathrm{X}, U) := Z^1(G, \mathrm{X}, U) \cap M_{<k}(G, \mathrm{X}, U)$ denote the space of (G, X, U)-cocycles of type $< k$.

EXAMPLE 4.2. A (G, X, U)-function is of type < 0 if and only if it is a (G, X, U)-coboundary, thus $M_{<0}(G, \mathrm{X}, U) = Z^1_{<0}(G, \mathrm{X}, U) = B^1(G, \mathrm{X}, U)$.

$$\begin{array}{ccc}
\boldsymbol{\Delta}\mathcal{P}_{<k+1}(\mathrm{X},U) & \longrightarrow & B^1(G,\mathrm{X},U) \\
\downarrow & & \downarrow \\
\mathcal{P}_{<k}(G,\mathrm{X},U) \cap Z^1(G,\mathrm{X},U) \longrightarrow Z^1_{<k}(G,\mathrm{X},U) \longrightarrow Z^1(G,\mathrm{X},U) \\
\downarrow \qquad\qquad\qquad\qquad \downarrow \qquad\qquad\qquad \downarrow \\
\mathcal{P}_{<k}(G,\mathrm{X},U) \longrightarrow M_{<k}(G,\mathrm{X},U) \longrightarrow M(G,\mathrm{X},U)
\end{array}$$

Figure 2: Inclusions between various groups of (G,X,U) functions.

We make some easy observations (see Figure 2):

LEMMA 4.3 (Basic facts about functions of type $< k$). *Let G be a countable abelian group, let $\mathrm{X} = (X, \mathcal{B}_X, \mu_X, (T_g)_{g \in G})$ be an ergodic G-system, let $U = (U, \cdot)$ be a compact abelian group, and let $k \geqslant 0$.*

(i) *Every (G,X,U)-function of type $< k$ is also a (G,X,U)-function of type $< k+1$.*

(ii) *The set of (G,X,U)-functions of type $< k$ is a subgroup of $M(G,\mathrm{X},U)$ that contains the group $B^1(G,\mathrm{X},U)$ of coboundaries. In particular, any function (G,X,U)-cohomologous to a function of type $< k$, is also of type $< k$.*

(iii) *Let f be a (G,X,U)-function. Then f is a (G,X,U)-phase polynomial of degree $< k$ if and only if $d^{[k]}f = 0$ $\mu^{[k]}$-almost everywhere. In particular, every (G,X,U)-phase polynomial of degree $< k$ is of type $< k$.*

(iv) *If f is a (G,X,U)-coboundary (resp. a (G,X,U)-cocycle), then $d^{[k]}f$ is a $(G,\mathrm{X}^{[k]},U)$-coboundary (resp. a $(G,\mathrm{X}^{[k]},U)$-cocycle). Equivalently, we have the commuting diagram*

$$\begin{array}{ccccccccc}
0 & \longrightarrow & B^1(G,\mathrm{X},U) & \longrightarrow & Z^1(G,\mathrm{X},U) & \longrightarrow & H^1(G,\mathrm{X},U) & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle d^{[k]}} & & \downarrow{\scriptstyle d^{[k]}} & & \downarrow{\scriptstyle d^{[k]}} & & \\
0 & \longrightarrow & B^1(G,\mathrm{X}^{[k]},U) & \longrightarrow & Z^1(G,\mathrm{X}^{[k]},U) & \longrightarrow & H^1(G,\mathrm{X}^{[k]},U) & \longrightarrow & 0
\end{array}$$

of short exact sequences.

Proof. We first show (i). Let f be a (G,X,U)-function of type $< k$, then we can find a $(\mathrm{X}^{[k]},U)$-function F such that $d^{[k]}f(g,\mathbf{x}) = \boldsymbol{\Delta}_{g^{[k]}} F(\mathbf{x})$ for all $g \in G$ and $\mu^{[k]}$-almost all $\mathbf{x}$. Expressing $\mathrm{X}^{[k+1]} = \mathrm{X}^{[k]} \times \mathrm{X}^{[k]}$ and using (A.2) we conclude that $d^{[k+1]}f(g,\mathbf{x},\mathbf{x}') = \boldsymbol{\Delta}_{g^{[k+1]}}[F(\mathbf{x})/F(\mathbf{x}')]$ for all $g \in G$ and $\mu^{[k+1]}$-almost all $(\mathbf{x},\mathbf{x}')$, and so f is of type $< k+1$ as desired.

From (i) we see in particular that coboundaries, being of type < 0, are of type $< k$. The claims in (ii) are then easily verified.

To show (iii), we induct on k. The claim is easy for $k = 0, 1$ (using ergodicity), so suppose that $k \geqslant 2$ and the claim has already been shown for $k - 1$.

Suppose f is a (G,X,U)-phase polynomial of degree k. For any $g \in G$, $\boldsymbol{\Delta}_g f$ is a phase polynomial of degree $< k - 1$, and so by induction hypothesis $d^{[k-1]}(\boldsymbol{\Delta}_g f) = 1$ $\mu^{[k-1]}$-a.e. Since $d^{[k-1]}(\boldsymbol{\Delta}_g f) = \boldsymbol{\Delta}_{g^{[k-1]}}(d^{[k-1]}f)$, we conclude that $d^{[k-1]}f$ is invariant under the action of the diagonal group $\mathrm{diag}(G^{[k-1]})$. By Remark A.13, $d^{[k-1]}f$ is

constant $\left(\mu^{[k-1]}\right)_s$-almost everywhere for P_{k-1}-a.e. s. By Remark A.13 again, we conclude that $d^{[k]}f = 1$ $\mu^{[k]}$-a.e., as desired. The converse claim follows by reversing all the above steps.

The claim (iv) for coboundaries follows from the identity $d^{[k]}\mathbf{\Delta}_g F = \mathbf{\Delta}_{g^{[k]}} d^{[k]}F$, valid for any (X, U)-function F and any $g \in G$. The claim (iv) for cocycles is clear from direct computation. $\square$

The relevance of the type $< k$ concept to us lies in the important observation that abelian extensions of order $< k+1$ arise from functions of type $< k$:

PROPOSITION 4.4. Let G be a countable abelian group, and let $X = (X, \mathcal{B}_X, \mu_X, (T_g)_{g \in G})$ be an ergodic G-system of order $< k$ for some $k \geqslant 1$. Let U be a compact abelian group, and let ρ be a (G, X, U)-cocycle such that the abelian extension $X \times_\rho U$ is an ergodic G-system of order $< k+1$. Then ρ is a (G, X, U)-cocycle of type $< k$. In particular, for every character $\chi \in \hat{U}$, $\chi \circ \rho$ is a (G, X, S^1)-cocycle of type $< k$.

Proof. See [HK, Prop. 6.4] (the result there is only stated for $\mathbb{Z}$-actions, but the proof generalizes without difficulty). We give a brief sketch of the proof here, to indicate why the type $< k$ condition arises naturally. For simplicity let us take $U = S^1$. Let ϕ be the $(X \times_\rho U, S^1)$-function $\phi(z, u) := u$, then the function $J : L^2(X^{[k]}) \to L^2((X \times_\rho U)^{[k]})$ defined by $Jf(\mathbf{z}, \mathbf{u}) = f(\mathbf{z})d^{[k]}\phi(\mathbf{z}, \mathbf{u})$ is an isometry whose range is a closed $\mathrm{diag}(G^{[k]})$-invariant space containing $d^{[k]}\phi$. On the other hand, since $X \times_\rho U$ has order $< k+1$, we see from Lemma A.32 that $\|\phi\|_{U^{k+1}(X \times_\rho U)} \neq 0$. Applying Corollary A.21, this implies that $\left(\pi_{\mathcal{I}_k(X \times_\rho U)}^{(X \times_\rho U)^{[k]}}\right)_* d^{[k]}\phi \neq 0$. By the ergodic theorem, we conclude that the image of J contains a $\mathrm{diag}(G^{[k]})$-invariant function Jf. Unpacking the definitions, this implies that $f(T_g^{[k]}\mathbf{z})d^{[k]}\rho(g, \mathbf{z}) = f(\mathbf{z})$ for $\mu^{[k]}$ a.e. $\mathbf{z}$, which roughly speaking asserts that ρ is of type $< k$ on the support of f, which is a set of positive measure. By various cocycle identities one can also conclude that ρ is of type $< k$ on various shifts of the support of f, which by ergodicity can be glued together to establish a global type $< k$ condition; see the proof of [HK, Prop. 6.4] for details. $\square$

In view of Proposition 4.4, Theorem 3.8 will now follow by applying the following result to the $(\mathbb{F}^\omega, X, S^1)$-cocycle $\chi \circ \rho$.

Theorem 4.5 (Third reduction). *Let $\mathbb{F}$ be a finite field, let $m, k \geqslant 1$, and let X be an ergodic $\mathbb{F}^\omega$-system of order $< k$. Let $f \in M_{<m}(\mathbb{F}^\omega, X, S^1)$ be a $(\mathbb{F}^\omega, X, S^1)$-function of type $< m$. Then f is $(\mathbb{F}^\omega, X, S^1)$-cohomologous to a $(\mathbb{F}^\omega, X, S^1)$-phase polynomial of degree $< O_{k,m}(1)$. In other words,*

$$M_{<m}(\mathbb{F}^\omega, X, S^1) \subset \mathcal{P}_{<O_{k,m}(1)}(\mathbb{F}^\omega, X, S^1) \cdot B^1(\mathbb{F}^\omega, X, S^1).$$

REMARK 4.6. When $k = 1$, the only ergodic $\mathbb{F}^\omega$-systems of order 0 are trivial, and Theorem 4.5 is easily verified in this case. Unwinding all the previous reductions, this already establishes the $k = 1$ case of Theorem 1.20; thus $\mathcal{Z}_{<2}(X)$ is precisely the Kronecker system spanned by the eigenfunctions of X (i.e. the elements of $\mathcal{P}_{<2}(X)$).

Of course, as is well known, this fact can also be established more directly by classical spectral theory methods; see for instance the discussion just before [HK, Lem. 4.2].

While we will only need Theorem 4.5 in the case when f is a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cocycle of type $< m$, for inductive purposes it is important that we generalize the claim to the case where f is a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-function of type $< m$.

We will prove Theorem 4.5 in later sections. For now, let us observe that we can use Theorem 1.20 to obtain some structural control on systems of order $< k$. We say that a group U is m-torsion for some $m \geqslant 1$ if we have $u^m = 1$ for all $u \in U$.

LEMMA 4.7 (Structure groups for $\mathbb{F}^\omega$-systems are finite torsion). *Let $k \geqslant 2$ be such that Theorem 4.5 holds for $k-1$. Let $\mathbb{F}$ be a finite field of characteristic p, let X be an ergodic $\mathbb{F}^\omega$-system of order $< k-1$, and let $\mathrm{X} \times_\rho U$ be an abelian extension of order $< k$ by a compact abelian group which is also ergodic. Then U is p^m-torsion for some $m = O_k(1)$. If p is sufficiently large depending on k, then we can take $m = 1$.*

Proof. Since characters separate points, it suffices to show that $\chi(U)$ is a p^m-torsion group for every character $\chi \in \hat{U}$. By Proposition 4.4, $\chi \circ \rho$ is a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cocycle of type $< k-1$, and thus by Theorem 4.5 we have $\chi \circ \rho = q_\chi \mathbf{\Delta} F$ for some $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-phase polynomial q_χ of degree $< O_k(1)$ and some (X, S^1)-function F. Clearly q_χ is also a cocycle. By Lemma D.3, q_χ takes values in the cyclic group C_{p^m} for some $m = O_k(1)$, and for p sufficiently large depending on k, one can take $m = 1$.

As χ is a character, $\chi(U)$ is a compact subgroup of S^1. The system $\mathrm{X} \times_{\chi \circ \rho} \chi(U)$ is then a factor of $\mathrm{X} \times_\rho U$ and is therefore ergodic. Since $\chi \circ \rho = q_\chi \mathbf{\Delta} F$, the system X' defined with the same set, σ-algebra, and action as the system $\mathrm{X} \times_{q_\chi} S^1$, but with the measure $\mu_\mathrm{X} \times \mu_{S^1}$ replaced with the measure

$$\int_{\mathrm{X}'} f(x, u) d\mu_{\mathrm{X}'} := \int_\mathrm{X} \int_{\chi(U)} f\big(x, F(x)u\big) d\mu_{\chi(U)}(u) d\mu_\mathrm{X}(x)\,,$$

is equivalent to $\mathrm{X} \times_{\chi \circ \rho} \chi(U)$ and thus also ergodic. On the other hand, since q_χ takes values in C_{p^m}, any set of the form $X \times A$, where A is invariant under C_{p^m}, will be invariant in X'. These two statements are only consistent with each other if $\chi(U)$ is a subgroup of C_{p^m}, and is thus p^m-torsion, as desired. $\qquad\square$

Theorem 4.8 (Structure theorem). *Let $k \geqslant 1$ be such that Theorem 4.5 holds for all smaller values of k. Let $\mathbb{F}$ be a finite field of characteristic p, and let X be an ergodic $\mathbb{F}^\omega$-system. Then $\mathcal{Z}_{<1}(\mathrm{X})$ is trivial, and for all $2 \leqslant j \leqslant k$, we can write $\mathcal{Z}_{<j}(\mathrm{X}) \equiv \mathcal{Z}_{<j-1}(\mathrm{X}) \times_{\rho_{j-1}} U_{j-1}$ where U_{j-1} is p^m-torsion for some $m = O_{k,p}(1)$, and ρ_{j-1} is a $(\mathbb{F}^\omega, \mathcal{Z}_{<j-1}(\mathrm{X}), U_{j-1})$-phase polynomial cocycle of degree $O_k(1)$. In particular, we have*

$$\mathcal{Z}_{<k}(\mathrm{X}) \equiv U_0 \times_{\rho_1} U_1 \times_{\rho_2} \ldots \times_{\rho_{k-1}} U_{k-1}$$

where U_0 is trivial.

Proof. Fix $\mathbb{F}$, X. The triviality of $\mathcal{Z}_{<1}(\mathrm{X})$ follows from ergodicity. Now suppose inductively that $k \geqslant 2$, and the claim has already been proven for smaller values of k. By Proposition 3.4, we can write $\mathcal{Z}_{<k}(\mathrm{X}) \equiv \mathcal{Z}_{<k-1}(\mathrm{X}) \times_{\rho_{k-1}} U_{k-1}$ for some compact

abelian group U_{k-1} and some $(\mathbb{F}^\omega, \mathcal{Z}_{<k-1}(\mathbf{X}), U_{k-1})$-cocycle ρ_{k-1}. By Lemma 4.7, the compact abelian group U_{k-1} is thus p^m-torsion for some $m = O_k(1)$ (with $m = 1$ if p is large enough). This implies (see [Mor, Ch. 5, Th. 18]) that U_{k-1} is topologically isomorphic to the direct product of some cyclic p-groups C_{p^m} for $m = O_k(1)$. Let $\chi : U_{k-1} \to C_{p^m}$ be one of the coordinate maps. By Theorem 4.5, $\chi \circ \rho_{k-1}$ is cohomologous to a $(\mathbb{F}^\omega, \mathcal{Z}_{<k-1}(\mathbf{X}), S^1)$-phase polynomial cocycle q_χ of degree $< O_k(1)$, which by Lemma D.3 also takes values some p-group C_{p^n} for some $n = O_k(1)$. By increasing n if necessary we can take $n \geqslant m$; but note that if p is sufficiently large depending on k then we can take $n = m = 1$.

By construction, there exists a $(\mathcal{Z}_{<k-1}(\mathbf{X}), S^1)$-function F_χ such that $\chi \circ \rho_{k-1} = q_\chi \mathbf{\Delta} F_\chi$. Since $(\chi \circ \rho_{k-1})^{p^m} = 1$, we conclude that $1 = q_\chi^{p^m} \mathbf{\Delta} F_\chi^{p^m}$.

We claim that there exists a $(\mathcal{Z}_{<k-1}(\mathbf{X}), S^1)$-phase polynomial $\tilde{F}_\chi$ of degree $< O_k(1)$ such that $F_\chi^{p^m} = \tilde{F}_\chi^{p^m}$. There are two cases. If p is sufficiently large depending on k, then $q_\chi^{p^m} = 1$, so by ergodicity $F_\chi^{p^m}$ is constant, and the claim is trivial. Now suppose instead that $p = O_k(1)$. Then $\mathbf{\Delta} F_\chi^{p^m}$ is a $(\mathbb{F}^\omega, \mathcal{Z}_{<k-1}(\mathbf{X}), S^1)$-phase polynomial of degree $< O_k(1)$, and so (by (3.1)) $F_\chi^{p^m}$ is a $(\mathcal{Z}_{<k-1}(\mathbf{X}), S^1)$-phase polynomial of degree $< O_k(1)$. By Lemma D.7, we can find a $(\mathcal{Z}_{<k-1}(\mathbf{X}), S^1)$-phase polynomial $\tilde{F}_\chi$ of degree $< O_k(1)$ such that $F_\chi^{p^m} = \tilde{F}_\chi^{p^m}$, as claimed.

Now write $q'_\chi := q_\chi \mathbf{\Delta} \tilde{F}_\chi$ and $F'_\chi := F_\chi / \tilde{F}_\chi$, then we have $\chi \circ \rho = q'_\chi \mathbf{\Delta} F'_\chi$. Since F'_χ takes values in C_{p^m}, and q'_χ does also. By construction, q'_χ is thus a $(\mathbb{F}^\omega, \mathbf{X}, C_{p^m})$-phase polynomial of degree $< O_k(1)$.

If we let $q' := (q'_\chi)_\chi$ and $F' := (F_\chi)_\chi$, where χ ranges over all the coordinate projection maps, then q' is a $(\mathbb{F}^\omega, \mathbf{X}, U)$-phase polynomial of degree $< O_k(1)$ and $\rho = q' \mathbf{\Delta} F'$. Thus ρ is cohomologous to a phase polynomial of degree $< O_k(1)$, and thus we may assume (see Remark 2.2) that ρ is equal to a phase polynomial of degree $< O_k(1)$. The claim then follows from the induction hypothesis. $\square$

Our remaining task is to prove Theorem 4.5.

5   Reduction to Solving a Conze–Lesigne Type Equation

To prove Theorem 4.5 we will use two lemmas to reduce matters to solving a certain equation of Conze–Lesigne type. The first lemma allows one to descend a type condition on an extension to a type condition on a base, worsening the type if necessary:

LEMMA 5.1 (Descent of type).   *Let G be a countable abelian group, let $\mathbf{Y}$ be a G-system, let $k, m \geqslant 1$, and let $\mathbf{X} = \mathbf{Y} \times_\rho U$ be an ergodic abelian extension of $\mathbf{Y}$ by a $(G, \mathbf{Y}, U)$-phase polynomial cocycle ρ of degree $< m$. Let $\pi : X \to Y$ be the factor map, and let f be a $(\mathbf{Y}, S^1)$-function such that $\pi^* f$ is of type $< k$. Then f is of type $< k + m + 1$.*

REMARK 5.2.   For cocycles, a more general (and stronger) statement appears in [HK, Cor. 7.8]. However, for technical reasons, it is necessary for us to work with more general functions than just cocycles. (But see Corollary 8.11 below.)

Proof. Note that $\pi^* d^{[k]} f = d^{[k]} \pi^* f$ is a $(G, \mathrm{X}^{[k]}, S^1)$-coboundary, hence a $(G, \mathrm{X}^{[k]}, S^1)$-cocycle, and so (by Lemma B.1) $d^{[k]} f$ is a $(G, \mathrm{Y}^{[k]}, S^1)$-cocycle.

It is convenient to move from k to $k+1$. By Lemma 4.3(i), $\pi^* f$ is of type $< k+1$, thus we have an $(\mathrm{X}^{[k+1]}, S^1)$-function F such that $d^{[k+1]}(\pi^* f)(g, \mathbf{x}) = \mathbf{\Delta}_{g^{[k+1]}} F(\mathbf{x})$ for all $g \in G$ and $\mu^{[k+1]}$-almost every $\mathbf{x} \in X^{[k+1]}$.

We view $\mathrm{X}^{[k+1]}$ as $\mathrm{Y}^{[k+1]} \times U^{[k+1]}$. Since the projection of $\mu^{[k+1]}$ on $\mathrm{Y}^{[k+1]}$ is $(\pi_* \mu)^{[k+1]}$, every ergodic component of $\mathrm{X}^{[k+1]}$ is an abelian extension of an ergodic component of $\mathrm{Y}^{[k+1]}$, with a Mackey group $M \leqslant U^{[k+1]}$ (see for example [FuB]). In other words, for each ergodic component of $\mathrm{Y}^{[k+1]}$, there exists a map $G : \mathrm{Y}^{[k+1]} \to U^{[k+1]}$ such that the $(\mathbb{F}^\omega, \mathrm{Y}^{[k+1]}, U^{[k+1]})$-cocycle

$$\tilde{\rho}(g, \mathbf{y}) := \big(\rho(g, \mathbf{y_w})\big)_{\mathbf{w} \in \mathbf{2}^{k+1}} \mathbf{\Delta}_{g^{[k+1]}} G(\mathbf{y})$$

takes values in M (i.e. it is a $(\mathbb{F}^\omega, \mathrm{Y}^{[k+1]}, M)$-cocycle). On each ergodic component of $\mathrm{X}^{[k+1]}$, if we set $K(\mathbf{y}, \mathbf{m}) := F(\mathbf{y}, G(\mathbf{y})\mathbf{m})$ then we have

$$d^{[k+1]} f(g, \mathbf{y}) = \frac{K\big(S_g^{[k+1]} \mathbf{y}, \tilde{\rho}(g, \mathbf{y})\mathbf{m}\big)}{K(\mathbf{y}, \mathbf{m})} \,.$$

If we expand into a Fourier series $K(\mathbf{y}, \mathbf{m}) = \sum_{\chi \in \hat{M}} a_\chi(\mathbf{y}) \chi(\mathbf{m})$ and compare Fourier coefficients, we conclude

$$a_\chi(S_{g^{[k+1]}} \mathbf{y}) \chi\big(\tilde{\rho}(g, \mathbf{y})\big) = d^{[k+1]} f(g, \mathbf{y}) a_\chi(\mathbf{y})$$

almost everywhere on any ergodic component. In particular, $|a_\chi|$ is invariant and thus constant a.e. on any ergodic component. We extend χ arbitrarily to a character $\chi \in \widehat{U^{[k+1]}} \equiv \hat{U}^{[k+1]}$.

On almost every component, at least one of the a_χ must be non-vanishing; since there are at most countably many characters (by the separability of X), we conclude that for some $\chi \in \hat{U}^{[k+1]}$, we have $|a_\chi| \neq 0$ on a positive measure collection of ergodic components. On each such component, if one sets $b := a_\chi / |a_\chi|$, we thus have

$$d^{[k+1]} f(g, \mathbf{y}) = \chi\big(\tilde{\rho}(g, \mathbf{y})\big) \mathbf{\Delta}_{g^{[k+1]}} b(\mathbf{y}) = \Big[\prod_{\mathbf{w} \in \mathbf{2}^{k+1}} \chi_\mathbf{w}\big(\rho(g, \mathbf{y_w})\big) \Big] \mathbf{\Delta}_{g^{[k+1]}} Gb(\mathbf{y}) \,,$$

where $\chi_\mathbf{w}$ are the components of χ. Thus the $(G, \mathrm{Y}^{[k+1]}, S^1)$-function

$$H(g, \mathbf{y}) := d^{[k+1]} f(g, \mathbf{y}) / \prod_{\mathbf{w} \in \mathbf{2}^{k+1}} \chi_\mathbf{w}\big(\rho(g, \mathbf{y_w})\big)$$

is a (G, A, S^1)-coboundary for some $\mathrm{diag}\big(G^{[k]}\big)$-invariant subset A of $\mathrm{Y}^{[k+1]}$ of positive measure.

We now claim that $(T_h)_\alpha^{[k+1]} H$ is also a (G, A, S^1)-coboundary for every positive side transformation $(T_h)_\alpha^{[k+1]} H$. But a computation shows that

$$\frac{(T_h)_\alpha^{[k+1]} H}{H}(g, y) = \Delta_{h^{[k]}} d^{[k]} f\big(g, \partial(\alpha)_*(\mathbf{y})\big) / \prod_{\mathbf{w} \in \alpha} \chi_\mathbf{w}\big(\Delta_h \rho(g, \mathbf{y_w})\big) \,.$$

Now we crucially use the fact that $d^{[k]} f$ and ρ are cocycles to write this as

$$\frac{(T_h)_\alpha^{[k+1]}H}{H}(g,y) = \Delta_{g^{[k]}}d^{[k]}f\big(h,\partial(\alpha)_*(\mathbf{y})\big) \Big/ \prod_{\mathbf{w}\in\alpha} \chi_{\mathbf{w}}\big(\Delta_g \rho(h,\mathbf{y_w})\big).$$

The right-hand side is then a $\big(G,\mathrm{Y}^{[k+1]},S^1\big)$-coboundary, and the claim follows.

Arguing as in the proof of Proposition 4.4, we can glue various translated coboundaries together and conclude that H is a (G,Y,S^1)-coboundary, thus $d^{[k+1]}f(g,\mathbf{y})$ is $\big(G,\mathrm{Y}^{[k+1]},S^1\big)$-cohomologous to $\prod_{\mathbf{w}\in\mathbf{2}^{k+1}} \chi_{\mathbf{w}}(\rho(g,\mathbf{y_w}))$. On the other hand, the factors $\chi_{\mathbf{w}}(\rho(g,\mathbf{y_w}))$ are of degree $< m$ and thus $d^{[m]}\chi_{\mathbf{w}}(\rho(g,\mathbf{y_w})) = 0$. It then follows that $d^{[m+k+1]}f(g,\mathbf{y})$ is a $\big(G,\mathrm{Y}^{[m+k+1]},S^1\big)$-coboundary, and so f is of type $< m+k+1$ as claimed. $\qquad\square$

The second lemma allows us to reduce the type of a function by differentiation in the vertical direction.

LEMMA 5.3 (Vertical differentiation lemma). *Let G be a countable abelian group, let $k \geqslant 1$, and let $\mathrm{X} = \mathrm{Y} \times_\rho U$ be a G-system of order $< k$ for some compact abelian U and a (G,Y,U)-cocycle ρ. Let f be a (G,X,S^1)-function of type $< m$ for some $m \geqslant 1$. Then for every $t \in U$, the (G,X,S^1)-function $\boldsymbol{\Delta}_t f$ is of type $< m - \min(k,m)$, where t acts on X by the action $V_t : (y,u) \mapsto (y,tu)$ for $y \in Y$ and $u \in U$.*

Proof. By hypothesis, there exists a $\big(\mathrm{X}^{[m]},S^1\big)$-function F such that $d^{[m]}f = \boldsymbol{\Delta}F$. Let α be a $m - \min(k,m)$-dimensional side of $\mathbf{2}^m$. By Lemma A.23(iv), the side transformation $(V_t)_\alpha^{[m]}$ leaves $\mu^{[m]}$ invariant, and commutes with the $\mathrm{diag}\big(G^{[k]}\big)$-action. Thus we have $d^{[m]}(V_t)_\alpha^{[m]}f = \boldsymbol{\Delta}(V_t)_\alpha^{[m]}F$. Dividing through, we conclude

$$d^{[m]}\frac{(V_t)_\alpha^{[m]}f}{f} = \boldsymbol{\Delta}\frac{(V_t)_\alpha^{[m]}F}{F}$$

and so $d^{[m]}\frac{(V_t)_\alpha^{[m]}f}{f}$ is a $\big(G,\mathrm{X}^{[m]},S^1\big)$-coboundary. This expression can be rewritten as

$$d^{[m]}\frac{(V_t)_\alpha^{[m]}f}{f} = \big(\partial(\alpha)_*\big)^*\big(d^{[m-\min(k,m)]}\boldsymbol{\Delta}_t f\big),$$

and so $\big(\partial(\alpha)_*\big)^*\big(d^{[m-\min(k,m)]}\boldsymbol{\Delta}_t f\big)$ is a $\big(G,\mathrm{X}^{[m]},S^1\big)$-coboundary, and thus a $\big(G,\mathrm{X}^{[m]},S^1\big)$-cocycle, which (by Lemma B.1) implies that $d^{[m-\min(k,m)]}\boldsymbol{\Delta}_t f$ is a $\big(G,\mathrm{X}^{[m-1]},S^1\big)$-cocycle. Applying Lemma B.10, we conclude that $d^{[m-\min(k,m)]}\boldsymbol{\Delta}_t f$ is in fact a $(G,\mathrm{X}^{[m-\min(k,m)]},S^1)$-coboundary, and so $\boldsymbol{\Delta}_t f$ is of type $< m - \min(k,m)$ as desired. $\qquad\square$

Because of these two lemmas, Theorem 4.5 will follow from

Theorem 5.4 (Fourth reduction). *Let $k \geqslant 2$ and $m \geqslant 1$, and assume that Theorem 4.5 has already been proven for all smaller values of k. Let $\mathbb{F}$ be a finite field, and let X be an ergodic $\mathbb{F}^\omega$-system of order $< k$. Write $\mathrm{X} = \mathcal{Z}_{<k-1}(\mathrm{X}) \times_\rho U$ for some abelian $(\mathbb{F}^\omega, \mathcal{Z}_{<k-1}(\mathrm{X}), U)$-cocycle ρ. Let $f \in M_{<m}(\mathbb{F}^\omega, \mathrm{X}, S^1)$ be a function of type $< m$, be such that*

$$\boldsymbol{\Delta}_t f \in \mathcal{P}_{<m}(\mathbb{F}^\omega, \mathrm{X}, S^1) \cdot B^1(\mathbb{F}^\omega, \mathrm{X}, S^1)$$

for all $t \in U$ (i.e. $\mathbf{\Delta}_t f$ is always cohomologous to a phase polynomial of degree $< m$). Then f is cohomologous to $P\pi^* \tilde{f}$ for some $P \in \mathcal{P}_{<O_{k,m}(1)}(\mathbb{F}^\omega, \mathrm{X}, S^1)$, and some $(\mathbb{F}^\omega, \mathcal{Z}_{<k-1}(\mathrm{X}), S^1)$-function $\tilde{f}$, where π is the factor map from X to $\mathcal{Z}_{<k-1}(\mathrm{X})$.

Proof of Theorem 4.5 assuming Theorem 5.4.   Let k, m, $\mathbb{F}$, X, f be as in Theorem 4.5. We can assume inductively that Theorem 4.5 has already been proven for smaller values of k.

We claim for each $0 \leqslant j \leqslant m$ that

$$\mathbf{\Delta}_{t_1} \ldots \mathbf{\Delta}_{t_j} f \in \mathcal{P}_{<O_{k,m,j}(1)}(\mathbb{F}^\omega, \mathrm{X}, S^1) \cdot B^1(\mathbb{F}^\omega, \mathrm{X}, S^1) \qquad (5.1)$$

for all $t_1, \ldots, t_j \in U$.

We establish this by downward induction on j. When $j = m$, the claim follows by m applications of Lemma 5.3 (and no $\mathcal{P}_{<O_{k,m,j}(1)}(\mathbb{F}^\omega, \mathrm{X}, S^1)$ term appears in this case). Now suppose that $0 \leqslant j < m$ and the claim (5.1) has already been shown for $j+1$. Applying Theorem 5.4, we see that for all $t_1, \ldots, t_j$, $\mathbf{\Delta}_{t_1} \ldots \mathbf{\Delta}_{t_j} f$ is $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cohomologous to $P_{t_1,\ldots,t_j} \pi^* \tilde{f}_{t_1,\ldots,t_j}$ for some $P_{t_1,\ldots,t_j} \in \mathcal{P}_{<O_{k,m,j}(1)}(\mathbb{F}^\omega, \mathrm{X}, S^1)$ and $\tilde{f} \in M(\mathbb{F}^\omega, \mathcal{Z}_{<k-1}(\mathrm{X}), S^1)$.

Since f is of type $< m$, and the U action commutes with the G action, we see that $\mathbf{\Delta}_{t_1} \ldots \mathbf{\Delta}_{t_j} f$ is also of type $< m$. By Lemma 4.3(ii) and (iii), we conclude that the $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-function $\pi^* \tilde{f}_{t_1,\ldots,t_j}$ is of type $< O_{k,m,j}(1)$. Applying Lemma 5.1, we conclude that the $(\mathbb{F}^\omega, \mathcal{Z}_{<k-1}(\mathrm{X}), S^1)$-function $\tilde{f}_{t_1,\ldots,t_j}$ is of type $< O_{k,m,j}(1)$. Applying the inductive hypothesis, we conclude that $\tilde{f}_{t_1,\ldots,t_j}$ is $(\mathbb{F}^\omega, \mathcal{Z}_{<k-1}(\mathrm{X}), S^1)$-cohomologous to a $(\mathbb{F}^\omega, \mathcal{Z}_{<k-1}(\mathrm{X}), S^1)$-phase polynomial of degree $< O_{k,m,j}(1)$. By functoriality, this implies that $\pi^* \tilde{f}_{t_1,\ldots,t_j}$ is $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cohomologous to a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-phase polynomial of degree $< O_{k,m,j}(1)$. Since $\mathbf{\Delta}_{t_1} \ldots \mathbf{\Delta}_{t_j} f$ was $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cohomologous to $P_{t_1,\ldots,t_j} \pi^* \tilde{f}_{t_1,\ldots,t_j}$, the claim (5.1) follows.

Theorem 4.5 now follows by specializing (5.1) to the case $j = 0$. $\square$

It remains to establish Theorem 5.4.

6   Reduction to a Finite U

The purpose of this section is to obtain the following reduction.

PROPOSITION 6.1 (Reduction to finite U).   *In order to prove Theorem 5.4, it suffices to do so in the case when U is finite.*

Proof. Fix k, and assume that Theorem 5.4 has already been proven in the case of finite U. Let $m, \mathbb{F}, \mathrm{X}, \rho, U, f$ be as in Theorem 5.4. By Theorem 4.8 we may assume that ρ is a $(\mathbb{F}^\omega, \mathcal{Z}_{<k-1}(\mathrm{X}), U)$-phase polynomial of degree $< O_{k,p}(1)$. By hypothesis, for each $t \in U$ there exists $q_t \in \mathcal{P}_{<m}(\mathbb{F}^\omega, \mathrm{X}, S^1)$ and $F_t \in M(\mathrm{X}, S^1)$ such that

$$\mathbf{\Delta}_t f = q_t \mathbf{\Delta} F_t . \qquad (6.1)$$

By Lemma C.4 we can take q_t, F_t to be measurable with respect to t.

The next step is to linearize q_t on an open subgroup of U, by arguing as follows. Let $t, u \in U$. Then the cocycle identity $\mathbf{\Delta}_{tu} f = (\mathbf{\Delta}_t(V_u f)) \mathbf{\Delta}_u f$ and (6.1) give

$$q_{tu} \mathbf{\Delta} F_{tu} = (V_u q_t) q_u \mathbf{\Delta}\big((V_u F_t) F_u\big)$$

and hence

$$\Delta \frac{F_{tu}}{(V_u F_t) F_u} \in \mathcal{P}_{<m}(\mathbb{F}^\omega, \mathrm{X}, S^1)$$

and thus (by (3.1)), the function

$$\phi_{t,u} := \frac{F_{tu}}{(V_u F_t) F_u} \tag{6.2}$$

is a (X, S^1)-phase polynomial of degree $< m+1$.

Let $\varepsilon > 0$ be a small number. By Lusin's theorem, the function $t \to F_t$ is equal to a uniformly continuous function (in $L^2(\mathrm{X})$) outside of an open set E of measure $O(\varepsilon)$ in U. In particular, there exists an open neighborhood U' of the identity in U such that F_{tu} and $V_u F_t$ both lie within $O(\varepsilon)$ of F_t whenever $u \in U'$ and $t, tu \notin E$. In particular, we see that $\phi_{t,u}$ lies within $O(\varepsilon)$ of $\overline{F_u}$ when $u \in U'$ and $t, tu \notin E$. Applying Lemma C.1 and taking ε small enough, we conclude that for each $u \in U'$ there exists $\phi'_u \in \mathcal{P}_{<m+1}(\mathrm{X}, S^1)$ such that $\phi_{t,u}/\phi'_u$ is constant whenever $t, tu \notin E$. Arguing as in the proof of Lemma C.4 we can ensure that $u \mapsto \phi'_u$ is measurable.

Now let $u, v \in U'$. From (6.2) we have the second-order cocycle identity $\phi_{uv,w} \phi_{u,v} = \phi_{u,vw} V_u \phi_{v,w}$ for all $w \in U$. If ε is small enough, we can find w so that $w, vw, uvw \notin E$. We conclude that

$$\phi_{u,v} = c(u,v) \frac{\phi'_u V_u \phi'_v}{\phi'_{uv}} \tag{6.3}$$

for some constant $c(u,v)$. If we then set $F'_u := F_u/\phi'_u$ and $q'_u := q_u \Delta \phi'_u$ for $u \in U$, then q'_u is a (G, X, S^1)-phase polynomial of degree $< m$, and we conclude from (6.1) that

$$\Delta_u f = q'_u \Delta F'_u \tag{6.4}$$

for all $u \in U$, while from (6.2), (6.3) we have $\Delta \frac{F'_{uv}}{(V_u F'_v) F'_u} = 1$ and thus (by (6.4))

$$\frac{q'_{uv}}{(V_u q'_v) q'_u} = 1 \tag{6.5}$$

for $u, v \in U'$.

Now we make a crucial use of the finite characteristic hypothesis. By Lemma 4.7, U is p^n-torsion for some $n = O_k(1)$. By Lemma D.1, we conclude that U' contains an open subgroup of U; by reducing U' if necessary, we may assume that U' is in fact equal to an open subgroup.

Now we pass from U to U' as follows. By Lemma D.2, we may write $U = U' \times W$ for some finite group W (shrinking U' if necessary). We can then factorise $\mathrm{X} = \mathcal{Z}_{<k-1}(\mathrm{X}) \times_\rho U$ as $\mathrm{X} = \mathrm{Y} \times_{\rho'} U'$, where $\mathrm{Y} := \mathcal{Z}_{<k-1}(\mathrm{X}) \times_{\rho'} W$ and ρ', ρ'' are the projections of ρ to U' and W respectively. Note that ρ' is a $(\mathbb{F}^\omega, \mathrm{Y}, U')$-phase polynomial of degree $< O_k(1)$ (since ρ is also).

Applying Lemma B.6 once for each $g \in \mathbb{F}^\omega$ and then pasting together, we may write $q'_u = \Delta_u Q$ for all $u \in U$ and some (G, X, S^1)-phase polynomial Q of degree $< O_{k,m}(1)$. If we let f' be the $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-function $f' := f/Q$, we thus conclude from (6.4) that $\Delta_u f' = \Delta F'_u$ for all $u \in U'$.

The compact abelian group U acts freely on X in a manner commuting with the G action, and so the compact abelian subgroup U' does also. We can thus apply Lemma B.8 and conclude that f' is $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cohomologous to a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-function f'' which is invariant with respect to some open subgroup U'' of U'.

Let $\sigma : U \to U/U''$ be the quotient map, let $\mathrm{X}' := \mathcal{Z}_{<k-1}(\mathrm{X}) \times_{\sigma \circ \rho} U/U''$, and let $\pi : \mathrm{X} \to \mathrm{X}'$ be the associated factor map. Then we can write $f'' = \pi^* \tilde{f}$ for some $(\mathbb{F}^\omega, \mathrm{X}', S^1)$-function $\tilde{f}$. By functoriality, X' is of order $< k$.

By construction, $\pi^* \tilde{f}$ is cohomologous to $f' = f/Q$. Since f has type $< m$ and Q is a phase polynomial of degree $< O_{k,m}(1)$, we see from Lemma 4.3(ii) and (iii) that $\pi^* \tilde{f}$ is a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-function of type $< O_{k,m}(1)$.

From (6.1) and the polynomial nature of Q, we know that $\pi^* \pmb{\Delta}_t \tilde{f} = \pmb{\Delta}_t \pi^* \tilde{f}$ is $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cohomologous to a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-phase polynomial of degree $< O_{k,m}(1)$. By Lemma B.11, $\pmb{\Delta}_t \tilde{f}$ is thus $(\mathbb{F}^\omega, \mathrm{X}', S^1)$-cohomologous to a $(\mathbb{F}^\omega, \mathrm{X}', S^1)$-phase polynomial of degree $< O_{k,m}(1)$, times a function of the form $\chi \circ \rho \circ \pi^{\mathrm{X}'}_{\mathcal{Z}_{<k-1}(\mathrm{X})}$ for some $\chi \in \hat{U}$. But recall that ρ is a $(\mathbb{F}^\omega, \mathcal{Z}_{<k-1}(\mathrm{X}), U)$-phase polynomial of degree $< O_k(1)$. We conclude that $\pmb{\Delta}_t \tilde{f}$ is $(\mathbb{F}^\omega, \mathrm{X}', S^1)$-cohomologous to a $(\mathbb{F}^\omega, \mathrm{X}', S^1)$-phase polynomial of degree $< O_{k,m}(1)$. We can now invoke Theorem 5.4 for the finite group U/U'' and conclude that $\tilde{f}$ is $(\mathbb{F}^\omega, \mathrm{X}', S^1)$-cohomologous to $P\big(\pi^{\mathrm{X}'}_{\mathcal{Z}_{<k-1}(\mathrm{X})}\big)^* \tilde{f}'$ for some $(\mathbb{F}^\omega, \mathrm{X}', S^1)$-phase polynomial P of degree $< O_{k,m}(1)$, and some $(\mathbb{F}^\omega, \mathcal{Z}_{<k-1}(\mathrm{X}), S^1)$-function $\tilde{f}'$. Since f is cohomologous to $Q\pi^* \tilde{f}$, we obtain the desired factorization of f required to establish Theorem 5.4. $\qquad\square$

7   The Finite Group Case

PROPOSITION 7.1 (The finite case).   *Theorem 5.4 is true when U is finite.*

Proof. Let $k, m, \mathbb{F}, \mathrm{X}, \rho, U, f$ be as in Theorem 5.4, and let p be the characteristic of $\mathbb{F}$. By Lemma 4.7, U is a finite abelian group which is p^n-torsion for some $n = O_k(1)$, and so by the classification of finite abelian groups we may assume that

$$U = C_{p^{n_1}} \times \ldots \times C_{p^{n_N}}$$

for some $1 \leqslant n_1, \ldots, n_N \leqslant O_k(1)$ and some finite (but unbounded) N. When p is large enough, depending on k, we can take $n_1 = \ldots = n_N = 1$.

Let $e_1, \ldots, e_N$ be the standard set of generators of U. By hypothesis, for every $1 \leqslant j \leqslant N$ we can find $p_j \in \mathcal{P}_{<m}(\mathbb{F}^\omega, \mathrm{X}, S^1)$ and $F_j \in M(\mathrm{X}, S^1)$ such that

$$\pmb{\Delta}_{e_j} f = p_j \pmb{\Delta} F_j . \tag{7.1}$$

The idea here is to express F_j as $\pmb{\Delta}_{e_j} F$ times a polynomial error, for some F independent of j; we will then "integrate" this to express f as $\pmb{\Delta} F$ times a polynomial error, times a function invariant under $e_1, \ldots, e_N$; this is basically what we need to establish Theorem 5.4.

We turn to the details. The first task is to measure two potential obstructions to F_j being expressible as $\pmb{\Delta}_{e_j} F$ (modulo polynomial errors), namely the obstruction coming from the torsion of U, and the obstruction coming from the multi-dimensionality of U.

Observe that we have the telescoping identity

$$\prod_{t=0}^{p^{n_j}-1} V_{e_j^t} \boldsymbol{\Delta}_{e_j} f = 1 \tag{7.2}$$

and thus by (7.1)

$$\boldsymbol{\Delta} \prod_{t=0}^{p^{n_j}-1} V_{e_j^t} F_j \in \mathcal{P}_{<m}(\mathbb{F}^\omega, \mathrm{X}, S^1)$$

and so (by (3.1)) we have

$$\prod_{t=0}^{p^{n_j}-1} V_{e_j^t} F_j \in \mathcal{P}_{<m+1}(\mathrm{X}, S^1)\,. \tag{7.3}$$

Note, conversely, that $\prod_{t=0}^{p^{n_j}-1} V_{e_j^t} F_j$ needed to be polynomial in order to have any chance to express F_j as $\boldsymbol{\Delta}_{e_j} F$ times a polynomial.

Suppose now that p is sufficiently large depending on k, so that $n_j = 1$. By Lemma D.3 we have $\prod_{t=0}^{p^{n_j}-1} V_{e_j^t} p_j = 1$, and hence by (7.2), (7.1), we may now strengthen (7.3) to

$$\boldsymbol{\Delta} \prod_{t=0}^{p^{n_j}-1} V_{e_j^t} F_j = 1\,.$$

In a similar fashion, from the commutation identity $\boldsymbol{\Delta}_{e_i} \boldsymbol{\Delta}_{e_j} f / \boldsymbol{\Delta}_{e_j} \boldsymbol{\Delta}_{e_i} f = 1$ for any $1 \leqslant i, j \leqslant N$, we see from Lemma B.5(i) and (7.1) that

$$\boldsymbol{\Delta} \left(\frac{\boldsymbol{\Delta}_{e_i} F_j}{\boldsymbol{\Delta}_{e_j} F_i} \right) \in \mathcal{P}_{<m-1}(\mathbb{F}^\omega, \mathrm{X}, S^1)$$

and hence by (3.1) we have

$$\frac{\boldsymbol{\Delta}_{e_i} F_j}{\boldsymbol{\Delta}_{e_j} F_i} \in \mathcal{P}_{<m}(\mathbb{F}^\omega, \mathrm{X}, S^1)\,. \tag{7.4}$$

Again, observe that $\boldsymbol{\Delta}_{e_i} F_j / \boldsymbol{\Delta}_{e_j} F_i$ had to be polynomial in order to have a chance to express F_i, F_j as $\boldsymbol{\Delta}_{e_i} F, \boldsymbol{\Delta}_{e_j} F$ modulo polynomials.

We now clean up the first condition (7.3). We claim that there exists a e_j-invariant phase polynomial $\Psi_j \in \mathcal{P}_{<O_{k,m}(1)}(\mathrm{X}, S^1)$ such that $\prod_{t=0}^{p^{n_j}-1} V_{e_j^t} F_j = \Psi_j^{p^{n_j}}$.

When p is sufficiently large depending on k, then as observed before, $\prod_{t=0}^{p^{n_j}-1} V_{e_j^t} F_j$ is constant, and the claim is trivial, so suppose instead then $p = O_k(1)$. By (7.3), the (X, S^1)-function $\prod_{t=0}^{p^{n_j}-1} V_{e_j^t} F_j$ is polynomial of degree $< m + 1$. By inspection, it is also e_j-invariant. Quotienting out by the e_j action, applying Corollary D.7, and pulling back, we obtain the claim in this case.

If we then write $\tilde{F}_j := F_j / \Psi_j$, then by construction we have

$$\prod_{t=0}^{p^{n_j}-1} V_{e_j^t} \tilde{F}_j = 1 \tag{7.5}$$

and

$$\boldsymbol{\Delta}_{e_j} f \in (\boldsymbol{\Delta} F_j) \cdot \mathcal{P}_{<O_{k,m}(1)}(\mathbb{F}^\omega, \mathrm{X}, S^1) \,, \tag{7.6}$$

while from (7.4) we have

$$\frac{\boldsymbol{\Delta}_{e_i} \tilde{F}_j}{\boldsymbol{\Delta}_{e_j} \tilde{F}_i} \in \mathcal{P}_{<O_{k,m}(1)}(\mathbb{F}^\omega, \mathrm{X}, S^1) \,. \tag{7.7}$$

Note from (7.5) and Lemma D.3 that the $\tilde{F}_i$ take values in C_{p^M} for some $M = O_{k,m}(1)$; when p is sufficiently large depending on k, m, then we can take $M = 1$.

Recall that in the N-dimensional Euclidean space $\mathbb{R}^N$, a (smooth) vector field $(\tilde{F}_i)_{1 \leqslant i \leqslant N}$ which obeys the curl-free condition $\frac{\partial F_i}{\partial x_j} - \frac{\partial F_j}{\partial x_i} = 0$ can be expressed as a gradient $\tilde{F}_i = \partial F / \partial x_i$, where F can be given explicitly by the formula

$$F\big((t_1, \ldots, t_N)\big) := \sum_{i=1}^N \int_0^{t_i} \tilde{F}_i\big((t_1, \ldots, t_{i-1}, t_i', 0, \ldots, 0)\big) dt_i' \,.$$

If in addition the $\tilde{F}_i$ are periodic modulo $\mathbb{Z}^N$, and obey the necessary condition $\int_0^1 \tilde{F}_i(x + t_i e_i) dt_i = 0$ for all $1 \leqslant i \leqslant N$ and $x \in \mathbb{R}^N$, then the function F defined above is also periodic modulo $\mathbb{Z}^N$ and descends to the torus $\mathbb{R}^N / \mathbb{Z}^N$.

We can perform exactly the same construction in our finite multiplicative setting. By Theorem 4.8, we may write $\mathrm{X} = \mathcal{Z}_{<k-1}(\mathrm{X}) \times_\rho U$ where ρ is a $(\mathbb{F}^\omega, \mathcal{Z}_{<k-1}(\mathrm{X}), U)$-phase polynomial cocycle of degree $< O_k(1)$. We adopt the notation

$$[t_1, \ldots, t_N] := e_1^{t_1} \ldots e_N^{t_N}$$

for any integers $t_1, \ldots, t_N$, thus for instance $[t_1, \ldots, t_N]$ is periodic in t_j with period p^{n_j}. We let F be the (X, S^1) function defined by the formula

$$F\big(y, [t_1, \ldots, t_N]\big) := \prod_{i=1}^N \prod_{0 \leqslant t_i' < t_i} \tilde{F}_i\big(y, [t_1, \ldots, t_{i-1}, t_i', 0, \ldots, 0]\big)$$

for $y \in \mathcal{Z}_{<k-1}(\mathrm{X})$ and $t_1, \ldots, t_N \in \mathbb{Z}$, with the convention that $\prod_{0 \leqslant t_i' < t_i} a_{t_i'} := \big(\prod_{t_i \leqslant t_i' < 0} a_{t_i'}\big)^{-1}$ when t_i is negative. Note from (7.5) that the right-hand side here is periodic in t_j with period p^{n_j}, and so F is well-defined. Since the $\tilde{F}_i$ take values in a cyclic group C_{p^M}, F does also.

Now we compute a derivative of F. We clearly have

$$\boldsymbol{\Delta}_{e_j} F\big(y, [t_1, \ldots, t_N]\big) = \prod_{i=1}^N \prod_{0 \leqslant t_i' < t_i} \boldsymbol{\Delta}_{e_j} \tilde{F}_i\big(y, [t_1, \ldots, t_{i-1}, t_i', 0, \ldots, 0]\big)$$

for any $1 \leqslant j \leqslant N$. On the other hand, we have the telescoping identity

$$\prod_{i=1}^N \prod_{0 \leqslant t_i' < t_i} \boldsymbol{\Delta}_{e_i} \tilde{F}_j\big(y, [t_1, \ldots, t_{i-1}, t_i', 0, \ldots, 0]\big) = \tilde{F}_j\big(y, [t_1, \ldots, t_N]\big) / \tilde{F}_j(y, 1)$$

and thus

$$\boldsymbol{\Delta}_{e_j} F\big(y, [t_1, \ldots, t_N]\big) = \frac{\tilde{F}_j(y, [t_1, \ldots, t_N])}{\tilde{F}_j(y, 1)} \prod_{i=1}^{N} \prod_{0 \leqslant t_i' < t_i} \omega_{ij}\big(y, [t_1, \ldots, t_{i-1}, t_i', 0, \ldots, 0]\big)$$

(7.8)

where $\omega_{ij} := \boldsymbol{\Delta}_{e_j} \tilde{F}_i / \boldsymbol{\Delta}_{e_i} \tilde{F}_j$.

Since ρ has degree $< O_k(1)$, the map $(y, u) \mapsto u$ is a (X, U)-phase polynomial of degree $< O_k(1)$, which implies for any fixed $1 \leqslant i \leqslant N$ and $0 \leqslant t_i' < p^{n_i}$ that the map $(y, [t_1, \ldots, t_N]) \mapsto [t_1, \ldots, t_{i-1}, t_i', 0, \ldots, 0]$ is also a (X, U)-phase polynomial of degree $< O_k(1)$, since the map from $[t_1, \ldots, t_N]$ to $[t_1, \ldots, t_{i-1}, t_i', 0, \ldots, 0]$ is a homomorphism. By Lemma B.5(iii), we conclude that the functions $(y, [t_1, \ldots, t_N]) \mapsto \tilde{F}_i(y, [t_1, \ldots, t_{i-1}, t_i', 0, \ldots, 0])$ are (X, S^1)-phase polynomials of degree $< O_{k,m}(1)$ for all $1 \leqslant i \leqslant N$. The map $(y, [t_1, \ldots, t_N]) \mapsto t_i$ is also a (X, S^1)-phase polynomial of degree $< O_{k,m}(1)$. Also, by (7.7), the ω_{ij} are (X, S^1)-phase polynomials of degree $< O_{k,m}(1)$. We now claim that

$$\big(y, [t_1, \ldots, t_N]\big) \mapsto \prod_{0 \leqslant t_i' < t_i} \omega_{ij}\big(y, [t_1, \ldots, t_{i-1}, t_i', 0, \ldots, 0]\big)$$

is also a (X, S^1)-phase polynomial of degree $< O_{k,m}(1)$.

When $p = O_{k,m}(1)$, this claim follows from Corollary D.6, so suppose now that p is sufficiently large depending on k, m. Then $\omega_{ij}(y, [t_1, \ldots, t_{i-1}, t_i', 0, \ldots, 0])$ is a phase polynomial of degree $O_{k,m}(1)$ in t_i' that takes values in C_p. By Taylor expansion we may thus write

$$\omega_{ij}\big(y, [t_1, \ldots, t_{i-1}, t_i', 0, \ldots, 0]\big) = \prod_{0 \leqslant j \leqslant O_{k,m}(1)} \big[\boldsymbol{\Delta}_{e_i}^j \omega_{ij}(y, [t_1, \ldots, t_{i-1}, 0, 0, \ldots, 0])\big]^{\binom{t_i'}{j}}$$

and thus

$$\prod_{0 \leqslant t_i' < t_i} \omega_{ij}\big(y, [t_1, \ldots, t_{i-1}, t_i', 0, \ldots, 0]\big)$$

$$= \prod_{0 \leqslant j \leqslant O_{k,m}(1)} \big[\boldsymbol{\Delta}_{e_i}^j \omega_{ij}(y, [t_1, \ldots, t_{i-1}, 0, 0, \ldots, 0])\big]^{\binom{t_i}{j+1}}.$$

The claim now follows from Lemma B.5.

Inserting the above claim into (7.8) we conclude that

$$\tilde{F}_j \in \boldsymbol{\Delta}_{e_j} F(\pi^* F_j') \cdot \mathcal{P}_{<O_{k,m}(1)}(\mathbb{F}^\omega, \mathrm{X}, S^1)$$

where $F_j'(y, u) := \tilde{F}_j(y, 1)$, and hence by (7.6)

$$\boldsymbol{\Delta}_{e_j} f \in (\boldsymbol{\Delta}_{e_j} \boldsymbol{\Delta} F) \cdot (\pi^* \boldsymbol{\Delta} F_j') \cdot \mathcal{P}_{<O_{k,m}(1)}(\mathbb{F}^\omega, \mathrm{X}, S^1).$$

Thus if we set $f' := f/\boldsymbol{\Delta} F$, then f is cohomologous to f' and

$$\boldsymbol{\Delta}_{e_j} f' \in (\pi^* \boldsymbol{\Delta} F_j') \cdot \mathcal{P}_{<O_{k,m}(1)}(\mathbb{F}^\omega, \mathrm{X}, S^1).$$

(7.9)

Now we need to work on the F_j' term. From the telescoping identity $\prod_{0 \leqslant t_j < p^{n_j}} V_{e_j}^{t_j} \boldsymbol{\Delta}_{e_j} f' = 1$ and (7.9), we have

$$\pi^* \boldsymbol{\Delta}(F_j')^{p^{n_j}} \in \mathcal{P}_{<O_{k,m}(1)}(\mathbb{F}^\omega, \mathrm{X}, S^1);$$

pushing forward by π and then using (3.1), we conclude

$$(F_j')^{p^{n_j}} \in \mathcal{P}_{<O_{k,m}(1)}\left(\mathbb{F}^\omega, \mathcal{Z}_{<k-1}(X), S^1\right).$$

In the case that p is sufficiently large depending on k, m, we see from Lemma D.3 that we have the improvement $\pi^* \mathbf{\Delta}(F_j')^{p^{n_j}} = 1$, and so $(F_j')^{p^{n_j}}$ is constant in this case.

We now claim that $F_j' = q_j F_j''$, where $q_j \in \mathcal{P}_{<O_{k,m}(1)}(\mathbb{F}^\omega, \mathcal{Z}_{<k-1}(X), S^1)$ and F_j'' takes values in $C_{p^{n_j}}$. For p sufficiently large depending on k, m, this is immediate from the previous discussion; for $p = O_{k,m}(1)$, the claim follows instead from Corollary D.7.

Inserting this claim back into (7.9) we obtain

$$\mathbf{\Delta}_{e_j} f' \in (\pi^* \mathbf{\Delta} F_j'') \cdot \mathcal{P}_{<O_{k,m}(1)}(\mathbb{F}^\omega, X, S^1). \tag{7.10}$$

Now let F^* be the (X, S^1)-function $F^*(y, [t_1, \ldots, t_N]) := \prod_{j=1}^{N} F_j''(y)^{t_j}$; this is well-defined since F_j'' takes values in $C_{p^{n_j}}$. We observe that $\pi^* \mathbf{\Delta} F_j'' = \mathbf{\Delta}_{e_j} \mathbf{\Delta} F^*$. Thus if $f'' := f'/F^*$, then f'' is (X, S^1)-cohomologous to f and $\mathbf{\Delta}_{e_j} f'' \in \mathcal{P}_{<O_{k,m}(1)}(\mathbb{F}^\omega, X, S^1)$ for all $1 \leqslant j \leqslant N$. By repeated use of the cocycle identity $\mathbf{\Delta}_{uu'} f'' = (\mathbf{\Delta}_u f) V_u (\mathbf{\Delta}_{u'} f)$ we conclude that $\mathbf{\Delta}_u f'' \in \mathcal{P}_{<O_{k,m}(1)}(\mathbb{F}^\omega, X, S^1)$ for all $u \in U$. Applying Lemma B.6 once for each $g \in \mathbb{F}^\omega$ we may thus write $\mathbf{\Delta}_u f'' = \mathbf{\Delta}_u P$ for some $P \in \mathcal{P}_{<O_{k,m}(1)}(\mathbb{F}^\omega, X, S^1)$. Thus f''/P is U-invariant, and so $f'' = P \pi^* \tilde{f}$ for some $(\mathbb{F}^\omega, \mathcal{Z}_{<k-1}(X), S^1)$-function $\tilde{f}$, and Proposition 7.1 follows. $\square$

From Proposition 7.1 and Proposition 6.1, we obtain Theorem 5.4, and thus Theorem 3.3.

8   The High-Characteristic Case

We now develop high-characteristic analogues of the above theory, establishing the sharp Theorem 1.19 instead of Theorem 1.20 in this setting. The arguments here will be similar to those used to prove Theorem 1.20; the main new difficulty is to be careful to not lose anything in the degree of various functions beyond what is absolutely necessary.

8.1   Preliminary reductions.

Just as Theorem 1.20 follows from Theorem 4.5, Theorem 1.19 will follow from

Theorem 8.1 (First reduction of high-characteristic case). *Let $\mathbb{F}$ be a finite field, let $1 \leqslant k \leqslant \operatorname{char}(\mathbb{F})$, and let X be an ergodic $\mathbb{F}^\omega$-system of order $< k - 1$. Let $f \in Z_{<k}^1(\mathbb{F}^\omega, X, S^1)$ be a $(\mathbb{F}^\omega, X, S^1)$-cocycle of type $< k$. Then f is $(\mathbb{F}^\omega, X, S^1)$-cohomologous to a $(\mathbb{F}^\omega, X, S^1)$-phase polynomial of degree $< k$.*

The deduction of Theorem 1.19 from Theorem 8.1 is completely analogous to the corresponding derivation of Theorem 1.20 from Theorem 4.5 and is omitted. It remains to establish Theorem 8.1. It should not be surprising that this will be accomplished by an induction on k. But it will also be convenient to induct on a secondary parameter j, measuring the order of X. For technical reasons, in this

induction the function f will no longer be a cocycle, but instead have two weaker properties, which we refer to as being a *quasi-cocycle* and a *line cocycle* respectively:

DEFINITION 8.2 (Quasi-cocycles).   Let G be a countable abelian group, let X be an ergodic G-system, and let f be a (G, X, S^1)-function. Let $k \geq 0$. We say that f is a (G, X, S^1)-*quasi-cocycle of order* $< k$ if, for every $g, g' \in G$, one has

$$f(g + g', x) = f(g, T_{g'}x)f(g', x)p_{g,g'}(x)$$

for some $p_{g,g'} \in \mathcal{P}_{<k}(\mathrm{X}, S^1)$.

EXAMPLES 8.3. A (G, X, S^1)-cocycle is precisely a (G, X, S^1)-quasi-cocycle of order < 0. Every quasi-cocycle of order $< k$ is of course a quasi-cocycle of order $< k + 1$; in particular, cocycles are quasi-cocycles of every order. Every (G, X, S^1)-phase polynomial of order $< k$ is also a (G, X, S^1)-quasi-cocycle of order $< k$. The space of (G, X, S^1)-quasi-cocycles of order $< k$ form a group. One can of course define this concept for compact abelian groups other than S^1, but we will only need the S^1 quasi-cocycles in our arguments.

DEFINITION 8.4 (Line cocycle). Let $\mathbb{F}$ be a finite field of characteristic p, let X be an ergodic $\mathbb{F}^\omega$-system, and let f be a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-function. We say that f is a *line cocycle* if for every $g \in \mathbb{F}^\omega$, $\prod_{j=0}^{p-1} f(g, T_{g^j}x) = 1$ for μ-a.e. x.

EXAMPLE 8.5. Every $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cocycle is a line cocycle, and the set of line cocycles forms a group. In particular, any function cohomologous to a line cocycle is again a line cocycle.

It is clear that Theorem 8.1 then follows from the $j = k$ case of

Theorem 8.6 (Second reduction of high-characteristic case).   *Let* $\mathbb{F}$ *be a finite field, let* $1 \leqslant j, k \leqslant \mathrm{char}(\mathbb{F})$, *and let* X *be an ergodic* $\mathbb{F}^\omega$-*system of order* $< j$. *Let* $f \in M_{<k}(\mathbb{F}^\omega, \mathrm{X}, S^1)$ *be a* $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-*function of type* $< k$ *which is also a line cocycle and a quasi-cocycle of order* $< k - 1$. *Then* f *is* $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-*cohomologous to a* $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-*phase polynomial* P *of degree* $< k$; *furthermore,* P *takes values in the cyclic group* $C_p = \{z \in S^1 : z^p = 1\}$.

8.2 Vertical differentiation. We now begin the proof of Theorem 8.6. We first deal with the easy case $k = 1$. In this case, f is a quasi-cocycle of order < 0, and is thus a cocycle. By hypothesis, $d^{[1]}f$ is a $(\mathbb{F}^\omega, \mathrm{X}^{[1]}, S^1)$-coboundary, which by Lemma B.9 implies that f is cohomologous to a constant $c(g)$, thus $f(g, x) = c(g)\boldsymbol{\Delta}_g F(x)$ for some (X, S^1)-function F. Since f is a cocycle, c is a character, and thus takes values in C_p, and the claim follows.

Now suppose that $2 \leqslant k \leqslant \mathrm{char}(\mathbb{F})$ and assume inductively that the claim has already been proven for smaller values of k. In particular, Theorem 8.1 holds for smaller values of k. On the other hand, by repeating the proof of Theorem 4.8, we have the following structure theorem:

COROLLARY 8.7 (Exact structure theorem, high characteristic).   *Let* $\mathbb{F}$ *be a finite field, and let* $1 \leqslant k \leqslant \mathrm{char}(\mathbb{F})$ *be such that Theorem 8.1 holds for all values smaller equal than* k. *Let* X *be an ergodic* $\mathbb{F}^\omega$-*system. Then* $\mathcal{Z}_{<1}(\mathrm{X})$ *is trivial, and for all*

$2 \leqslant j \leqslant k$, we can write $\mathcal{Z}_{<j}(\mathrm{X}) \equiv \mathcal{Z}_{<j-1}(\mathrm{X}) \times_{\rho_{j-1}} U_{j-1}$, where U_{j-1} is char($\mathbb{F}$)-torsion and ρ_{j-1} is a $(\mathcal{Z}_{<j-1}(\mathrm{X}), U_{j-1})$-phase polynomial of degree $< j - 1$. In particular, we have

$$\mathcal{Z}_{<k}(\mathrm{X}) \equiv U_0 \times_{\rho_1} U_1 \times_{\rho_2} \ldots \times_{\rho_{k-1}} U_{k-1}$$

where U_0 is trivial.

Note that the torsion of the groups U_j here is just char($\mathbb{F}$) rather than a power of char($\mathbb{F}$), due to the high-characteristic hypothesis (see Lemma 4.7). By hypothesis, Corollary 8.7 is applicable for our fixed value of k.

When $j = 1$, then X is a point, and the claim is trivial. Now suppose $2 \leqslant j \leqslant$ char($\mathbb{F}$) and assume inductively that the claim has already been proven for the same value of k and all smaller values of j.

We first deal with the low-order case $j \leqslant k$, returning to the high-order case $j > k$ later. We use Corollary 8.7 to write $\mathrm{X} = \mathcal{Z}_{<j-1}(\mathrm{X}) \times_{\rho_{j-1}} U_{j-1}$.

Let $t \in U_{j-1}$. We observe the following properties of the $(\mathbb{F}^{\omega}, \mathrm{X}, S^1)$-function $\pmb{\Delta}_t f$:

LEMMA 8.8 (Exact differentiation). $\pmb{\Delta}_t f$ is a line cocycle, is of type $< k - j + 1$, and is a quasi-cocycle of order $< k - j$.

Proof. The first claim follows from the fact that f is a line cycle, and that the action V_t of t on X commutes with the $\mathbb{F}^{\omega}$ action. The second claim follows from Lemma 5.3.

Finally, we prove the quasi-cocycle claim. As the action V_t of t commutes with the action of $\mathbb{F}^{\omega}$, it suffices from Definition 8.2 to show that $\pmb{\Delta}_t$ maps $\mathcal{P}_{<k-1}(\mathrm{X}, S^1)$ to $\mathcal{P}_{<k-j}(\mathrm{X}, S^1)$.

Let $f \in \mathcal{P}_{<k-1}(\mathrm{X}, S^1)$, then by Lemma 4.3(iii), $d^{[k-1]} f = 0 \ \mu^{[k-1]}$-a.e. Let α be a $(k-j)$-face of $\mathbf{2}^{k-1}$. By Lemma A.23(iv), $(V_t)_{\alpha}^{[k-1]}$ preserves $\mu^{[k-1]}$. We conclude that $(V_t)_{\alpha}^{[k-1]}(d^{[k-1]} f) = 0 \ \mu^{[k-1]}$-a.e. Dividing these two equations, we conclude that

$$(V_t)_{\alpha}^{[k-1]}(d^{[k-1]} f) / (d^{[k-1]} f) = (\partial(\alpha)_*)^*(d^{[k-j]} \pmb{\Delta}_t f) = 0$$

$\mu^{[k-1]}$-a.e. Since $\partial(\alpha)_*$ pushes forward $\mu^{[k]}$ to $\mu^{[k-j]}$ (see Lemma A.15), we conclude that $d^{[k-j]} \pmb{\Delta}_t f = 0 \ \mu^{[k-j]}$-a.e.. Applying Lemma 4.3(iii) again we conclude that $\pmb{\Delta}_t f \in \mathcal{P}_{<k-j}(\mathrm{X}, S^1)$ as required. $\qquad \square$

By the induction hypothesis, $\pmb{\Delta}_t f$ is $(\mathbb{F}^{\omega}, \mathrm{X}, S^1)$-cohomologous to a $(\mathbb{F}^{\omega}, \mathrm{X}, S^1)$-phase polynomial q_t of degree $< k - j + 1$ which takes values in C_p. Since $\pmb{\Delta}_t f$ is a line cocycle and a quasi-cocycle of order $< k - j$, q_t is also.

8.3 Reduction to the finite U case. We now argue (as in Proposition 6.1) that in order to conclude the proof of Theorem 8.6, it suffices to do so in the case when the vertical structure group U_{j-1} is finite.

To do this, we follow the procedure as in Proposition 6.1, to make q_t a U'-cocycle for an open subgroup U' of U_{j-1}. Note that the modifications done to q_t in this procedure are by $\mathbb{F}^{\omega}$-cocycles of degree $< k - j + 1$, so none of the properties

of q_t are damaged (it is still a $k-j+1$ side cocycle, a line cocycle, and of the correct degree).

As in Proposition 6.1, we can write $U_{j-1} = U' \times W$ for some finite W, and write $\mathrm{X} = \mathrm{Y} \times_{\rho'} U'$, where $\mathrm{Y} := \mathcal{Z}_{<j-1}(\mathrm{X}) \times_{\rho''} W$ and ρ', ρ'' are the projections of ρ_{j-1}. Note that as ρ_{j-1} is of degree $< j$, ρ' is also. We write $x \in \mathrm{X}$ as $x = (y, u)$ for $y \in \mathrm{Y}$ and $u \in U'$.

We now invoke the following variant of Lemma B.6 which is more efficient with the degree.

PROPOSITION 8.9 (Exact integration).   *Let G be a countable abelian group, let $j \geqslant 0$, let U be a compact abelian group, and let $\mathrm{X} = \mathrm{Y} \times_\rho U$ be an ergodic G-system with $\mathrm{Y} \geqslant \mathcal{Z}_{<j}(\mathrm{X})$, where ρ a (G, Y, U)-phase polynomial cocycle of degree $< j$. For any $t \in U$, let $p_t(x)$ be a (X, S^1)-phase polynomial of degree $< l$, and suppose that for any $t, s \in U$, $p_{t \cdot s}(x) = p_t(V_s x) \cdot p_s(x)$. Then there exists a (X, S^1)-phase polynomial $Q(x)$ of degree $< l + j$ such that $\mathbf{\Delta}_t Q(x) = p_t(x)$. Furthermore, we can take $Q(y, u u_0) := p_u(y, u_0)$ for some $u_0 \in U$.*

REMARK 8.10. In the converse direction, one can show (by using the properties of the nilpotent group $\mathcal{G}^{[k]}$ studied in [HK]) that if Q has degree $< l + j$, then $\mathbf{\Delta}_t Q$ has degree $< l$. This may explain the terminology "exact".

We will prove this proposition in section 8.6. Assuming it for now, we see (by applying it once for each g) that we can write $q_t = \mathbf{\Delta}_t Q$ for all $t \in U'$ and some $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-phase polynomial of degree $< k$, such that $Q(g, y, u u_0) = q_u(g, y, u_0)$ for all $y \in \mathrm{Y}$, $u \in U'$, and some $u_0 \in U'$. Since q_t takes values in C_p, we see that Q does also.

We now claim that Q is a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-quasi-cocycle of order $< k-1$. Indeed, for any $g, h \in G$ and $x = (y, u u_0) \in \mathrm{X}$, we have

$$
\frac{Q(g+h, x)}{Q(g, x) Q(h, T_g x)} = \frac{Q(g+h, x)}{Q(g, x) Q(h, x) \mathbf{\Delta}_g Q(h, x)}
$$

$$
= \frac{q_u(g+h, y, u_0)}{q_u(g, y, u_0) q_u(h, y, u_0) \mathbf{\Delta}_g Q(h, x)}
$$

$$
= \frac{q_u(g+h, y, u_0)}{q_u(g, y, u_0) q_u(h, T_g y, \rho_{j-1}(g, y) u_0)} \frac{q_u(h, T_g y, \rho_{j-1}(g, y) u_0)}{q_u(h, y, u_0) \mathbf{\Delta}_g Q(h, x)}
$$

$$
= P_{u,g,h}(y, u_0) \frac{\mathbf{\Delta}_g q_u(h, x)}{\mathbf{\Delta}_g Q(h, x)}
$$

where $P_{u,g,h}(x) := q_u(g+h, x)/q_u(g, x) q_u(h, T_g x)$. Fix g, h. Since q_u and Q are $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-phase polynomials of degree $< k-j+1$ and $< k$ respectively, we see that $\mathbf{\Delta}_g q_u(h, x)/\mathbf{\Delta}_g Q(h, x)$ is a phase polynomial in x of degree $< k-1$. Also, as q_u is a quasi-cocycle of order $< k-j$, we see that $P_{u,g,h}$ is a (X, S^1)-phase polynomial of degree $< k-j$. Since q_u is a cocycle in u, $P_{u,g,h}$ is also. Applying Proposition 8.9 we conclude that $P_{u,g,h}(y, u_0)$ is a (X, S^1)-phase polynomial of degree $< k-1$. Putting this all together we see that $Q(g+h, x)/Q(g, x) Q(h, T_g x)$ is a (X, S^1)-phase polynomial of degree $< k-1$, and the claim follows.

Write $f' := f/Q$, then (as in Proposition 6.1) f' is a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-function with the property that $\boldsymbol{\Delta}_u f'$ is a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-coboundary for all $u \in U'$. Applying Lemma B.8 just as in Proposition 6.1, we conclude that f' is $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cohomologous to an $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-function f'' which is invariant with respect to some open subgroup U'' of U'. Thus we can write $f'' = \pi^* \tilde{f}$ for some $(\mathbb{F}^\omega, \mathrm{X}', S^1)$-function $\tilde{f}$, where $\mathrm{X}' := \mathcal{Z}_{<j}(\mathrm{X}) \times_{\sigma \circ \rho_{<j-1}} U_{j-1}/U''$, $\sigma : U_{j-1} \to U_{j-1}/U''$ is the quotient map, and $\pi : \mathrm{X} \to \mathrm{X}'$ is the factor map.

Since Q takes values in C_p and is of degree $< p$, we see from Lemma D.3 that Q is a line cocycle. Since $\pi^* \tilde{f}$ is cohomologous to f/Q, we conclude that $\pi^* \tilde{f}$, and hence $\tilde{f}$, are also line cocycles. Similarly, since f, Q are $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-quasi-cocycles of order $< k-1$, $\pi^* \tilde{f}$ is also, which implies in turn that $\tilde{f}$ is a $(\mathbb{F}^\omega, \mathrm{X}', S^1)$-quasi-cocycle of order $< k-1$ (cf. Lemma B.1).

From Lemma 4.3(iii) see that Q has type $< k$. Since f also has type $< k$, we conclude that $\pi^* \tilde{f}$ has type $< k$ also. We now use the following variant of Lemma 5.1, which does not concede any losses in the type:

PROPOSITION 8.11 (Exact descent). *Let G be a discrete countable abelian group and $k \geqslant 0$. Let X be an ergodic G-system of order $< k$. Let Y be a factor of X, with factor map $\pi : \mathrm{X} \to \mathrm{Y}$. Suppose that f is a (G, Y, S^1)-quasi-cocycle of order $< k$. If $\pi^* f$ is of type $< k$, then f is of type $< k$.*

We will prove this proposition in section 8.7. Assuming it for now, we conclude that $\tilde{f}$ is of type $< k$. Observe that X' is an extension of $\mathcal{Z}_{<j-1}(\mathrm{X}) = \mathcal{Z}_{<j-1}(\mathrm{X}')$ by a finite abelian group by a cocycle of degree $< j-1$. If Theorem 8.6 (for this choice of k and j) has already been established in the case when U_{j-1} is finite, then $\tilde{f}$ is $(\mathbb{F}^\omega, \mathrm{X}', S^1)$-cohomologous to a $(\mathbb{F}^\omega, \mathrm{X}', S^1)$-phase polynomial of degree $< k$. Pulling this back by π, we conclude that f'' is $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cohomologous to a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-phase polynomial of degree $< k$. Since f'' is $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cohomologous to f/Q, and Q is also a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-phase polynomial of degree $< k$, we conclude that f is $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cohomologous to a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-phase polynomial of degree $< k$, and Theorem 8.6 follows.

The remaining tasks (in the low-order case $j \leqslant k$) are to verify Theorem 8.6 in the case of finite U_{j-1}, and to also verify Proposition 8.9 and Proposition 8.11.

8.4 The finite group case.
We now establish Theorem 8.6 in the case when U_{j-1} is finite. This is the analogue of Proposition 7.1, but our arguments here are somewhat simpler thanks to the high characteristic (which allows us to use the full power of Lemma D.3).

Since U_{j-1} is finite and p-torsion, we can write $U_{j-1} = C_p^L$ for some finite L. We will now induct on the dimension L. The case $L = 0$ is trivial, so suppose inductively that $L \geqslant 1$ and that the claim has already been proven for $L - 1$. We write $U_{j-1} = C_p^{L-1} \times \langle e \rangle$, where e is a generator of C_p. Recall that $\boldsymbol{\Delta}_e f$ is $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cohomologous to a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-phase polynomial q_e of degree $< k-j+1$ which takes values in C_p. We now extend this from e to $\langle e \rangle$ in a manner which is a cocycle with respect to this parameter.

Fix q_e, and then define the $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-function q_{e^s} for all $0 \leqslant s < p$ by the formula $q_{e^s}(g, x) := \prod_{i=0}^{s-1} q_e(g, V_e^i x)$. Since q_e is a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-phase polynomial of degree $< k - j + 1$, q_{e^s} is also. Since $\blacktriangle_e f$ is $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cohomologous to q_e, we see from the cocycle identity

$$\blacktriangle_{e^s} f(g, x) = \prod_{i=0}^{s-1} \blacktriangle_e f(g, V_e^i x) \tag{8.1}$$

that $\blacktriangle_u f$ is $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cohomologous to q_u for all $u \in \langle e \rangle$.

By repeated application of Lemma 8.8, we know that q_e has degree $< p$ with respect to differentiation $\blacktriangle_e$ in the e direction. By Lemma D.3, we conclude that $\prod_{i=0}^{p-1} q_e(g, V_e^i x) = 1$, and thus the q_{e^s} form a cocycle in the e^s variable, in the sense that

$$q_{uv}(g, x) = q_u(g, x) q_v(g, V_u x) \tag{8.2}$$

for all $g \in \mathbb{F}^\omega$, $u, v \in \langle e \rangle$, and μ-a.e. x. Applying Proposition 8.9, we may find a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-phase polynomial Q of degree $< k$ such that $q_u = \blacktriangle_u Q$ for all $u \in \langle e \rangle$. Since q_e takes values in C_p, Q does also, and thus (by Lemma D.3) is a line cocycle.

By repeating the arguments in the previous section, we also see that Q is a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-quasi-cocycle of order $< k - 1$.

Recall that $\blacktriangle_e f$ is $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cohomologous to $q_e = \blacktriangle_e Q$, thus we can find an (X, S^1)-function F_e such that $\blacktriangle_e f = (\blacktriangle_e Q) \blacktriangle F_e$. Using the cocycle identity (8.1) for f and Q we conclude that $\blacktriangle \prod_{i=0}^{p-1} V_e^i F_e = 1$ and thus by ergodicity, $\prod_{i=0}^{p-1} V_e^i F_e$ is equal to some constant in S^1. Taking p^{th} roots, we can express this constant as c^p for some $c \in S^1$; dividing the F_e by this constant (which does not affect any of the properties of F_e), we may take $c = 1$, thus $\prod_{i=0}^{p-1} V_e^i F_e = 1$. If we then define $F_{e^s} := \prod_{i=0}^{s-1} V_e^i F_e$ for $0 \leqslant s < p$, then the F_{e^s} form a cocycle in the e^s variable in the sense of (8.2). Applying Lemma B.4, this cocycle is a coboundary, thus we can write $F_e = \blacktriangle_e F$ for some (X, S^1)-function F. We conclude that $\blacktriangle_e \frac{f}{Q \blacktriangle F} = 1$, thus f/Q is cohomologous to a function which is $\langle e \rangle$-invariant. We can now argue as in the preceding subsection (with $\langle e \rangle$ playing the role of U'') to deduce Theorem 8.6 for $U_{j-1} = C_p^L$ from the corresponding claim for $U_{j-1} = C_p^{L-1}$, which we have by induction. This establishes Theorem 8.6 in the low order case $j \leqslant k$, contingent on Propositions 8.9 and 8.11.

8.5 The high-order case.

We now modify the above arguments to deal with the high-order case $j > k$. We need a key definition: we say that a $\mathbf{2}^k$-tuple $(f_\mathbf{w})_{\mathbf{w} \in \mathbf{2}^k}$ of $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-functions $f_\mathbf{w}$ for $\mathbf{w} \in \mathbf{2}^k$ is a *good* $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-*tuple* if the following properties hold:

(i) Each $f_\mathbf{w}$ is a line cocycle and a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-quasi-cocycle of order $< k - 1$.
(ii) For each face α of $\mathbf{2}^k$, $\prod_{\mathbf{w} \in \alpha} f_\mathbf{w}$ is a $(\mathbb{F}^\omega, \mathrm{X}^{[k]}, S^1)$-cocycle.
(iii) $\prod_{\mathbf{w} \in \mathbf{2}^k} f_\mathbf{w}$ is a $(\mathbb{F}^\omega, \mathrm{X}^{[k]}, S^1)$-coboundary.

From the hypothesis we see that $(f^{\mathrm{sgn}(\mathbf{w})})_{\mathbf{w} \in \mathbf{2}^k}$ is a good $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-tuple. Observe also that if $(f_\mathbf{w})_{\mathbf{w} \in \mathbf{2}^k}$ is a good $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-tuple, and we replace one or more of the $f_\mathbf{w}$ by a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cohomologous function, then we still obtain a good $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-tuple.

We now claim the following proposition:

PROPOSITION 8.12 (Descent of good tuples).   *Let $1 \leqslant k < j \leqslant \mathrm{char}(\mathbb{F})$, and let* $\mathrm{X} = \mathcal{Z}_{<j-1}(\mathrm{X}) \times_{\rho_{j-1}} U_{j-1}$ *be an ergodic $\mathbb{F}^\omega$-system of order $< j$. Suppose that $(f_{\mathbf{w}})_{\mathbf{w} \in \mathbf{2}^k}$ is a good $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-tuple. Then there exists a good $(\mathbb{F}^\omega, \mathcal{Z}_{<j-1}(\mathrm{X}), S^1)$-tuple $(\tilde{f}_{\mathbf{w}})_{\mathbf{w} \in \mathbf{2}^k}$ such that for every $\mathbf{w}$, $f_{\mathbf{w}}$ is $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cohomologous to $(\pi^{\mathrm{X}}_{\mathcal{Z}_{<j-1}(\mathrm{X})})^* \tilde{f}_{\mathbf{w}}$, where $\pi^{\mathrm{X}}_{\mathcal{Z}_{<j-1}(\mathrm{X})} : \mathrm{X} \to \mathcal{Z}_{<j-1}(\mathrm{X})$ is the factor map.*

Proof. As in previous arguments, we first reduce to the case when U_{j-1} is finite, and then establish the finite case.

If $t \in U_{j-1}$, then by Lemma A.23(iv), the transformation $(V_t)^{[k]}_{\{\mathbf{w}\}}$ preserves $\mu^{[k]}$ for every $\mathbf{w} \in \mathbf{2}^k$. Arguing as in the proof of Lemma 8.8, we thus conclude that $\Delta_t f_{\mathbf{w}}$ is a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-coboundary for every $\mathbf{w} \in \mathbf{2}^k$. By repeatedly applying Lemma B.8, we conclude that there exists an open subgroup U'' of U_{j-1} such that each $f_{\mathbf{w}}$ is $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cohomologous to an U''-invariant function, which we can write as $\pi^* \tilde{f}_{\mathbf{w}}$ for some $(\mathbb{F}^\omega, \mathrm{X}', S^1)$-function $\tilde{f}_{\mathbf{w}}$, where $\mathrm{X}' := \mathcal{Z}_{<j}(\mathrm{X}) \times_{\sigma \circ \rho_{<j-1}} U_{j-1}/U''$, $\sigma : U_{j-1} \to U_{j-1}/U''$ is the quotient map, and $\pi : \mathrm{X} \to \mathrm{X}'$ is the factor map, thus we have $f_{\mathbf{w}} = (\pi^* \tilde{f}_{\mathbf{w}}) \Delta F_{\mathbf{w}}$ for some (X, S^1)-functions $\mathbb{F}_w$.

Arguing as in section 8.3 (and using Lemma B.1), we see that the tuple $(\tilde{f}_{\mathbf{w}})_{\mathbf{w} \in \mathbf{2}^k}$ obeys properties (i) and (ii) of being a good $(\mathbb{F}^\omega, \mathrm{X}', S^1)$-tuple. Unfortunately, it need not obey (iii); we know that $\bigotimes_{\mathbf{w}} \pi^* \tilde{f}_{\mathbf{w}}$ is a $(\mathbb{F}^\omega, \mathrm{X}^{[k]}, S^1)$-coboundary, but this does not automatically imply that $\bigotimes_{\mathbf{w}} \tilde{f}_{\mathbf{w}}$ is a $(\mathbb{F}^\omega, (\mathrm{X}')^{[k]}, S^1)$-coboundary. However, in the high-order case $j > k$, we see from Lemma A.36 that $\mathrm{X}^{[k]}$ is an abelian extension $\mathrm{X}^{[k]} = \mathcal{Z}_{<j-1}(\mathrm{X})^{[k]} \times_{\rho^{[k]}_{<j-1}} U^{[k]}_{j-1}$, where the $(\mathbb{F}^\omega, \mathcal{Z}_{<j-1}(\mathrm{X})^{[k]}, U^{[k]}_{j-1})$-cocycle $\rho^{[k]}_{<j-1} = \bigotimes_{\mathbf{w} \in \mathbf{2}^k} \rho_{<j-1}$ is the tensor product of 2^k copies of the $(\mathbb{F}^\omega, \mathcal{Z}_{<j-1}(\mathrm{X}), U_{j-1})$-cocycle ρ, and similarly for $(\mathrm{X}')^{[k]}$. Applying Lemma B.11, we conclude that $\bigotimes_{\mathbf{w} \in \mathbf{2}^k} \tilde{f}_{\mathbf{w}}$ is $(\mathbb{F}^\omega, (\mathrm{X}')^{[k]}, S^1)$-cohomologous to $\chi^{[k]} \circ \rho^{[k]} \circ (\pi^{\mathrm{X}'}_{\mathcal{Z}_{<j-1}(\mathrm{X})})^{[k]}$ for some character $\chi^{[k]} \in \widehat{U^{[k]}_{j-1}}$, where $\pi^{\mathrm{X}'}_{\mathcal{Z}_{<j-1}(\mathrm{X})} : \mathrm{X}' \to \mathcal{Z}_{<j-1}(\mathrm{X})$ is the factor map. We can factorize the latter as a tensor product

$$\chi^{[k]} \circ \rho^{[k]} \circ (\pi^{\mathrm{X}'}_{\mathcal{Z}_{<j-1}(\mathrm{X})})^{[k]} = \bigotimes_{\mathbf{w} \in \mathbf{2}^k} \chi_{\mathbf{w}} \circ \rho_{<j-1} \circ \pi^{\mathrm{X}'}_{\mathcal{Z}_{<j-1}(\mathrm{X})} =: \bigotimes_{\mathbf{w} \in \mathbf{2}^k} p_{\mathbf{w}}$$

for some characters $\chi_{\mathbf{w}} \in \hat{U}_{j-1}$. We thus see that $\bigotimes_{\mathbf{w}} f'_{\mathbf{w}}$ is a $(\mathbb{F}^\omega, (\mathrm{X}')^{[k]}, S^1)$-coboundary. Since $\rho_{<j-1} \circ \pi^{\mathrm{X}}_{\mathcal{Z}_{<j-1}(\mathrm{X})}$ is a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-coboundary (being the derivative of the coordinate function $(y, u) \mapsto u$), we see that $\pi^* p_{\mathbf{w}}$ is also a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-coboundary. Thus $f_{\mathbf{w}}$ is $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cohomologous to $\pi^* f'_{\mathbf{w}}$. Also, since $(\tilde{f}_{\mathbf{w}})_{\mathbf{w} \in \mathbf{2}^k}$ obeys properties (i), (ii), of being a $(\mathbb{F}^\omega, \mathrm{X}', S^1)$-good tuple, $(f'_{\mathbf{w}})_{\mathbf{w} \in \mathbf{2}^k}$ does also. Since we have already established (iii), we conclude that $(f'_{\mathbf{w}})_{\mathbf{w} \in \mathbf{2}^k}$ is a $(\mathbb{F}^\omega, \mathrm{X}', S^1)$-good tuple. Thus Proposition 8.12 for U_{j-1} will follow from that for U_{j-1}/U'', thus reducing matters to the case when U_{j-1} is finite.

We now repeat the arguments from section 8.4. As in that section, we can write $U_{j-1} = C_p^L = C_p^{L-1} \times \langle e \rangle$ for some $e \in C_p$ and induct on L. Arguing as in the start of this proof (i.e. using the invariance of $\mu^{[k]}$ with respect to $(V_e)^{[k]}_{\{\mathbf{w}\}}$), we

know that $\Delta_e f_{\mathbf{w}}$ is a $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-coboundary, thus we can find a (X, S^1)-function $F_{e,\mathbf{w}}$ such that $\Delta_e f_{\mathbf{w}} = \Delta F_{e,\mathbf{w}}$. Using the cocycle identity (8.1), this implies that $\Delta \prod_{i=0}^{p-1} V_e^i F_{e,\mathbf{w}} = 1$ and so (by ergodicity) $\prod_{i=0}^{p-1} V_e^i F_{e,\mathbf{w}}$ is constant. Dividing $F_{e,\mathbf{w}}$ by the pth root of this constant as in section 8.4 we may thus take $\prod_{i=0}^{p-1} V_e^i F_{e,\mathbf{w}} = 1$. Applying Lemma B.4 as in section 8.4, we can write $F_{e,\mathbf{w}} = \Delta_e F_{\mathbf{w}}$ for some (X, S^1)-function $F_{\mathbf{w}}$. We conclude that $\Delta_e(f_{\mathbf{w}}/\Delta F_{\mathbf{w}}) = 1$, i.e. $f_{\mathbf{w}}$ is cohomologous to an e-invariant function. From this we see (as in the reduction to the finite U_{j-1} case) that Proposition 8.12 for C_p^L will follow from Proposition 8.12 for C_p^{L-1}. Since this proposition is trivial when $L = 0$, the claim follows. $\qquad\square$

Applying this proposition iteratively, starting with the good $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-tuple $(f^{\mathrm{sgn}(\mathbf{w})})_{\mathbf{w} \in \mathbf{2}^k}$ and decrementing j, we see that there exists a good $(\mathbb{F}^\omega, \mathcal{Z}_{<k}(\mathrm{X}), S^1)$-tuple $(\tilde{f}_{\mathbf{w}})_{\mathbf{w} \in \mathbf{2}^k}$ such that $f^{\mathrm{sgn}(\mathbf{w})}$ is $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cohomologous to $\left(\pi_{\mathcal{Z}_{<k}(\mathrm{X})}^{\mathrm{X}}\right)^* \tilde{f}_{\mathbf{w}}$ for all $\mathbf{w} \in \mathbf{2}^k$.

We now invoke the following lemma:

LEMMA 8.13 (Cauchy–Schwarz–Gowers for finite type). *Let G be a countable abelian group, let $k \geqslant 0$, let X be an ergodic G-system of order $< k$, and for each $\mathbf{w} \in \mathbf{2}^k$, let $f_{\mathbf{w}}$ be a (G, X, S^1)-quasi-cocycle of order $< k$ such that $\bigotimes_{\mathbf{w}} f_{\mathbf{w}}$ is a $\left(G, \mathrm{X}^{[k]}, S^1\right)$-coboundary. Then each $f_{\mathbf{w}}$ is of type $< k$.*

We prove this lemma later in section 8.7. Assuming it for now, we conclude in particular that $\tilde{f}_{-\mathbf{1}}$ is of type $< k$. It is also a line cocycle and a $(\mathbb{F}^\omega, \mathcal{Z}_{<k}(\mathrm{X}), S^1)$-quasi-cocycle of order $< k - 1$. Since we have already established Theorem 8.6 in the case $j = k$, we conclude that $\tilde{f}_{-\mathbf{1}}$ is is $(\mathbb{F}^\omega, \mathcal{Z}_{<k}(\mathrm{X}), S^1)$-cohomologous to a $(\mathbb{F}^\omega, \mathcal{Z}_{<k}(\mathrm{X}), S^1)$-phase polynomial P of degree $< k$ that takes values in C_p. Since f is $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cohomologous to $\left(\pi_{\mathcal{Z}_{<k}(\mathrm{X})}^{\mathrm{X}}\right)^* \tilde{f}_{-\mathbf{1}}$, we conclude that f is $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-cohomologous to the $(\mathbb{F}^\omega, \mathrm{X}, S^1)$-phase polynomial $\left(\pi_{\mathcal{Z}_{<k}(\mathrm{X})}^{\mathrm{X}}\right)^* P$, and Theorem 8.6 then follows.

Our remaining tasks are to prove Proposition 8.9, Proposition 8.11, and Lemma 8.13.

8.6 Exact integration.

In this subsection we establish Proposition 8.9. We begin with the analogue of Lemma B.5.

LEMMA 8.14 (Refined composition of polynomials). *Let G be a countable abelian group, let $j \geqslant 0$, let U be a compact abelian group, and let $\mathrm{X} = \mathrm{Y} \times_\rho U$ be an ergodic G-system such that $\mathrm{Y} \geqslant \mathcal{Z}_{>j}(\mathrm{X})$. For any $t \in U$, let p_t be a (X, S^1)-phase polynomial of degree $< l$ for some $l \geqslant 1$, and suppose that for any $t, s \in U$, we have the cocycle property*

$$p_{t \cdot s}(x) = p_t(V_s x) \cdot p_s(x). \tag{8.3}$$

Let r be a (X, U)-phase polynomial of degree $< m$ for some $1 \leqslant m \leqslant j + 1$. For any $x \in X$, write $x = (y, u)$, $y \in \mathrm{Y}$, $u \in U$.

(i) *The function $(y, u) \mapsto p_{r(y,u)}(y, u)$ is a (X, S^1)-phase polynomial of degree $< l + m - 1$.*

(ii) *More generally, if q is a (X, U)-phase polynomial of degree $< j + 1$, then $(y, u) \mapsto p_{r(y,u)}(y, uq(y, u))$ is a (X, S^1)-phase polynomial of degree $< l + m - 1$.*

Proof. We prove (i) and (ii) simultaneously by induction on l. If $l = 1$, then p_t is a constant (by ergodicity), and so the map $t \mapsto p_t$ is a homomorphism, and the claim (i) then easily follows. Also in this case (ii) is clearly equivalent to (i).

Now suppose inductively that $l \geqslant 2$, and that the claim has already been proven for $l - 1$. We now induct on m. When $m = 1$ then r is constant, and (i) is clear.

Now we show (for any m) that (i) implies (ii). Suppose first that $l \leqslant j$. Then p_t is measurable with respect to $\mathrm{Abr}_{<l}(\mathrm{X}) \leqslant \mathrm{Abr}_{<j}(\mathrm{X})$, and hence measurable with respect to $\mathcal{Z}_{<j}(\mathrm{X})$ (and hence Y) by Lemma A.35. In particular, p_t is U-invariant, and (i) and (ii) are clearly equivalent. Now suppose instead that $l > j$. The action of U preserves $\mathcal{Z}_{<j}(\mathrm{X})$, and thus by Lemma A.23(iv), the measure $\mu^{[l]}$ is invariant under the face transformations $(V_u)^{[l]}_\alpha$ for any $(l - j)$-face α and any $u \in U$. As U is abelian, this implies that $\mu^{[l]}$ is invariant under any $(V_{u_\mathbf{w}})_{\mathbf{w} \in \mathbf{2}^l}$, where the $u_\mathbf{w}$ obey the conditions $\prod_{\mathbf{w} \in \alpha} u_\mathbf{w}^{\mathrm{sgn}(\mathbf{w})} = 1$ for every $(j + 1)$-face α. Since q has degree < 1, we conclude (by Lemma 4.3(iii)) that $\mu^{[l]}$ is invariant under $(V_{q(y_\mathbf{w}, u_\mathbf{w})})_{\mathbf{w} \in \mathbf{2}^l}$ for $\mu^{[l]}$-a.e. $(y_\mathbf{w}, u_\mathbf{w})_{\mathbf{w} \in \mathbf{2}^l}$. On the other hand, since p_t has degree $< l$ for every $t \in U$, we have $\prod_{\mathbf{w} \in \mathbf{2}^l} p_t(y_\mathbf{w}, u_\mathbf{w})^{\mathrm{sgn}(\mathbf{w})} = 1$ $\mu^{[l]}$-a.e. for such t by Lemma 4.3(iii). We conclude that $\prod_{\mathbf{w} \in \mathbf{2}^l} p_t(y_\mathbf{w}, u_\mathbf{w} q(y_\mathbf{w}, u_\mathbf{w}))^{\mathrm{sgn}(\mathbf{w})} = 1$ $\mu^{[l]}$-a.e. By one last application of Lemma 4.3(iii) we obtain that $(y, u) \mapsto p_t(y, uq(y, u))$ is of degree $< l$. The claim (ii) now follows from (i).

Finally, we assume that (i) (and hence (ii)) have been proven for $m - 1$, and then establish (i) for m. Let F be the (X, S^1)-function $F(x) := p_{r(x)}(x)$. From (8.3) we have

$$F(T_g x) = p_{r(x) \Delta_g r(x)}(T_g x) = p_{r(x)}(T_g x) p_{\Delta_g r(x)}(V_{r(x)} T_g x),$$

and thus (since the action of U commutes with that of G)

$$\Delta_g F(x) = \big(\Delta_g p_{r(x)}\big)(x)\big(T_g p_{\Delta_g r(x)}\big)\big(V_{r(x)} x\big).$$

Observe that for each $t \in U$, $\Delta_g p_t$ is a (X, S^1)-phase polynomial of degree $< l - 1$, and that $\Delta_g p_t$ obeys (8.3). By the induction hypothesis (i) with l replaced by $l - 1$, $(\Delta_g p_{r(x)})(x)$ is a (X, S^1)-phase polynomial of degree $< l + m - 2$. Similarly, $T_g p_t$ is a (X, S^1)-phase polynomial of degree $< l$ that also obeys (8.3), $\Delta_g r$ is a (X, U)-phase polynomial of degree $< m - 1$, and r has degree $< j + 1$ by hypothesis on m. Applying the induction hypothesis (ii) with m replaced by $m - 1$, we conclude that $(T_g p_{\Delta_g r(x)})(V_{r(x)} x)$ is a (X, S^1)-phase polynomial of degree $< l + m - 2$. Putting this all together, we see that $\Delta_g F$ has degree $< l + m - 2$ for all g, and thus (by (3.1)) F has degree $< l + m - 1$, establishing (i) as required. □

Now we can prove Proposition 8.9.

Proof of Proposition 8.9.   Write $x = (y, u)$, $y \in \mathrm{Y}$, $u \in U$. Take $Q(y, uu_0) := p_u(y, u_0)$ for $u_0 \in U$ a generic point (actually we can take $u_0 = 1$ since polynomials are continuous). Now as in Lemma B.6, $\Delta_t Q(x) = p_t(x)$. It remains to show that Q has degree $< l + j$. But this follows from Lemma 8.14(ii) with $r(y, u) := u/u_0$, $q(y, u) := u_0/u$, and $m := j + 1$; note that as ρ_j has degree $< j$, r and q have degree $< j + 1$. □

8.7   Exact descent.   We now prove Proposition 8.11 and Lemma 8.13. As noted already in Remark 5.2, an exact descent result for cocycles already appears in [HK, Cor. 7.8]; Proposition 8.11 can be viewed as an extension of that result to quasi-cocycles.

Our main tool for both of these tasks is the following equivalent characterization of the finite type condition.

LEMMA 8.15 (Characterization of finite type).   *Let G be a countable abelian group, let $k \geqslant 0$, let X be an ergodic G-system of order $< k$, and let f be a (G, X, S^1)-quasi-cocycle of order $< k - 1$. Let Φ_n be a Følner sequence for G. Then the following two statements are equivalent:*

(i)   *f is of type $< k$ (i.e. $d^{[k]} f$ is a $(G, X^{[k]}, S^1)$-coboundary).*
(ii)   *For all $\mathbf{x}$ in a set of positive $\mu^{[k]}$-measure, $\limsup_{n \to \infty} \left| \mathbb{E}_{g \in \Phi_n} d^{[k]} f(g, \mathbf{x}) \right| \neq 0$.*

Proof. We first show that (i) implies (ii). By hypothesis, we have $d^{[k]} f(g, \mathbf{x}) = F(T_g^{[k]} \mathbf{x}) \overline{F(\mathbf{x})}$ for all $g \in G$, $\mu^{[k]}$-a.e. $\mathbf{x}$, and some $(X^{[k]}, S^1)$-function F. We rearrange this as

$$f(g, x_{-\mathbf{1}}) = F\big((T_g x_{\mathbf{w}})_{\mathbf{w} \in 2^k}\big) \overline{F\big((x_{\mathbf{w}})_{\mathbf{w} \in 2^k}\big)} \prod_{\mathbf{w} \in 2^k \setminus -\mathbf{1}} f_{\mathbf{w}}(g, x_{\mathbf{w}})$$

where each $f_{\mathbf{w}}$ is either equal to f or its complex conjugate.

As X is of order $< k$, we see from Definition A.29 that any measurable function on $X^{[k]}$ that depends only on the first coordinate $x_{-\mathbf{1}}$, is equal $\mu^{[k]}$-a.e. to a function that is independent of this coordinate. Since these two classes of functions together generate all measurable functions on $X^{[k]}$, we conclude that all measurable functions on $X^{[k]}$ are equal $\mu^{[k]}$-a.e. to a function independent of the first coordinate. In particular, we may assume without loss of generality that F is independent of the first coordinate.

Let $\varepsilon > 0$ be a small number. By definition of the product σ-algebra, we can approximate F up to an error which is $O(\varepsilon)$ in $L^2(X^{[k]})$ by a function $\tilde{F}$ bounded in magnitude by 1 of the form $\tilde{F}((x_{\mathbf{w}})_{\mathbf{w} \in 2^k}) = \sum_{j=1}^{N} \prod_{\mathbf{w} \in 2^k \setminus -\mathbf{1}} F_{j, \mathbf{w}}(x_{\mathbf{w}})$ for some functions $f_{j, \mathbf{w}} \in L^\infty(X)$ with norm $\|f_{j, \mathbf{w}}\|_{L^\infty(X)} \leqslant 1$ and some finite N. We conclude that $f(g, x_{-\mathbf{1}})$ differs by $O(\varepsilon)$ in $L^2(X^{[k]})$ from the function

$$\sum_{j=1}^{N} \sum_{j'=1}^{N} \prod_{\mathbf{w} \in 2^k \setminus -\mathbf{1}} F_{j, \mathbf{w}}(T_g x_{\mathbf{w}}) \overline{F_{j', \mathbf{w}}(x_{\mathbf{w}})} f_{\mathbf{w}}(g, x_{\mathbf{w}}) \, .$$

On the other hand, f has magnitude 1. We conclude (for ε small enough) that

$$\left| \sum_{j=1}^{N} \sum_{j'=1}^{N} \int_{X^{[k]}} f(g, x_{-\mathbf{1}}) \prod_{\mathbf{w} \in 2^k \setminus -\mathbf{1}} \overline{F_{j, \mathbf{w}}(T_g x_{\mathbf{w}})} F_{j', \mathbf{w}}(x_{\mathbf{w}}) \overline{f_{\mathbf{w}}(g, x_{\mathbf{w}})} \, d\mu^{[k]} \right| \geqslant \frac{1}{2}$$

and thus by the pigeonhole principle that

$$\left| \int_{X^{[k]}} f(g, x_{-\mathbf{1}}) \prod_{\mathbf{w} \in 2^k \setminus -\mathbf{1}} \overline{F_{j, \mathbf{w}}(T_g x_{\mathbf{w}})} F_{j', \mathbf{w}}(x_{\mathbf{w}}) \overline{f_{\mathbf{w}}(g, x_{\mathbf{w}})} \, d\mu^{[k]} \right| \geqslant \frac{1}{2N^2}$$

for some j, j' (depending on g). Applying the Cauchy–Schwarz–Gowers inequality (A.4) we conclude that $\|f(g, \cdot)\|_{U^k(\mathrm{X})} \geqslant 1/2N^2$ for all $g \in G$; the point here is that the lower bound is uniform in g. Applying Lemma A.18, we conclude that

$$\int_{\mathrm{X}^{[k]}} d^{[k]} f(g, \mathbf{x}) \, d\mu^{[k]} \geqslant \left(\frac{1}{2N^2}\right)^{2^k}$$

for all $g \in G$. Averaging this over a Følner set Φ_n, we conclude in particular that

$$\int_{\mathrm{X}^{[k]}} \left|\mathbb{E}_{g \in \Phi_n} d^{[k]} f(g, \mathbf{x})\right| d\mu^{[k]} \geqslant \left(\frac{1}{2N^2}\right)^{2^k}$$

for all n. From the monotone convergence theorem we conclude that

$$\int_{\mathrm{X}^{[k]}} \limsup_{n \to \infty} \left|\mathbb{E}_{g \in \Phi_n} d^{[k]} f(g, \mathbf{x})\right| d\mu^{[k]} \geqslant \left(\frac{1}{2N^2}\right)^{2^k}$$

and (ii) follows.

Now we show that (ii) implies (i). We will use some arguments related to those used to prove Proposition 4.4. We first observe from Definition 8.2 and Lemma 4.3(iii) that $d^{[k]} f$ is a $\left(G, \mathrm{X}^{[k]}, S^1\right)$-cocycle. This allows us to build the circle extension $\mathrm{X}^{[k]} \times_{d^{[k]} f} S^1$ of $\mathrm{X}^{[k]}$. Applying the ergodic theorem to the vertical function $(\mathbf{x}, u) \mapsto u$ in this extension, we conclude that the limit $F(\mathbf{x}) := \lim_{n \to \infty} \mathbb{E}_{g \in \Phi_n} d^{[k]} f(g, \mathbf{x})$ exists $\mu^{[k]}$-a.e., and is invariant under the diagonal action of G. By (ii), F is non-zero on a set A of positive measure in $\mathrm{X}^{[k]}$, which we can take to be invariant under the action of G. In particular, A corresponds to some set $B \in \mathcal{I}_k(\mathrm{X})$ with $P_k(B) > 0$, where P_k is the restriction of $\mu^{[k]}$ to $\mathcal{I}_k(\mathrm{X})$.

Since $d^{[k]} f$ is a $\left(G, \mathrm{X}^{[k]}, S^1\right)$-cocycle, we have $d^{[k]} f(g + g', \mathbf{x}) = \left(d^{[k]} f(g, \mathbf{x})\right) d^{[k]} f\left(g', T_g^{[k]} \mathbf{x}\right)$. Averaging g' over the Følner sequence Φ_n and taking limits, we conclude that $F(\mathbf{x}) = \left(d^{[k]} f(g, \mathbf{x})\right) F\left(T_g^{[k]} \mathbf{x}\right)$. This implies that $d^{[k]} f$ is a $\left(G, B, S^1\right)$-coboundary.

Now let α be any face of $\mathbf{2}^k$. From Definition 8.2 and Lemma 4.3(iii) we see that $d^{[k-1]} f$ is also a $\left(G, \mathrm{X}^{[k-1]}, S^1\right)$-cocycle, and thus $(\partial(\alpha)_*)^* d^{[k-1]} f$ is a $\left(G, \mathrm{X}^{[k]}, S^1\right)$-cocycle. This implies that

$$\frac{(T_h)_\alpha^{[k]} d^{[k]} f(g, \mathbf{x})}{d^{[k]} f(g, \mathbf{x})} = \boldsymbol{\Delta}_{g^{[k]}} \left(\partial(\alpha)_*\right)^* d^{[k-1]} f(h, \mathbf{x})$$

for every $h \in G$, and so $(T_h)_\alpha^{[k]} d^{[k]} f$ is $\left(G, \mathrm{X}^{[k]}, S^1\right)$-cohomologous to $d^{[k]} f$. In particular, $d^{[k]} f$ is a $\left(G, (T_h)_\alpha^{[k]} B, S^1\right)$-coboundary. Using the same gluing argument used in the proof of Proposition 4.4 we conclude that $d^{[k]} f$ is a $\left(G, \mathrm{X}^{[k]}, S^1\right)$-coboundary, and the claim follows. $\qquad \square$

We can now prove Proposition 8.11 and Lemma 8.13.

Proof of Proposition 8.11. By hypothesis and Lemma 8.15, we have

$$\limsup_{n \to \infty} \left|\mathbb{E}_{g \in \Phi_n} d^{[k]} \pi^* f(g, \mathbf{x})\right| \neq 0$$

for all $\mathbf{x}$ in a set of positive measure in $\mathrm{X}^{[k]}$.

By Lemma A.22, $Y^{[k]}$ is a factor of $X^{[k]}$. Let $\pi^{[k]}$ be the factor map, then we have $d^{[k]}\pi^{*}f = \left(\pi^{[k]}\right)^{*}d^{[k]}f$. We conclude that $\limsup_{n\to\infty}\left|\mathbb{E}_{g\in\Phi_n}d^{[k]}f(g,\mathbf{y})\right| \neq 0$ for all $\mathbf{y}$ in a set of positive measure in $Y^{[k]}$.

Since X is of order $< k$, Y is also. The claim then follows from another application of Lemma 8.15. $\qquad\square$

Proof of Lemma 8.13. By hypothesis, we have $\bigotimes f_{\mathbf{w}}(g,\mathbf{x}) = F\left(T_g^{[k]}\mathbf{x}\right)\overline{F(\mathbf{x})}$ for some $\left(X^{[k]},S^1\right)$-function F. Arguing exactly as in the proof of Lemma 8.15, we see that for all $\mathbf{x}$ in a set of positive $\mu^{[k]}$-measure, $\limsup_{n\to\infty}\left|\mathbb{E}_{g\in\Phi_n}d^{[k]}f_{-\mathbf{1}}(g,\mathbf{x})\right| \neq 0$. Applying Lemma 8.15, we conclude that $f_{-\mathbf{1}}$ is of type $< k$ as desired. The corresponding claims for the other $f_{\mathbf{w}}$ are established similarly. $\qquad\square$

The proof of Theorem 1.19 is now complete.

Appendices

A   General Theory of Gowers–Host–Kra Seminorms

In this appendix we set out the general "soft" theory of the Gowers–Host–Kra seminorms. The theory for $\mathbb{Z}$-systems is discussed in detail in [HK]. This theory carries over without any difficulties to other discrete abelian groups, such as $\mathbb{F}^{\omega}$, but for the convenience of the reader we reproduce the theory from [HK] here.

Throughout this appendix, $(G,+)$ is a fixed countable abelian group, the system $X = (X, \mathcal{B}_X, \mu_X, (T_g)_{g\in G})$ is a fixed G-system, and $k \geqslant 1$ is a fixed integer.

A.1   The cubic construction.   The first step is to construct the cubic spaces and measures from [HK], generalized from $\mathbb{Z}$-actions to G-actions. We slightly modify the terminology and notation of [HK] to make it compatible with standard notation of the theory of cubic complexes (see section E).

DEFINITION A.1 (The discrete cube and its faces).   Let $\mathbf{2}^k$ denote the set $\{-1,1\}^k$, which we equip with the lexicographic order. One can identify $\mathbf{2}^k$ with the vertices of the standard cube $\mathbf{I}^k := \{(x_1,\ldots,x_k)\in\mathbb{R}^k : -1 \leqslant x_i \leqslant 1\}$. Under this identification we will refer to $\mathbf{2}^k$ as the standard *discrete* (k-dimensional) cube. We will denote the elements of $\mathbf{2}^k$ by $\mathbf{w} = (w_1,\ldots,w_k)$. For each $j = 1,\ldots,k$, the sets $\beta_j^+ := \{\mathbf{w}\in\mathbf{2}^k : w_j = 1\}$, $\beta_j^- := \{\mathbf{w}\in\mathbf{2}^k : w_j = -1\}$ correspond to the sets of vertices on opposite sides of the cube $\mathbf{I}^k$. We will refer to β_j^+, β_j^- as *opposite sides* or *parallel sides* of $\mathbf{2}^k$, and to the β_j^+ as the *positive sides*.

More generally, for any $0 \leqslant l \leqslant k$, define an *l-dimensional face* or *l-face* to be any set formed by intersecting $k - l$ distinct non-parallel sides. Thus $\mathbf{2}^k$ has one k-face, $2k$ faces of dimension $k-1$ (i.e. the sides $\beta_j^{\pm}$), $k2^{k-l}$ faces of dimension $(k-l)$, and so forth down to 2^k faces of dimension 0 (which are the vertices of the discrete cube).

Let α be an l-face. Enumerating the elements of α in lexicographic order gives a natural bijection $\partial(\alpha) : \alpha \to \mathbf{2}^l$, which we call the *coordinate map* of α, which maps the faces of $\mathbf{2}^k$, which are subsets of α, to the faces of $\mathbf{2}^l$.

DEFINITION A.2 (Cubic complexes).   Let S be an arbitrary set. We write $S^{[k]} :=$ $S^{\mathbf{2}^k}$ for the set of functions $\mathbf{s} : \mathbf{2}^k \to S$. For each $0 \leqslant l \leqslant k$ and each l-face α, we let $\partial(\alpha)_* : S^{[k]} \to S^{[l]}$ be the pushforward map given by the formula $\partial(\alpha)_*(\mathbf{s})(\mathbf{w}) :=$ $\mathbf{s}(\partial(\alpha)^{-1}(\mathbf{w}))$ for all $\mathbf{w} \in \mathbf{2}^l$. For any $1 \leqslant j \leqslant k$ and sign $\pm$, we abbreviate the *cubic boundary map* $\partial(\beta_j^\pm)_* : S^{[k]} \to S^{[k-1]}$ as $\partial_j^\pm = \partial_{j,k}^\pm$.

EXAMPLE A.3. If $k = 2$, then

$$\partial_1^- \big(s_{(-1,-1)}, s_{(-1,1)}, s_{(1,-1)}, s_{(1,1)} \big) = \big(s_{(-1,-1)}, s_{(-1,1)} \big)$$
$$\partial_1^+ \big(s_{(-1,-1)}, s_{(-1,-1)}, s_{(1,-1)}, s_{(1,1)} \big) = \big(s_{(1,-1)}, s_{(1,1)} \big)$$
$$\partial_2^- \big(s_{(-1,-1)}, s_{(-1,1)}, s_{(1,-1)}, s_{(1,1)} \big) = \big(s_{(-1,-1)}, s_{(1,-1)} \big)$$
$$\partial_2^+ \big(s_{(-1,-1)}, s_{(-1,1)}, s_{(1,-1)}, s_{(1,1)} \big) = \big(s_{(-1,1)}, s_{(1,1)} \big).$$

REMARK A.4.   For future reference we make the trivial observation that the map $\mathbf{s} \mapsto (\partial_{k+1}^- \mathbf{s}, \partial_{k+1}^+ \mathbf{s})$ is a bijection between $S^{[k+1]}$ and $S^{[k]} \times S^{[k]}$.

DEFINITION A.5 (Face groups).   Let G be a (possibly non-abelian) group with identity id_G, and let α be a face of $\mathbf{2}^k$. For every $g \in G$, we let $g_\alpha^{[k]} \in G^{[k]}$ denote the group element whose components $\big(g_\alpha^{[k]} \big)_\mathbf{w}$ for $\mathbf{w} \in \mathbf{2}^k$ are defined to equal g when $\mathbf{w} \in \alpha$, and to equal id_G otherwise. The map $g \mapsto g_\alpha^{[k]}$ is a bijection from G to the *face group* $G_\alpha^{[k]} := \big\{ g_\alpha^{[k]} : g \in G \big\} \leqslant G^{[k]}$. When α is a side (resp. a positive side) we refer to $G_\alpha^{[k]}$ as a *side group* (resp. a *positive side group*); when $\alpha = \mathbf{2}^k$ is the entire cube we refer to $G_{\mathbf{2}^k}^{[k]}$ as the *diagonal group* and denote it as $\mathrm{diag}\big(G^{[k]}\big)$, and abbreviate $g_{\mathbf{2}^k}^{[k]}$ as $g^{[k]}$. We also write $\partial^{[k]}G$ (resp. $\partial_+^{[k]}$) for the subgroup of $G^{[k]}$ generated by all the side groups (resp. all the positive side groups).

EXAMPLE A.6. For $k = 2$, the group $\partial^{[2]}G$ is generated by

$$\bigcup_{g \in G} \big\{ (\mathrm{id}_G, \mathrm{id}_G, g, g), (g, g, \mathrm{id}_G, \mathrm{id}_G), (\mathrm{id}_G, g, \mathrm{id}_G, g), (g, \mathrm{id}_G, g, \mathrm{id}_G) \big\}_{g \in G}$$

while the group $\partial_+^{[2]}G$ is generated by $\bigcup_{g \in G} \{ (\mathrm{id}_G, \mathrm{id}_G, g, g), (\mathrm{id}_G, g, \mathrm{id}_G, g) \}_{g \in G}$.

REMARK A.7.   For future reference we observe that the side group $\partial^{[k]}G$ is the group generated by the positive side group $\partial_+^{[k]}G$ and the diagonal group $\mathrm{diag}\big(G^{[k]}\big)$.

DEFINITION A.8 (Face actions).   Let G be a group acting on a space X by transformations $T_g : X \to X$ for $g \in G$. Then $G^{[k]}$ acts on $X^{[k]}$ in the obvious manner, with the action $T_\mathbf{g}^{[k]}$ of a group element $\mathbf{g} = (g_\mathbf{w})_{\mathbf{w} \in \mathbf{2}^k} \in G^{[k]}$ mapping each point $(x_\mathbf{w})_{\mathbf{w} \in \mathbf{2}^k} \in X^{[k]}$ to $(T_{g_\mathbf{w}}(x_\mathbf{w}))_{\mathbf{w} \in \mathbf{2}^k}$. If α is a face, we abbreviate the *face transformation* $T_{g_\alpha^{[k]}}^{[k]}$ as $(T_g)_\alpha^{[k]}$, thus $\big((T_g)_\alpha^{[k]} \big)_{g \in G}$ is an action of G on $X^{[k]}$. If α is a side (resp. a positive side), we refer to $(T_g)_\alpha^{[k]}$ as a *side transformation* (resp. *positive side transformation*), and if $\alpha = \mathbf{2}^k$ is the entire cube, we refer to $(T_g)_{\mathbf{2}^k}^{[k]}$ as a *diagonal transformation* and abbreviate it further as $(T_g)^{[k]}$.

EXAMPLE A.9. If $k = 2$, then

$$(T_g)^{[2]}_{\beta_1^-}\big(x_{(-1,-1)}, x_{(-1,1)}, x_{(1,-1)}, x_{(1,1)}\big) = \big(T_g x_{(-1,-1)}, T_g x_{(-1,1)}, x_{(1,-1)}, x_{(1,1)}\big)$$

and similarly

$$(T_g)^{[2]}_{\beta_2^+}\big(x_{(-1,-1)}, x_{(-1,1)}, x_{(1,-1)}, x_{(1,1)}\big) = \big(x_{(-1,-1)}, T_g x_{(-1,1)}, x_{(1,-1)}, T_g x_{(1,1)}\big).$$

Finally,

$$(T_g)^{[2]}\big(x_{(-1,-1)}, x_{(-1,1)}, x_{(1,-1)}, x_{(1,1)}\big) = \big(T_g x_{(-1,-1)}, T_g x_{(-1,1)}, T_g x_{(1,-1)}, T_g x_{(1,1)}\big).$$

REMARK A.10.   For future reference we observe that all positive side maps preserve the $(-1, \ldots, -1)$ coordinate of $X^{[k]}$.

We recall the notion of a relative product:

DEFINITION A.11. Let G be a countable group, and let X_1, X_2 be two G-systems with a common factor Y. Let $\pi_1 : X_1 \to Y$ and $\pi_2 : X_2 \to Y$ be the factor maps. For $i = 1, 2$, let $\mu_{X_i,y}$ represent the disintegration of μ_{X_i} with respect to Y, thus

$$\int_{X_i} (\pi_i^* f) F \, d\mu_{X_i} = \int_{Y_i} f(y) \left(\int_{X_i} F \, d\mu_{X_i,y} \right) d\mu_Y(y)$$

for all $f \in L^\infty(Y)$, $F \in L^\infty(X_i)$. It is well known that this disintegration exists and is unique up to almost everywhere equivalence.

Let $\mu_{X_1} \times_Y \mu_{X_2}$ denote the measure defined by the formula

$$\mu_{X_1} \times_Y \mu_{X_2}(A) := \int \mu_{X_1,y} \times_Y \mu_{X_2,y}(A) d\mu_Y \tag{A.1}$$

for $A \in \mathcal{B}_{X_1} \times \mathcal{B}_{X_2}$; this is known as the *relatively independent joining* of μ_{X_1} and μ_{X_2} over μ_Y. We refer to the G-system

$$X_1 \times_Y X_2 := \big(X_1 \times X_2, \mathcal{B}_{X_1} \times \mathcal{B}_{X_2}, \mu_{X_1} \times_Y \mu_{X_2}, \mathrm{diag}(G \times G)\big)$$

where $\mathrm{diag}(G \times G)$ denotes the diagonal action of G, as the *relative product* of X_1 and X_2 with respect to Y.

Now we can introduce the cubic measure spaces of Host and Kra.

DEFINITION A.12 (Cubic measure spaces [HK, §3]).   Let $X = (X, \mathcal{B}_X, \mu_X, (T_g)_{g \in G})$ be a G-system. For each $k \geqslant 0$, we endow $X^{[k]}$ with the product σ-algebra $\mathcal{B}^{[k]} := (\mathcal{B}_X)^{\mathbf{2}^k}$, and define the cubic measures $\mu^{[k]}$ and σ-algebras $\mathcal{I}_k \subset \mathcal{B}^{[k]}$ inductively as follows:

- Set $\mathcal{I}_0$ to be the σ-algebra of invariant sets in X, and $\mu^{[0]}$ to be μ_X.
- Once $\mu^{[k]}$ and $\mathcal{I}_k$ are defined, we identify $X^{[k+1]}$ with $X^{[k]} \times X^{[k]}$ as per Remark A.4, and define $\mu^{[k+1]} := \mu^{[k]} \times_{\mathcal{I}_k} \mu^{[k]}$ to be the relatively independent joining of $\mu^{[k]}$ with itself over $\mathcal{I}_k$. We then let $\mathcal{I}_{k+1} = \mathcal{I}_{k+1}(X)$ be the σ-algebra of invariant sets of the system $X^{[k+1]}$.

We then define the G-system $X^{[k]}$ as $X^{[k]} := (X^{[k]}, \mathcal{B}^{[k]}, \mu^{[k]}, (T_g^{[k]})_{g \in G})$, i.e $\mathcal{B}_{X^{[k]}} = \mathcal{B}^{[k]}$ and $\mu_{X^{[k]}} = \mu^{[k]}$.

REMARK A.13. From construction we see that

$$\mu^{[k+1]} = \int_{X^{[k]}} \left(\mu^{[k]}\right)_s \times \left(\mu^{[k]}\right)_s dP_k(s) \tag{A.2}$$

where

$$\mu^{[k]} = \int \left(\mu^{[k]}\right)_s dP_k(s) \tag{A.3}$$

is the ergodic decomposition of $\mu^{[k]}$ with respect to the diagonal action $\left(T_g^{[k]}\right)_{g \in G}$ of G.

The cubic measures have a useful symmetry property:

LEMMA A.14 (Symmetry of cubic measures [HK, Prop. 3.7]). *The measure $\mu^{[k]}$ is invariant under all the symmetries of the cube $\mathbf{2}^k$ (which act in an obvious manner on $X^{[k]}$).*

The cubic measures also behave well with respect to passage to subcubes.

LEMMA A.15 (Cubic complex structure [HK, Cor. 3.8]). *Let $0 \leqslant l \leqslant k$, and let α be an l-face of $\mathbf{2}^k$. Then the map $\partial(\alpha)_* : X^{[k]} \to X^{[l]}$ is a factor map from $X^{[k]}$ to $X^{[l]}$.*

EXAMPLE A.16. Let $X = (X, \mathcal{B}_X, \mu_X, (T_g)_{g \in G})$ be an ergodic G-system. For $k = 1$ we have $\mu^{[1]} = \mu_X \times \mu_X$, $X^{[1]} = (X \times X, \mathcal{B}_X \times \mathcal{B}_X, \mu_X \times \mu_X, (T_g \times T_g)_{g \in G})$, and $\mathcal{I}_1$ is the σ-algebra of measurable subsets of $X \times X$ that are invariant under the action of $T_g^{[1]} = T_g \times T_g$, for all $g \in G$. Now if $\mu_X \times \mu_X = \mu^{[1]} = \int \left(\mu^{[1]}\right)_s dP_1(s)$ is the ergodic decomposition of $\mu_X \times \mu_X$ with respect to the action of $\left(T_g^{[1]}\right)_{g \in G}$, then $\mu^{[2]} = \int \left(\mu^{[1]}\right)_s \times \left(\mu^{[1]}\right)_s dP_1(s)$.

EXAMPLE A.17. Let G be a finite abelian group, and let X be G with the translation action, normalized counting measure, and the discrete σ-algebra. Then $\mu^{[k]}$ is the normalized counting measure on the space of cubes

$$\left\{(x + h_1 w_1 + \cdots + h_k w_k)_{(w_1, \ldots, w_k) \in \mathbf{2}^k} : x, h_1, \ldots, h_k \in G\right\}$$

and $\mathcal{I}_k$ is the σ-algebra consisting of subsets of $G^{[k]}$ which are invariant under the diagonal translations $(x_{\mathbf{w}})_{\mathbf{w} \in \mathbf{2}^k} \mapsto (x_{\mathbf{w}} + h)_{\mathbf{w} \in \mathbf{2}^k}$ for $h \in G$. Thus, the probability space $\left(X^{[k]}, \mathcal{B}^{[k]}, \mu^{[k]}\right)$ is measure isomorphic to the space $G^{k+1} = \{(x, h_1, \ldots, h_k) : x, h_1, \ldots, h_k \in G\}$ with the discrete σ-algebra and normalized counting measure, whilst $\left(X^{[k]}, \mathcal{I}_k, \mu^{[k]}\right)$ is measure isomorphic to the space $G^k = \{(h_1, \ldots, h_k) : h_1, \ldots, h_k \in G\}$ with the discrete σ-algebra and normalized counting measure.

A.2 Existence of the seminorms. The objective of this section is to establish that the Gowers–Host–Kra seminorms from Definition 1.3 are in fact well-defined, and to relate them to the cubic measures just constructed.

For $\mathbf{w} = (w_1, \ldots, w_k) \in \mathbf{2}^k$ denote $\mathrm{sgn}(\mathbf{w}) := \prod_{i=1}^k w_i \in \{-1, 1\}$. For any functions $f_{\mathbf{w}} : X \to \mathbb{C}$, $\mathbf{w} \in \mathbf{2}^k$ we denote by $\bigotimes_{\mathbf{w} \in \mathbf{2}^k} f_{\mathbf{w}} : X^{[k]} \to \mathbb{C}$ the tensor product

$$\bigotimes_{\mathbf{w} \in \mathbf{2}^k} f_{\mathbf{w}}\left((x_{\mathbf{w}})_{\mathbf{w} \in \mathbf{2}^k}\right) := \prod_{\mathbf{w} \in 2^{\mathbf{w}}} f_{\mathbf{w}}(x_{\mathbf{w}}).$$

LEMMA A.18. *Let $f \in L^\infty(X)$. The limits in Definition 1.3 exist and do not depend on the choice of Følner sequences. Furthermore, if $f_{\mathbf{w}} := f$ when $\mathrm{sgn}(\mathbf{w}) = 1$, and $f_{\mathbf{w}} := \bar{f}$ when $\mathrm{sgn}(\mathbf{w}) = -1$ ($\bar{f}$ denotes the complex conjugate of f), then*

$$\|f\|_{U^k(X)}^{2^k} = \big(\pi_{\mathrm{pt}}^{X^{[k]}}\big)_* \Big(\bigotimes_{\mathbf{w} \in \mathbf{2}^k} f_{\mathbf{w}} \Big) = \int_{X^{[k]}} \bigotimes_{\mathbf{w} \in \mathbf{2}^k} f_{\mathbf{w}} \, d\mu^{[k]}.$$

Proof. We prove this by induction on k. For $k = 1$ this follows from the mean ergodic theorem. Assume the induction hypothesis holds for $k - 1$. Then

$$\mathbb{E}_{h \in \Phi_n^k} \|\boldsymbol{\Delta}_h f\|_{U^{k-1}(X)}^{2^{k-1}} = \pi_{pt*}^{X^{[k-1]}} \Big(\bigotimes_{\mathbf{w} \in \mathbf{2}^{k-1}} f_{\mathbf{w}} \cdot T_h \bar{f}_{\mathbf{w}} \Big)$$

$$= \mathbb{E}_{h \in \Phi_n^k} \int \Big(\bigotimes_{\mathbf{w} \in \mathbf{2}^{k-1}} f_{\mathbf{w}} \Big) \cdot \Big(\bigotimes_{\mathbf{w} \in \mathbf{2}^{k-1}} T_h \bar{f}_{\mathbf{w}} \Big) d\mu^{[k-1]}$$

$$= \mathbb{E}_{h \in \Phi_n^k} \int \Big(\bigotimes_{\mathbf{w} \in \mathbf{2}^{k-1}} f_{\mathbf{w}} \Big) \cdot T_h^{[k]} \Big(\bigotimes_{\mathbf{w} \in \mathbf{2}^{k-1}} \bar{f}_{\mathbf{w}} \Big) d\mu^{[k-1]}.$$

Since $\mu^{[k-1]}$ is invariant with respect to the action of $\big(T_g^{[k]}\big)_{g \in G}$, by the ergodic theorem the above averages converge to

$$\int \pi_{\mathcal{I}_{k-1}*}^{X^{[k-1]}} \Big(\bigotimes_{\mathbf{w} \in \mathbf{2}^{k-1}} f_{\mathbf{w}} \Big) \cdot \pi_{\mathcal{I}_{k-1}*}^{X^{[k-1]}} \Big(\bigotimes_{\mathbf{w} \in \mathbf{2}^{k-1}} \bar{f}_{\mathbf{w}} \Big) d\mu^{[k-1]}$$

$$= \int \Big(\bigotimes_{\mathbf{w} \in \mathbf{2}^{k-1}} f_{\mathbf{w}} \Big) \Big(\bigotimes_{\mathbf{w} \in \mathbf{2}^{k-1}} \bar{f}_{\mathbf{w}} \Big) d\mu^{[k]}$$

$$= \Big(\bigotimes_{\mathbf{w} \in \mathbf{2}^k} f_{\mathbf{w}} \Big) d\mu^{[k]} = \pi_{pt*}^{X^{[k]}} \Big(\bigotimes_{\mathbf{w} \in \mathbf{2}^k} f_{\mathbf{w}} \Big)$$

thus closing the induction. $\square$

REMARK A.19. The above expression is used in [HK] as the *definition* of the $U^k(X)$ norm.

We record some basic properties of the Gowers–Host–Kra seminorms:

LEMMA A.20 (Basic properties of U^k [HK, Lem. 3.9] (see also [G1, Lem. 3.8,3.9])).
 (i) *For any $\mathbf{2}^k$-tuple $(f_{\mathbf{w}})_{\mathbf{w} \in \mathbf{2}^k}$ of functions $f_{\mathbf{w}} \in L^\infty(X)$, we have the Cauchy–Schwarz–Gowers inequality*

$$\Big| \int_{X^{[k]}} \bigotimes_{\mathbf{w} \in \mathbf{2}^k} f_{\mathbf{w}} \, d\mu^{[k]} \Big| \leqslant \prod_{\mathbf{w} \in \mathbf{2}^k} \|f_{\mathbf{w}}\|_{U^k(X)}. \tag{A.4}$$

 (ii) *The function $f \mapsto \|f\|_{U^k(X)}$ is a seminorm on $L^\infty(X)$.*
 (iii) *We have the monotonicity property*

$$\|f\|_{U^k(X)} \leqslant \|f\|_{U^{k+1}(X)}. \tag{A.5}$$

COROLLARY A.21. *Let $f \in L^\infty(X)$. Then $\|f\|_{U_k(X)} = 0$ if and only if $\big(\pi_{\mathcal{I}_{k-1}}^{X^{[k-1]}}\big)_* \prod_{\mathbf{w} \in \mathbf{2}^{k-1}} f_{\mathbf{w}} = 0$ $\mu^{[k-1]}$-a.e., where $f_{\mathbf{w}}$ is as in Lemma A.18.*

Proof. Using (A.2) and Lemma A.18, we can write

$$\|f\|_{U_k(\mathrm{X})}^{2^k} = \int_{X^{[k-1]}} \left| \left(\pi_{\mathcal{I}_{k-1}}^{X^{[k]}}\right)_* \prod_{\mathbf{w} \in \mathbf{2}^{k-1}} f_{\mathbf{w}} \right|^2 d\mu^{[k-1]}$$

and the claim follows.   □

The above construction is functorial:

LEMMA A.22 (Functoriality [HK, Lem. 4.5]).   *If* $\mathrm{Y} = (Y, \mathcal{B}_Y, \mu_Y, \pi_Y^X)$ *is a factor of* X, *then* $\mathrm{Y}^{[k]} = \left(Y^{[k]}, \mathcal{B}_Y^{[k]}, \mu_Y^{[k]}, (\pi_Y^X)^{[k]}\right)$ *is a factor of* $\mathrm{X}^{[k]}$, *and for every* $f \in L^\infty(\mathrm{Y})$ *one has* $\|(\pi_Y^X)^* f\|_{U^k(\mathrm{X})} = \|f\|_{U^k(\mathrm{Y})}$.

If X is ergodic with respect to G, it is certainly not the case in general that $\mathrm{X}^{[k]}$ is ergodic with respect to the diagonal action $\mathrm{diag}(G^{[k]})$. Nevertheless there are some other important ergodicity-preserving properties of the above construction:

LEMMA A.23 (Ergodicity-preserving properties [HK]).   *Let* X *be an ergodic* G-*system, and let* $k \geqslant 1$.

  (i) $\mathrm{X}^{[k]}$ *is ergodic with respect to the action of* $\partial^{[k]} G$.
 (ii) *The measure* P_k *is ergodic with respect to the action of* $\partial_+^{[k]} G$.
(iii) *For any measure-preserving transformation* $u : X \to X$ *that commutes with the* G-*action, and any side* α, *the side transformation* $u_\alpha^{[k]}$ *preserves* $\mu^{[k]}$.
(iv) *More generally, if* $u : X \to X$ *is a measure-preserving transformation that commutes with the* G-*action and leaves* $\mathcal{Z}_{<l}(\mathrm{X})$ *invariant for some* $1 \leqslant l \leqslant k$, *and* α *is a* $k - l$-*dimensional face of* $\mathbf{2}^k$, *then* $u_\alpha^{[k]}$ *preserves* $\mu^{[k]}$.

Proof. For (i), see [HK, Cor. 3.5]; for (ii), see [HK, Cor. 3.6]. For (iii) and (iv), see [HK, Lem. 5.5].   □

A.3   Dual functions.   A "soft" way to describe the universal characteristic factors $\mathcal{Z}_{<k}(\mathrm{X})$ is via the convenient device of *dual functions*.

DEFINITION A.24 (Dual functions).   Let $\mathrm{X} = (X, \mathcal{B}_X, \mu_X, (T_g)_{g \in G})$ be a G-system, and let $\{\Phi_n^i\}_{n=1}^\infty$, $i = 1, \ldots, k$, be k Følner sequences in G. Define the nonlinear operators $\mathcal{D}_k : L^\infty(\mathrm{X}) \to L^\infty(\mathrm{X})$ inductively by setting $\mathcal{D}_1 f := 1$ and $\mathcal{D}_k f := \lim_{n \to \infty} \mathbb{E}_{h \in \Phi_n^k} T_h f \cdot \mathcal{D}_{k-1}(\overline{\Delta_h f})$ for $k > 1$, where the limit is in $L^2(\mathrm{X})$.

REMARK A.25.   The limits above exist, as in Lemma A.18, by repeated applications of the ergodic theorem. Indeed, we easily verify that

$$\mathcal{D}_k f = \left(\pi_{\mathrm{X}}^{\mathrm{X}^{[k]}}\right)_* \left(\bigotimes_{\mathbf{w} \in \mathbf{2}^k \setminus \{-\mathbf{1}\}} \overline{f_{\mathbf{w}}} \right) \tag{A.6}$$

where $-\mathbf{1} := (-1, \ldots, -1)$ and the factor map is given by $(x_{\mathbf{w}})_{\mathbf{w} \in \mathbf{2}^k} \mapsto x_{-\mathbf{1}}$ (i.e. the pushforward map $\partial(\{-\mathbf{1}\})_*$). As a consequence we have

$$\|f\|_{U^k(\mathrm{X})}^{2^k} = \int_X f \overline{\mathcal{D}_k f} \, d\mu_X \tag{A.7}$$

and similarly (by repeated applications of the Cauchy–Schwarz inequality) that

$$\int_X f_1 \overline{\mathcal{D}_k f_2} \, d\mu_X \leqslant \|f_1\|_{U_k(\mathrm{X})} \cdot \|f_2\|_{U_k(\mathrm{X})}^{2^k - 1} . \tag{A.8}$$

EXAMPLE A.26. When $k = 2$, we have

$$\mathcal{D}_2 f := \lim_{m\to\infty} \lim_{n\to\infty} \mathbb{E}_{h_2\in\Phi_m^2} \mathbb{E}_{h_1\in\Phi_n^1} (T_{h_2}f)(T_{h_1}f)\overline{T_{h_1+h_2}f}\,.$$

The limit above is a repeated limit, but *a posteriori*, using Theorem 1.20, one can show that the double (simultaneous) limit exists as well, and both limits coincide, by modifying the proof of [HK, Th. 1.2], and similarly for higher values of k. We omit the details.

EXAMPLE A.27. If G is a finite abelian group, and X is G with the translation action, then for any $f : G \to \mathbb{C}$, the dual function $\mathcal{D}_k f : G \to \mathbb{C}$ is given by the formula

$$\mathcal{D}_k f(x) = \mathbb{E}_{h_1,\dots,h_k\in G} \prod_{(w_1,\dots,w_k)\in\{0,1\}^k\setminus\{0\}^k} \mathcal{C}^{w_1+\dots+w_k-1} f(x + w_1 h_1 + \dots + w_k h_k)$$

where $\mathcal{C} : z \mapsto \overline{z}$ is the complex conjugation operator. Dual functions in this setting play an important role in the finitary theory of arithmetic progressions and similar patterns; see [GrT1].

REMARK A.28. From Lemma A.22 we have the functoriality property $\mathcal{D}_k((\pi_Y^X)^* f) = (\pi_Y^X)^* \mathcal{D}_k f$ whenever Y is a factor of X and $f \in L^\infty(Y)$.

DEFINITION A.29 (Universal characteristic factor [HK, Def. 4.1]).   We let $\mathcal{Z}_{<k} = \mathcal{Z}_{<k}(X)$ be the sub-σ-algebra of $\mathcal{B}_X$ consisting of all sets $B \in \mathcal{B}_X$ such that $\left(\pi_X^{X^{[k]}}\right)^{-1}(B)$ is $\mu^{[k]}$-a.e. equivalent to a set A in $X^{[k]}$ which does not depend on the first coordinate $x_{-\mathbf{1}}$ (or equivalently, the set is invariant under the face transformation $g_{\{-\mathbf{1}\}}^{[k]}$ for all $g \in G$).

EXAMPLE A.30. Let X be an ergodic G-system. Then $\mathcal{Z}_{<1}(X)$ is trivial. One can use classical arguments to show that $\mathcal{Z}_{<2}(X)$ is the Kronecker factor (i.e. the factor generated by the eigenfunctions of X); see the discussion just before [HK, Lem. 4.2].

REMARK A.31.   It is easy to see that $\mathcal{Z}_{<k}(X)$ is an invariant sub-σ-algebra and therefore a factor. As the function on the right-hand side of (A.6) does not depend on the first coordinate, we see that the dual function $\mathcal{D}_k f$ lies in $\mathcal{Z}_{<k}(X)$ for all $f \in L^\infty(X)$.

LEMMA A.32 ($\mathcal{Z}_{<k}$ is universal [HK, Lem. 4.3]) .   For any $f \in L^\infty(X)$, we have $\|f\|_{U^k(X)} = 0$ if and only if $(\pi_{Z_{<k}(X)}^X)_* f = 0$.

REMARK A.33. From this lemma and Remark A.31 it is not hard to show that $\mathcal{Z}_{<k}(X)$ is in fact generated by the dual functions $\mathcal{D}_k f$ for $f \in L^\infty(X)$, although we will not use this fact here.

Note that Lemma A.32 immediately implies Proposition 1.10 in the introduction. From this lemma and (A.5) we also have

$$\mathcal{Z}_{<j}\big(\mathcal{Z}_{<k}(X)\big) = \mathcal{Z}_{<j}(X) \tag{A.9}$$

for $0 < j \leqslant k$ (cf. [HK, Cor. 4.4]).

From Lemma A.32 and Lemma A.22 one can show that universal characteristic factors are functorial:

LEMMA A.34 (Functoriality [HK, Prop. 4.6]).   *If* Y $= (Y, \mathcal{B}_Y, \mu_Y, \pi_Y^X)$ *is a factor of* X, *then* $\mathcal{Z}_{<k}(Y)$ *is a factor of* $\mathcal{Z}_{<k}(X)$ *for any* $k \geqslant 1$. *In fact, for any* $f \in L^\infty(Y)$, $(\pi_Y^X)^* f$ *is* $\mathcal{Z}_{<k}(X)$-*measurable if and only if* f *is* $\mathcal{Z}_{<k}(Y)$-*measurable.*

If $\phi \in L^\infty(X)$ is a phase polynomial of degree $< k$, an easy induction on k using Definition 1.13 and Definition A.24 shows that $\phi = \mathcal{D}_k \phi$. As a consequence of this and Remark A.31 we obtain the easy direction of Theorem 1.20, valid for any discrete abelian group G:

LEMMA A.35.   *Every phase polynomial of degree at most* $< k$ *is* $\mathcal{Z}_{<k}$-*measurable, or in other words* $\mathcal{P}_{<k}(X) \subset L^2(\mathcal{Z}_{<k}(X))$ *(or* $\mathrm{Abr}_{<k}(X) \leqslant \mathcal{Z}_{<k}(X)$*).*

Suppose that $f_\mathbf{w}$ is a function in $L^\infty(X)$ for each $\mathbf{w} \in \mathbf{2}^k$. From Definition A.29 and Lemma A.14 we see that the quantity $\int_{X^{[k]}} \bigotimes_{\mathbf{w} \in \mathbf{2}^k} f_\mathbf{w} \, d\mu^{[k]}$ does not change if we replace one of the $f_\mathbf{w}$ by $\mathbb{E}(f_\mathbf{w} | \mathcal{Z}_{<k})$. In particular we have

$$\int_{X^{[k]}} \bigotimes_{\mathbf{w} \in \mathbf{2}^k} f_\mathbf{w} \, d\mu^{[k]} = \int_{X^{[k]}} \bigotimes_{\mathbf{w} \in \mathbf{2}^k} \mathbb{E}(f_\mathbf{w} | \mathcal{Z}_{<k}) \, d\mu^{[k]}.$$

Thus for instance $\mathcal{D}_k f = \mathcal{D}_k \mathbb{E}(f | \mathcal{Z}_{<k})$ for all $f \in L^\infty(X)$ (and indeed $\mathcal{Z}_{<k}$ can be characterized as the minimal factor with this property). Another corollary of the above formula is

LEMMA A.36 [HK, Prop. 4.7(1)].   *The measure* $\mu^{[k]}$ *is a conditional relative product over* $\mathcal{Z}_{<k}^{[k]}(X)$. *In other words, one has a representation of the form* $\mu^{[k]} = \int_{\mathcal{Z}_{<k}^{[k]}(X)} \left(\bigotimes_{\mathbf{w} \in \mathbf{2}^k} \nu_{z,\mathbf{w}} \right) d\sigma(z)$, *where* σ *is the restriction of* $\mu^{[k]}$ *to* $\mathcal{Z}_{<k}^{[k]}$, *and for each* $z \in \mathcal{Z}_{<k}^{[k]}(X)$ *and* $\mathbf{w} \in \mathbf{2}^k$, $\nu_{z,\mathbf{w}}$ *is a probability measure on* X *which depends measurably on* z.

B   Abelian Cohomology

In this section we collect some basic facts about abelian cohomology (as defined in section 2) that we will need in the paper. Much of this machinery is essentially from [HK], but for the convenience of the reader (and given that we are generalizing from $\mathbb{Z}$-actions to more general countable abelian actions) we present the details here.

The reader may wish to review the definitions in Definition 2.1 before proceeding with the rest of this section.

We begin with the following trivial but useful lemma:

LEMMA B.1 (Cocycles and pullbacks).   *Let* G *be a locally compact group, let* X *be a* G-*system, and let* Y *be a factor of* X *with factor map* π. *Let* f *be a* (G, Y, U)-*function for some abelian group* U. *Then* f *is a* (G, Y, U)-*cocycle if and only if* $\pi^* f$ *is a* (G, X, U)-*cocycle.*

Next, we recall that cohomology is trivial for free actions:

DEFINITION B.2 (Free action).   *Let* X $= (X, \mathcal{B}_X, \mu_X, (T_g)_{g \in G})$ *be a* G-*system. The action of* G *is said to be* free *if* X *is measure-equivalent to a system of the form* $Y \times G$, *where the action of a group element* $g \in G$ *is given by the map* $(y, h) \mapsto (y, gh)$.

REMARK B.3. If G acts freely on X, then so does any compact abelian subgroup of G.

LEMMA B.4 (Free actions of compact abelian groups have no cohomology [HK, Lem. C.8]). *Let G be a compact abelian group, and let* X *be a G-system in which the action of G is free. Then every (G, X, S^1)-cocycle is a (G, X, S^1)-coboundary. In other words, $Z^1(G, \mathrm{X}, S^1) = B^1(G, \mathrm{X}, S^1)$, or equivalently $H^1(G, \mathrm{X}, S^1) = 0$.*

There is an analogue of Lemma B.4 in the polynomial category. To state it, we first need a useful algebraic lemma.

LEMMA B.5 (Composition of polynomials is again polynomial). *Let G be a countable abelian group, let U, V be abelian groups, and let* X $= $ Y $\times_\rho U$ *be an ergodic abelian extension of a G-system* Y *by a* (Y, U)-*phase polynomial ρ of degree $< k$ for some $k \geqslant 1$.*

  (i) *If p is a (X, V)-phase polynomial of degree $< d$, and $u \in U$, then $\boldsymbol{\Delta}_u p$ is a (X, V)-phase polynomial of degree $< d - 1$.*

 (ii) *If p is a (X, V)-phase polynomial of degree $< d$, and $v_1, \ldots, v_j$ are a collection of (X, U)-phase polynomials of degrees $< d_1, \ldots, < d_j$, then the (X, V)-function $P(y, u) := (\boldsymbol{\Delta}_{v_1(y,u)} \ldots \boldsymbol{\Delta}_{v_j(y,u)} p)(y, u)$ is a (X, V)-phase polynomial of degree $O_{d,j,d_1,\ldots,d_j}(1)$.*

(iii) *If p is a (X, V)-phase polynomial of degree $< d$, $v_1, \ldots, v_j$ are a collection of (X, U)-phase polynomials of degrees $< d_1, \ldots, < d_j$, and s is a (X, U)-phase polynomial of degree $< d'$, then the (X, V)-function*

$$P(y, u) := \big(\boldsymbol{\Delta}_{v_1(y,u)} \ldots \boldsymbol{\Delta}_{v_j(y,u)} p\big)\big(y, s(y, u)\big)$$

   is a (X, V)-phase polynomial of degree $O_{d,j,d_1,\ldots,d_j,d',k}(1)$.

 (iv) *For each $u \in U$, let q_u be a (X, V)-phase polynomial of degree $< m$ which obeys the U-cocycle equation*

$$q_{uv} = (V_u q_v) q_u \tag{B.1}$$

   for all $u, v \in U$. Suppose that $v_1, \ldots, v_j$ are a collection of (X, U)-phase polynomials of degrees $< d_1, \ldots, < d_j$, and r, s are (X, U)-phase polynomials of degree $< d', < d''$ respectively, then the map

$$P(y, u) := \big(\boldsymbol{\Delta}_{v_1(y,u)} \ldots \boldsymbol{\Delta}_{v_j(y,u)} q_{r(y,u)}\big)\big(y, s(y, u)\big)$$

   x is a (X, V)-polynomial of degree $O_{d,j,d',d'',k}(1)$.

Proof. We prove (i) by induction on d. Indeed, the claim is trivial for $d = 1$ by ergodicity, and for $d > 1$ we have by induction that $\boldsymbol{\Delta}_g \boldsymbol{\Delta}_u p = \boldsymbol{\Delta}_u \boldsymbol{\Delta}_g p$ is a phase polynomial of degree $< d - 2$ for all $g \in G$, and thus by (3.1) $\boldsymbol{\Delta}_u p$ is a phase polynomial of degree $< d - 1$ as claimed.

We prove (ii) by a triple induction. First, the claim is trivial when $d = 1$, so suppose that $d > 1$ and the claim has already been shown for $d - 1$. For $j \geqslant d$ the claim follows from the previous claim, so now assume that $j < d$ and the claim has already been proven for $j + 1$. Finally, the claim is clear when $d_1 + \ldots + d_j = 0$, so suppose inductively that $d_1 + \ldots + d_j > 0$ and the claim has already been proven for smaller

values of $d_1 + \ldots + d_j$. Consider the derivative $\mathbf{\Delta}_g P$ for some $g \in G$. Some computation shows that this expression can be written as $(\mathbf{\Delta}_{v_1(y,u)} \ldots \mathbf{\Delta}_{v_j(y,u)} \mathbf{\Delta}_g p)(y, u)$, times a product of finitely many expressions of the form $(\mathbf{\Delta}_{v'_1(y,u)} \ldots \mathbf{\Delta}_{v'_{j'}(y,u)} p)(y, u)$ where j' is either larger than j, or $j = j'$ and the total degree of $v'_1, \ldots, v'_j$ is less than that of $v_1, \ldots, v_j$. Using the various induction hypotheses we conclude that $\mathbf{\Delta}_g P$ is a (X, S^1)-phase polynomial of degree $O_{d,d_1,\ldots,d_j}(1)$, and the claim follows from (3.1).

Claim (iii) is proven by the same inductive argument as the previous claim; the non-linear nature of $s(y, u)$ introduced some new terms when one differentiates, but all such terms increase the number j of vertical derivatives (and only involve polynomials in the subscripts, thanks to the polynomial nature of s and ρ) and so can be safely handled by the induction hypothesis.

Finally, we prove claim (iv). The case $d' = 0$ follows from the previous claim, so suppose that $d' > 0$ and the claim has already been proven for the smaller values of d'. For $g \in G$, we take a derivative $\mathbf{\Delta}_g P$. One obtains essentially the same terms that appeared in the previous claim, plus (thanks to the cocycle equation (B.1)) some additional terms involving $q_{\Delta_g r(y,u)}$. But such terms can be dealt with by the induction hypothesis. $\qquad\square$

Lemma B.6 (Polynomial integration lemma). *Let G be a countable abelian group, let $m, k \geqslant 1$, and let $X = Y \times_\rho U$ be an ergodic abelian extension of a G-system Y by a (G, Y, U)-phase polynomial cocycle ρ of degree $< k$. Let V be a locally compact abelian group. For each $u \in U$, let q_u be a (X, V)-phase polynomial of degree $< m$ which obeys the U-cocycle equation (B.1) for all $u, v \in U$. Then there exists a (X, V)-phase polynomial Q of degree $< O_{m,k}(1)$ such that $q_u = \mathbf{\Delta}_u Q$ for all $u \in U$.*

Proof. Let u_0 be a generic element of U, and define the (X, V)-function Q by the formula $Q(y, vu_0) := q_v(y, u_0)$. for all $y \in Y$ and $v \in U$. Observe that for any $u \in U$, we have

$$\mathbf{\Delta}_u Q(y, vu_0) = \frac{q_{uv}(y, u_0)}{q_v(y, u_0)} = q_u(y, vu_0)$$

thanks to (B.1). Thus we have $q_u = \mathbf{\Delta}_u Q$ for all $u \in U$.

The fact that Q is a (X, V)-phase polynomial of degree $O_{m,k}(1)$ follows from the $j = 0$ case of Lemma B.5(iv). $\qquad\square$

Remark B.7. One can improve the degree bounds in Lemma B.5 and Lemma B.6 if one assumes that $Y = \mathcal{Z}_{<j}(X)$ for some j; see Lemma 8.14 and Proposition 8.9.

We will also need another result in a similar spirit.

Lemma B.8 (Straightening nearly translation-invariant cocycles [HK, Lem. C.9]). *Let G be a countable abelian group, let X be an ergodic G-system, let $K = (K, \cdot)$ be a compact abelian group acting freely on X and commuting with the G action, and let ρ be an (G, X, H)-function for some compact abelian H such that $\mathbf{\Delta}_k \rho : (g, x) \mapsto \rho(g, T_k x)/\rho(g, x)$ is a coboundary for all $k \in K$. Then ρ is cohomologous to a (G, X, H)-function which is invariant with respect to some open subgroup U of K.*

We will also take advantage of a useful splitting lemma.

LEMMA B.9 (Splitting lemma [HK, Lem. C.5] (see also [MoS])).   *Let G be a countable abelian group, let* X *be an ergodic G-system. Let $f : \mathrm{X} \to S^1$ be a (G, X, S^1)-cocycle such that $d^{[1]}f$ is a $(G, \mathrm{X}^{[1]}, S^1)$-coboundary. Then f is (G, X, S^1)-cohomologous to a constant cocycle (i.e. a cocycle that is independent of the* X *coordinate). In other words, we have an exact sequence*

$$H^1(G, \mathrm{pt}, S^1) \longrightarrow H^1(G, \mathrm{X}, S^1) \longrightarrow H^1\big(G, \mathrm{X}^{[1]}, S^1\big).$$

where the first map is the map induced by the factor map $\pi_{\mathrm{pt}}^{\mathrm{X}}$, and the second map is the map induced by the derivative map $f \mapsto d^{[1]}f$.

LEMMA B.10 (Cohomology of $X^{[l]}$ injects into cohomology of $X^{[k]}$ [HK, Lem. C.7]).   *Let G be a countable abelian group, let* X *be an ergodic G-system, let U be a compact abelian group, and let $k \geqslant l \geqslant 0$. Let α be an l-face of $\mathbf{2}^k$, thus by Lemma A.15 $\partial(\alpha)_* : \mathrm{X}^{[k]} \to \mathrm{X}^{[l]}$ is a factor map. Let f be a $\big(G, \mathrm{X}^{[l]}, U\big)$-cocycle such that the $\big(G, \mathrm{X}^{[k]}, U\big)$-cocycle $(\partial(\alpha)_*)^* f$ is a $\big(G, \mathrm{X}^{[k]}, U\big)$-coboundary. Then f is also a $(G, \mathrm{X}^{[l]}, U)$-coboundary. In other words, we have an exact sequence*

$$0 \longrightarrow H^1\big(G, \mathrm{X}^{[l]}, U\big) \longrightarrow H^1\big(G, \mathrm{X}^{[k]}, U\big).$$

Now we see how cohomology on an abelian extension relates to cohomology on the base space.

LEMMA B.11 (Descent lemma).   *Let $k \geqslant 1$, let G be a countable abelian group, and let* $\mathrm{X} = (X, \mathcal{B}_X, \mu_X, (T_g)_{g \in G})$ *be an ergodic G-system. Let ρ be an abelian (G, X, U)-cocycle for some compact abelian U. Let $\phi : U \to V$ be a surjective measurable homomorphism from U to another compact abelian group V, and let $\pi_{\mathrm{X} \times_{\phi \circ \rho} V}^{\mathrm{X} \times_\rho U}$ be the associated factor map $(x, u) \mapsto (x, \phi(u))$. Suppose that f is a $(G, \mathrm{X} \times_{\phi \circ \rho} V, S^1)$-function is such that $\big(\pi_{\mathrm{X} \times_{\phi \circ \rho} V}^{\mathrm{X} \times_\rho U}\big)^* f$ is $(G, \mathrm{X} \times_\rho U, S^1)$-cohomologous to $\big(\pi_{\mathrm{X} \times_{\phi \circ \rho} V}^{\mathrm{X} \times_\rho U}\big)^* p$ for some $(G, \mathrm{X} \times_{\phi \circ \rho} V, S^1)$-function p. Then f is $(G, \mathrm{X} \times_{\phi \circ \rho} V, S^1)$-cohomologous to $p\big(\chi \circ \rho \circ \pi_{\mathrm{X}}^{\mathrm{X} \times_{\phi \circ \rho} V}\big)$ for some character $\chi \in \hat{U}$.*

Proof. Let us first consider the case when V (and ϕ) is trivial, so that f is now a (G, X, S^1)-function. By hypothesis, there exists a $(\mathrm{X} \times_\rho U, S^1)$-function F such that

$$f(g, x) = p(g, x) \frac{F(T_g x, \rho(g, x) u)}{F(x, u)}$$

for all $g \in G$ and almost every $x \in X$, $u \in U$. We rearrange this as

$$F\big(T_g x, \rho(g, x) u\big) = F(x, u) \overline{p}(g, x) f(g, x). \tag{B.2}$$

We perform a Fourier expansion in U, obtaining

$$F(x, u) = \sum_{\chi \in \hat{U}} F_\chi(x) \chi(u)$$

for some $F_\chi \in L^\infty(\mathrm{X})$, not all identically zero. Comparing Fourier coefficients in (B.2), we conclude that $F_\chi(T_g x) \chi \circ \rho(g, x) = F_\chi(x) \overline{p}(g, x) f(g, x)$ for all $g \in G$ and $\chi \in \hat{U}$, and μ_X-almost every $x \in X$. In particular this shows that the function $|F_\chi|$

is G-invariant and thus (by ergodicity) constant. Thus there exists $\chi \in \hat{U}$ such that $|F_\chi|$ is almost everywhere non-vanishing. We can then write

$$f(g, x) = p(g, x) \chi \circ \rho(g, x) \frac{F_\chi(T_g x)}{F_\chi(x)} \, .$$

This completes the proof in the case when V is trivial.

To handle the general case we perform a lifting trick. Observe that the system $\mathrm{X}' := \mathrm{X} \times_{\rho \oplus \phi \circ \rho} U \times V$ is a simultaneous extension of both $\mathrm{X} \times_\rho U$ and $\mathrm{X} \times_{\phi \circ \rho} V$. Since $\left(\pi^{\mathrm{X} \times_\rho U}_{\mathrm{X} \times_{\phi \circ \rho} V}\right)^* f$ is $(G, \mathrm{X} \times_\rho U, S^1)$-cohomologous to $\left(\pi^{\mathrm{X} \times_\rho U}_{\mathrm{X} \times_{\phi \circ \rho} V}\right)^* p$, we see that $\left(\pi^{\mathrm{X}'}_{\mathrm{X} \times_{\phi \circ \rho} V}\right)^* f$ is (G, X', S^1)-cohomologous to $\left(\pi^{\mathrm{X}'}_{\mathrm{X} \times_{\phi \circ \rho} V}\right)^* p$. Writing X' as an extension of $\mathrm{X} \times_{\phi \circ \rho} V$ by the cocycle $\rho \circ \pi^{\mathrm{X} \times_{\phi \circ \rho}}_{\mathrm{X}}$, and applying the previous result, we obtain the claim. $\qquad \square$

C   A Measurable Selection Lemma

Suppose that G is a countable abelian group and X is an ergodic G-system. In our arguments we will frequently have a family of (G, X, S^1)-functions h_u, parameterized in some measurable fashion by a parameter u in a compact abelian group U, such that h_u takes values in $\mathcal{P}_k(G, \mathrm{X}, S^1) \cdot B^1(G, \mathrm{X}, S^1)$. In other words (by the axiom of choice), for each u, we may find $\psi_u \in \mathcal{P}_k(G, \mathrm{X}, S^1)$ and $F_u \in M(\mathrm{X}, S^1)$ such that for each $u \in U$, we have the equation

$$h_u = \psi_u \boldsymbol{\Delta} f_u \, . \tag{C.1}$$

Unfortunately, due to the use of the axiom of choice, it is not necessarily the case that the functions ψ_u and F_u that arise here are measurable. Fortunately, one can resolve this problem by using some separation properties of phase polynomials and the hypothesis that X is separable. The basic tool here is the following (cf. [GrT5, Lem. 7.1]):

LEMMA C.1 (Separation lemma).   Let G be a countable abelian group, let $\mathrm{X} = (X, \mathcal{B}_X, \mu_X, (T_g)_{g \in G})$ be an ergodic G-system, let $k \geqslant 1$, and let $\phi, \psi \in \mathcal{P}_{<k}(\mathrm{X})$ be such that ϕ/ψ is non-constant. Then $\|\phi - \psi\|_{L^2(\mathrm{X})} \geqslant \sqrt{2}/2^{k-2}$.

REMARK C.2. The constant $\sqrt{2}/2^{k-2}$ can be improved slightly, but for our purposes any quantity that is independent of X would suffice here.

Proof. By dividing ϕ, ψ by ψ we may assume $\psi = 1$.

The claim is vacuous when $k = 1$. When $k = 2$ we argue as follows. For any $h \in G$ we have

$$\int_X \phi \, d\mu_X = \int_X T_h \phi \, d\mu = \int_X (\boldsymbol{\Delta}_h \phi) \phi \, d\mu_X \, . \tag{C.2}$$

If ϕ is a phase polynomial of degree < 2, then $\boldsymbol{\Delta}_h \phi$ is constant; if ϕ is non-constant, then (by ergodicity) $\boldsymbol{\Delta}_h \phi$ is not identically 1 for at least one h. Thus $\int_X \phi \, d\mu_X = 0$ and so $\|\phi - 1\|_{L^2(X)} = \sqrt{2}$, and the claim follows.

Now suppose inductively that $k \geqslant 3$ and the claim has already been proven smaller values of k. Suppose for contradiction that there was a non-constant

$\phi \in \mathcal{P}_{<k}(\mathrm{X})$ such that $\|\phi - 1\|_{L^2(\mathrm{X})} < \sqrt{2}/2^{k-2}$. Arguing as in (C.2) we conclude that $\|(\boldsymbol{\Delta}_h \phi)\phi - 1\|_{L^2(\mathrm{X})} < \sqrt{2}/2^{k-2}$ for all h, and thus by the triangle inequality

$$\|\boldsymbol{\Delta}_h \phi - 1\|_{L^2(\mathrm{X})} = \left\|(\boldsymbol{\Delta}_h \phi)\phi - \phi\right\|_{L^2(\mathrm{X})} < \sqrt{2}/2^{k-3}.$$

But $\boldsymbol{\Delta}_h \phi \in \mathcal{P}_{<k-1}(\mathrm{X})$. By the induction hypothesis we conclude that $\boldsymbol{\Delta}_h \phi$ is constant for every h, or in other words that $\phi \in \mathcal{P}_{<2}(\mathrm{X})$. The claim then again follows from the induction hypothesis. □

Since $L^2(\mathrm{X})$ is separable, we conclude

COROLLARY C.3 (At most countably many polynomials modulo constants). *Let G be a countable abelian group, let $\mathrm{X} = (X, \mathcal{B}_X, \mu_X, (T_g)_{g \in G})$ be an ergodic G-system, let $k \geqslant 1$. Then the collection $\mathcal{P}_{<k}(\mathrm{X})$, after quotienting out by constants, is at most countable.*

We are now ready to establish the measurable selection lemma.

LEMMA C.4 (Measurable selection lemma). *Let G be a countable abelian group. Let X be an ergodic G-system, and let $k \geqslant 1$. Let U be a compact abelian group. If $u \to h_u$ is a Borel measurable map from U to $\mathcal{P}_{<k}(G, \mathrm{X}, S^1) \cdot B^1(G, \mathrm{X}, S^1) \subset M(G, \mathrm{X}, S^1)$ (where we give the latter the topology of convergence in measure), then there is a Borel measurable choice of f_u, ψ_u (as functions from U to $M(\mathrm{X}, S^1)$ and $\mathcal{P}_{<k}(G, \mathrm{X}, S^1)$ respectively) obeying (C.1).*

Proof. For each $u \in U$, write

$$\Omega_u := \left\{ f \in M(\mathrm{X}, S^1) : h_u/\boldsymbol{\Delta} f \in \mathcal{P}_{<k}(G, \mathrm{X}, S^1) \right\},$$

then Ω_u is non-empty for each u by hypothesis. Also, if $f_u, f'_u \in \Omega_u$, then $\boldsymbol{\Delta}(f_u/f'_u) \in \mathcal{P}_{<k}(G, \mathrm{X}, S^1)$, and so (by (3.1)) $f_u/f'_u \in \mathcal{P}_{<k+1}(\mathrm{X})$; reversing this argument, we conclude that Ω_u is a coset of $\mathcal{P}_{<k+1}(\mathrm{X})$ in $M(\mathrm{X}, S^1)$ for each u.

As $L^2(\mathrm{X})$ is separable, one can find a countable sequence $F_1, F_2, \ldots \in M(\mathrm{X}, S^1)$ which is dense in $M(\mathrm{X}, S^1)$. For each u, let n_u be the first integer such that there exists $k \geqslant 1$ and $f_u \in \Omega_u$ with $\|f_u - F_{n_u}\|_{L^2(\mathrm{X})} < \sqrt{2}/2^{k+1}$; this integer exists by density, and clearly depends in a measurable fashion on u. By Lemma C.1 and the triangle inequality, we see that the $f_u \in \Omega_u$ which lie within $\sqrt{2}/2^{k+1}$ of F_{n_u} are all constant multiples of each other. There is thus a unique $f_u \in \Omega_u$ which minimizes the distance $\|f_u - F_{n_u}\|_{L^2(\mathrm{X})}$. Selecting this f_u (and then solving for ψ_u using (C.1)) we obtain the claim. □

REMARK C.5. One can also establish this result using a general measure selection result of Dixmier (see e.g. [BK, Th. 1.2.4]) together with Lusin's theorem and Corollary C.3; we omit the details. One can also appeal to the descriptive set theory of Polish groups, see e.g. [HK, App. A].

D   Finite Characteristic Algebra

In this appendix we collect some algebraic facts that exploit the finite characteristic of the underlying field $\mathbb{F}$ (or the finite torsion of various abelian groups).

D.1   Compact abelian torsion groups.   Recall that a group U is *m-torsion* if we have $u^m = 1$ for all $u \in U$.

LEMMA D.1 (Open sets of torsion groups contain open subgroups).   *Let U be a compact abelian m-torsion group for some $m \geqslant 1$. Let V be an open neighborhood of the identity in U. Then V contains an open subgroup W of U.*

Proof.   We will use a Fourier-analytic method. As V is an open neighborhood of the origin, one can find another open neighborhood V' of the origin such that $V' - V' \subset V$.

Let μ be the Haar measure on U, then $\mu(V') > 0$. Let $\varepsilon > 0$ be a small number (depending on $\mu(V')$) to be chosen later. By Fourier analysis, we can approximate the indicator function $1_{V'}$ to within ε in $L^2(U)$-norm by some linear combination F of finitely many characters $\chi_1, \ldots, \chi_n \in \hat{U}$, where n is finite but potentially unbounded. Since U is m-torsion, each character χ_j takes on at most m values, with each level set of χ_j being a coset of an open subgroup of U. If we let W be the intersection of the kernels of all the χ_j, then W is also an open subgroup of U, and F is constant on every coset of W. Since F approximates $1_{V'}$ to within ε, we conclude (if ε is sufficiently small depending on $\mu(V')$) that there exists a coset of W on which V' has density greater than $1/2$. But then this forces $W \subset V' - V'$ and hence $W \subset U$, as desired. $\square$

LEMMA D.2 (Splitting lemma).   *Let U be a compact abelian m-torsion group for some $m \geqslant 1$. Let W be an open subgroup of U. Then there exists a splitting $U = W' \times Y$, where W' is an open subgroup of W, and Y is a finite abelian m-torsion group.*

Proof.   It is known (see e.g. [Mor, Ch. 5, Th. 18]) that a compact abelian m-torsion group U is topologically isomorphic to the direct product of cyclic m-torsion groups. (In particular, the bounded torsion allows us to avoid having to deal with procyclic groups which are not direct products of cyclic groups.) Thus W must contain a cylinder neighbourhood W' of the origin, i.e. a cofinite sub-product of these cyclic groups. Since one clearly has the desired splitting $U = W' \times Y$, the claim follows. $\square$

D.2   Polynomials are discretely valued.   An important fact about phase polynomials over $\mathbb{F}^\omega$, which is not true for polynomials over some other groups G (such as the integers $\mathbb{Z}$), is that such polynomials only take finitely many values. More precisely, if we let $C_n := \{z \in \mathbb{C} : z^n = 1\}$ denote the cyclic group of nth roots of unity, we have

LEMMA D.3 (Phase polynomials over $\mathbb{F}^\omega$ are discretely valued).   *Let $\mathbb{F}$ be a finite field of characteristic p, and let $\mathrm{X} = (X, \mathcal{B}_X, \mu_X, (T_g)_{g \in \mathbb{F}^\omega})$ be an ergodic $\mathbb{F}^\omega$-system.*

  (i) *If $f \in \mathcal{P}_{<k}(\mathrm{X}, S^1)$ for some $k \geqslant p$, then $f^p \in \mathcal{P}_{<k-p+1}(\mathrm{X}, S^1)$.*
  (ii) *If $f \in \mathcal{P}_{<k}(\mathrm{X}, S^1)$ for some $k \geqslant 1$, then (after multiplying f by a constant), f takes values in $C_{p^{\lfloor (k-2)/(p-1) \rfloor + 1}}$. In other words, $\mathcal{P}_{<k}(\mathrm{X}) = S^1 \cdot \mathcal{P}_{<k}(\mathrm{X}, C_{p^{\lfloor (k-2)/(p-1) \rfloor + 1}})$.*
  (iii) *If $f \in \mathcal{P}_{<k}(\mathrm{X}, S^1)$ for some $k \geqslant 1$, and f takes values in $C_{p^{\lfloor (k-2)/(p-1) \rfloor + 1}}$, then for any $g \in \mathbb{F}^\omega$, $\prod_{i=0}^{p-1} T_g^i f$ takes values in $C_{p^{\lfloor k/p \rfloor}}$.*

(iv) If $f \in \mathcal{P}_{<k}(\mathbb{F}^\omega, X, S^1)$ is a cocycle for some $k \geqslant 1$, then f takes values in $C_{p^{\lfloor k/p \rfloor + 1}}$.

Proof. To prove (i), it suffices to verify it in the case $k = p$, since the higher cases then follow by induction and from the identity $\mathbf{\Delta}_g(f^p) = (\mathbf{\Delta}_g f)^p$. Taking logarithms, it suffices to show that if $F : X \to \mathbb{R}/\mathbb{Z}$ is a polynomial of degree $< p$, then pf is constant.

Let $g \in G$. Since $T_g^p f = f$ and $T_g = 1 + \Delta_g$, we conclude using the binomial formula that $\sum_{i=0}^{p} \binom{p}{i} \Delta_g^i f = f$. Since f has degree $< p$, $\Delta_g^p f = 0$. We conclude that

$$p\Delta_g f + \binom{p}{2}\Delta_g f + \ldots + p\Delta_g^{p-1} f = 0$$

which we rewrite as

$$\left(1 + \frac{p-1}{2}\Delta_g + \cdots + \Delta_g^{p-2}\right)\Delta_g p f = 0\,.$$

Inverting the expression in brackets using Neumann series (and using the fact that Δ_g^{p-1} annihilates $\Delta_g p f$) we conclude that $\Delta_g p f = 0$ for any g, thus by ergodicity pf is constant as claimed.

To prove (ii), we first observe that it suffices to prove the claim for k of the form $k = pm + 1$ for integer m. But the claim is trivial for $m = 1$, and from (i), we see that the claim for m implies the claim for $m + 1$, and so (ii) follows by induction.

To prove (iii), it suffices to do so for k of the form $k = pm - 1$ for some integer m, since the claim is trivial otherwise. By (i), the claim for m implies the claim for $m-1$, so it suffices to verify the case $m = p - 1$. Taking logarithms, it suffices to show that if $F : X \to \mathbb{R}/\mathbb{Z}$ is a polynomial of degree $< p - 1$ with $pF = 0$, then $\sum_{i=0}^{p-1} T_g^i F = 0$. But writing $T_g = 1 + \Delta_g$ we obtain the identity $\sum_{i=0}^{p-1} T_g^i = \sum_{i=0}^{p-1} \binom{p}{i+1}\Delta_g^i$ and the claim follows, since $\binom{p}{i+1}$ is a multiple of p for all $0 \leqslant i < p - 1$, and Δ_g^{p-1} annihilates F.

To prove (iv), observe from the cocycle equation that $\prod_{i=0}^{p-1} f(g, T_g^i x) = 1$ for all $g \in \mathbb{F}^\omega$ and almost all $x \in X$. On the other hand, from (ii) we know that for fixed g, $f(g, \cdot)$ is equal to a constant c_g times a polynomial taking values in $C_{p^{\lfloor (k-2)/(p-1) \rfloor + 1}}$, so by (iii), $\prod_{i=0}^{p-1} f(g, T_g^i x)$ is equal to c_g^p times a quantity in $C_{p^{\lfloor k/p \rfloor}}$. Thus $c_g \in C_{p^{\lfloor k/p \rfloor + 1}}$, and the claim follows. $\square$

REMARK D.4. The claims are sharp. For instance, in the characteristic 2 space $\mathbb{F}_2^\omega$ (which acts on itself by translations), the function $\phi_k : (x_1, x_2, \ldots) \mapsto e^{2\pi i (\sum_{j=1}^\infty |x_j|)/2^k}$ for $k \geqslant 1$, where $x \mapsto |x|$ is the obvious map from $\mathbb{F}_2$ to $\{0, 1\}$, is a phase polynomial of degree k which takes values in C_{2^k}, but in no smaller group.

D.3   Roots of phase polynomials.   We now develop some machinery that will allow us to take m^{th} roots of phase polynomials and still obtain a phase polynomial. It will be convenient to use the notation $O_{a_1,\ldots,a_k}(1)$ to denote a quantity bounded in magnitude by some constant $C(a_1, \ldots, a_k)$ depending only on $a_1, \ldots, a_k$. Throughout this appendix, $\mathbb{F}$ is a finite field of characteristic p, and X is an $\mathbb{F}^\omega$-system.

For any cyclic p-group $\mathbb{Z}/p^m\mathbb{Z}$, and any $0 \leqslant j \leqslant m - 1$, let $b_j = b_{j,m} : \mathbb{Z}/p^m\mathbb{Z} \to \{0, 1, \ldots, p - 1\}$ be the jth digit map, thus $x = \sum_{j=0}^{m-1} b_j(x)p^j$ for all $x \in \mathbb{Z}/p^m\mathbb{Z}$.

Recall that a map $P : X \to H$ into an *additive* group H is a polynomial of degree $< d$ if we have $\Delta_{g_1} \ldots \Delta_{g_d} P = 0$ for all $g_1, \ldots, g_d \in G$.

PROPOSITION D.5 (Digits of polynomials are polynomials). Let $P : X \to \mathbb{Z}/p^m\mathbb{Z}$ be *a polynomial of degree* $< d$, *and let* $\mathbb{Z}/p^l\mathbb{Z}$ *be a cyclic group. Embed* $\{0, 1, \ldots, p-1\}$ *into* $\mathbb{Z}/p^l\mathbb{Z}$ *in the obvious manner. Then for any* $0 \leqslant j < m-1$, $b_j(P)$ *is a polynomial of degree* $< O_{l,d,p,j}(1)$.

Proof. We may assume inductively that the claim is already proven for smaller values of d; for the same value of d and smaller values of l; or the same value of d and l and smaller values of j. We abbreviate $O_{l,d,p,j}(1)$ as $O(1)$.

From primary school arithmetic we know that we have a formula of the form

$$b_j(x+y) = b_j(x) + b_j(y) + c_j\big(b_0(x), \ldots, b_{j-1}(x), b_0(y), \ldots, b_{j-1}(y)\big)$$
$$- pc_{j+1}\big(b_0(x), \ldots, b_j(x), b_0(y), \ldots, b_j(y)\big)$$

for some "carry bit" functions $c_j : \{0, 1, \ldots, p-1\}^{2j} \to \{0, 1\}$. Applying this with P and $\Delta_g P$ for some group element g we conclude

$$b_j(T_g P) = b_j(P) + b_j(\Delta_g P) + c_j\big(b_0(P), \ldots, b_{j-1}(P), b_0(\Delta_g P), \ldots, b_{j-1}(\Delta_g P)\big)$$
$$- pc_{j+1}\big(b_0(P), \ldots, b_j(P), b_0(\Delta_g P), \ldots, b_j(\Delta_g P)\big)$$

and so

$$\Delta_g\big(b_j(P)\big) = b_j(\Delta_g P) + c_j\big(b_0(P), \ldots, b_{j-1}(P), b_0(\Delta_g P), \ldots, b_{j-1}(\Delta_g P)i\big)$$
$$- pc_{j+1}\big(b_0(P), \ldots, b_j(P), b_0(\Delta_g P), \ldots, b_j(\Delta_g P)\big).$$

By the induction hypothesis on d, $b_j(\Delta_g P)$ is already a polynomial of some degree $O(1)$. By the induction hypothesis on j, $b_0(P), \ldots, b_{j-1}(P), b_0(\Delta_g P), \ldots, b_{j-1}(\Delta_g P)$ are also polynomials of degree $O(1)$. The carry function c_j, by Lagrange interpolation, can be expressed as a polynomial in $\mathbb{Z}/p^l\mathbb{Z}$ of its arguments; the key point here is that as the arguments lie in $\{0, \ldots, p-1\}$, the denominators in the Lagrange interpolation formula contain no factors of p and are thus invertible. We thus conclude that $c_j(b_0(P), \ldots, b_{j-1}(P), b_0(\Delta_g P), \ldots, b_{j-1}(\Delta_g P))$ is also a polynomial of degree $O(1)$.

Finally, by the induction hypothesis on l, we know that

$$b_0(P), \ldots, b_j(P), b_0(\Delta_g P), \ldots, b_j(\Delta_g P)$$

are polynomials of degree $O(1)$ in $\mathbb{Z}/p^{l-1}\mathbb{Z}$, and thus by arguing as before

$$c_{j+1}(b_0(P), \ldots, b_j(P), b_0(\Delta_g P), \ldots, b_j(\Delta_g P))$$

is a polynomial of degree $O(1)$ in $\mathbb{Z}/p^{l-1}\mathbb{Z}$. This implies that

$$pc_{j+1}\big(b_0(P), \ldots, b_j(P), b_0(\Delta_g P), \ldots, b_j(\Delta_g P)\big)$$

is a polynomial of degree $O(1)$ in $\mathbb{Z}/p^l\mathbb{Z}$.

Putting all this together we see that $\Delta_g(b_j(P))$ is a polynomial of degree $O(1)$ for all g, and hence $b_j(P)$ is a polynomial of degree $O(1)$, thus closing the induction. $\square$

COROLLARY D.6 (Functions of phase polynomials are phase polynomial).   *Let $\phi_1, \ldots, \phi_m$ be (X, S^1)-phase polynomials of degree $< d$ for some $d, m \geqslant 1$, let $n \geqslant 1$, and let $F(\phi_1, \ldots, \phi_m)$ be some function of $\phi_1, \ldots, \phi_m$ taking values in the cyclic group C_{p^n}. Then $F(\phi_1, \ldots, \phi_m)$ is a (X, S^1)-phase polynomial of degree $O_{p,d,m,n}(1)$.*

Proof. We have the freedom to rotate each of the ϕ_j by a constant. By Lemma D.3, this allows us to assume that all the $\phi_1, \ldots, \phi_m$ take values in C_{p^d} (say), thus $\phi_j = e(P_j/p^d)$ for some additive polynomials $P_j : X \to C_{p^d}$ of degree $< d$. Now observe that F can be viewed as a $e(G((b_i(P_j))_{0 \leqslant i < d; 1 \leqslant j \leqslant m})/p^n)$ for some function $G : \{0, \ldots, p-1\}^{md} \to \mathbb{Z}/p^n\mathbb{Z}$. By Lagrange interpolation, G can be viewed as the restriction of a polynomial from $(\mathbb{Z}/p^n\mathbb{Z})^{md}$ to $\mathbb{Z}/p^n\mathbb{Z}$ with degree $O_{d,m,p,n}(1)$. The claim now follows from Proposition D.5.   □

We isolate one special case of Corollary D.6:

COROLLARY D.7 (Phase polynomials have phase polynomial roots).   *Let ϕ be a (X, S^1)-phase polynomial of degree $< d$ for some $d \geqslant 1$, and let $n \geqslant 1$. Then there exists a phase polynomial ψ of degree $O_{d,n,p}(1)$ such that $\psi^n = \phi$.*

Proof. By rotating ϕ by a constant and using Lemma D.3, we may assume that ϕ takes values in C_{p^m} for some $m = O_{d,p}(1)$. If n is not divisible by p, then n is invertible in C_{p^m} and the claim is immediate, so it suffices to check the case when n is a power of p. But then the claim follows immediately from Corollary D.6.   □

Another interesting consequence of Corollary D.6 (or Proposition D.5) is that phase polynomials can always be expressed in terms of C_p-valued polynomials of higher degree.

COROLLARY D.8 (Representation of phase polynomials by C_p-valued phase polynomials).   *Let ϕ be a (X, S^1)-phase polynomial of degree $< d$ for some $d \geqslant 1$. Then ϕ can be expressed as a function of $O_{d,p}(1)$ many C_p-valued (X, S^1)-phase polynomials of degree $< O_{d,p}(1)$.*

E   Connection with Cubic Complexes

In this appendix we point out some connections between the notions of polynomiality and type in this paper with the theory of cubic complexes as used in topology, as set out in [MP], in analogy with the more well-known simplicial complexes used in that field. This material is not used elsewhere in this paper.

Abstractly, a cubic complex is a sequence of spaces $X^{[k]}, X^{[k-1]}, \ldots, X^{[0]}$, together with maps $\partial(\alpha)_* : X^{[k]} \to X^{[l]}$ for every l-face α of $\mathbf{2}^k$, such that one has the relation $\partial(\alpha)_* \partial(\beta)_* = \partial(\partial(\beta)(\alpha))_*$ whenever $k \geqslant l \geqslant m$, β is an l-face of $\mathbf{2}^k$, and α is an m-face of $\mathbf{2}^l$ (so that $\partial(\beta)(\alpha)$ is an m-face of $\mathbf{2}^k$. Note that the concrete cubic complex defined in Definition A.2 is of this form.

Let X be an ergodic G-system for some countable abelian group G, and let $U = (U, +)$ a locally compact abelian group, which we now express additively for compatibility with [MP]. By Lemma 4.3(iii), a (G, X, U)-function f is an a (G, X, U)-polynomial of degree $< k$ if and only if $d^{[k]} f = 0$ $\mu^{[k]}$-almost everywhere. In the

language of [MP], this is equivalent to f being of *degree* $< k$ in the sense of cubic complexes.

In the case that U is uniquely divisible by $1, \ldots, k$ (i.e. for every $1 \leqslant j \leqslant k$ and $u \in U$ there is a unique solution $u/j \in U$ to the equation $j(u/j) = u$), one can express the operator $d^{[k]} : M(G, \mathrm{X}, U) \to M(G, \mathrm{X}^{[k]}, U)$ as the composition of the k differentiation operators $d : M(G, \mathrm{X}^{[j-1]}, U) \to M(G, \mathrm{X}^{[j]}, U)$ defined by $df(g, x) := \frac{1}{j} \sum_{i=1}^{j} (f(g, \partial_i^+ \mathbf{x}) - f(g, \partial_i^- \mathbf{x}))$.

References

[A]   L.M. ABRAMOV, Metric automorphisms with quasi-discrete spectrum, Izv. Akad. Nauk SSSR Ser. Mat. 26 (1962), 513–530; English transl.: Amer. Math. Soc. Transl. (2) 39 (1964), 37–56.

[BK]   H. BECKER, A. KECHRIS, The Descriptive Set Theory of Polish Group Actions, Lond. Math. Soc. Lecture Notes Series 232, Cambridge University Press, (1996).

[BeHK]   V. BERGELSON, B. HOST, B. KRA (appendix by I. RUZSA), Multiple recurrence and nilsequences, Inventiones Math. 160:2 (2005), 261–303.

[BeL]   V. BERGELSON, S. LEIBMAN, Distribution of values of bounded generalized polynomials, Acta. Math 198:2 (2007), 155–230.

[CL]   J.-P. CONZE, E. LESIGNE, Théorèmes ergodiques pour des mesures diagonales, Bull. Soc. Math. France 112 (1984), 143–175.

[FHK]   N. FRANTZIKINAKIS, B. HOST, B. KRA, hMultiple recurrence and convergence for sequences related to the prime numbers, Crelles Journal (J. Reine Angew. Math.) 611 (2007), 131–144.

[FW]   N. FRANTZIKINAKIS, M. WIERDL, A Hardy field extension of Szemeredi's theorem, preprint.

[Fu1]   H. FURSTENBERG, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, J. Analyse Math 31 (1977), 204–256.

[Fu2]   H. FURSTENBERG, Recurrence in Ergodic Theory and Combinatorial Number Theory, Princeton University Press, New Jersey, 1981.

[FuB]   H. FURSTENBERG, B. WEISS, A mean ergodic theorem for $\frac{1}{N} \sum_{n=1}^{N} f(T^n x) g(T^{n^2} x)$, Convergence in Ergodic Theory and Probability (Columbus, OH, 1993), Ohio State Univ. Math. Res. Inst. Publ., 5 de Gruyter, Berlin (1996), 193–227.

[G1]   T. GOWERS, A new proof of Szemeredi's theorem, Geom. Func. Anal. 11 (2001), 465–588.

[GW]   T. GOWERS, J. WOLF, The true complexity of a system of linear equations, preprint.

[GrT1]   B. GREEN, T. TAO, The primes contain arbitrarily long arithmetic progressions, Annals of Math. 167 (2008), 481–547.

[GrT2]   B. GREEN, T. TAO, An inverse theorem for the Gowers $U^3(G)$ norm, preprint.

[GrT3]   B. GREEN, T. TAO, Linear equations in primes, Annals of Math., to appear.

[GrT4]   B. GREEN, T. TAO, New bounds for Szemerédi's Theorem, I: Progressions of length 4 in finite field geometries, preprint.

[GrT5]   B. GREEN, T. TAO, The distribution of polynomials over finite fields, and applications to the Gowers Inverse conjecture, preprint.

[HK]   B. Host, B. Kra, Non-conventional ergodic averages and nilmanifolds, Ann.
        Math. 161:1 (2005), 397–488.

[L]   E. Lesigne, Théorèmes ergodiques pour une translation sur une nilvarieté, Ergod.
        Th. Dynam. Sys. 9:1 (1989), 115–126.

[LoMS] S. Lovett, R. Meshulam, A. Samorodnitsky, Inverse Conjecture for the Gow-
        ers norm is false , preprint.

[MP]   S. Matveev, M. Polyak, Cubic complexes and finite type invariants, Geometry
        and Topology Monographs, Vol. 4: Invariant Knots and 3-manifolds (Kyoto 2001),
        215–233.

[MoS]   C. Moore, K. Schmidt, Coboundaries and homomorphisms for non-singular ac-
        tions and a problem of H. Helson, Proc. London Math. Soc. (3) 40 (1980), 443–475.

[Mor]   S.A. Morris, Pontryagin Duality and the Structure of Locally Compact Abelian
        Groups, London Math. Soc. Lecture Note Series 29. Cambridge University Press,
        1977.

[P]   W. Parry, Dynamical systems on nilmanifolds, Bull. London Math. Soc. 2 (1970),
        37–40.

[S]   A. Samorodnitsky, Low-degree tests at large distances, preprint.

[Sz]   E. Szemerédi, On sets of integers containing no k elements in arithmetic progres-
        sion, Acta Arith. 27 (1975), 299–345.

[TZ]   T. Tao, T. Ziegler, The inverse conjecture for the Gowers norm over finite fields
        via the correspondence principle, preprint.

[Z]   T. Ziegler, Universal characteristic factors and Furstenberg averages, J. Amer.
        Math. Soc. 20 (2007), 53–97.

Vitaly Bergelson, Department of Mathematics, Ohio State University, Columbus OH
43210-1174, USA   vitaly@math.ohio-state.edu

Terence Tao, UCLA Department of Mathematics, Los Angeles, CA 90095-1555, USA
                                                        tao@math.ucla.edu

Tamar Ziegler, Department of Mathematics, Technion, Haifa 32000, Israel
                                                        tamarzr@tx.technion.ac.il