IDEMPOTENT ULTRAFILTERS AND POLYNOMIAL RECURRENCE

CHRISTIAN SCHNELL

ABSTRACT. We give a new proof of a polynomial recurrence result due to Bergelson, Furstenberg, and McCutcheon, using idempotent ultrafilters instead of IP-limits.

INTRODUCTION

In the thirty or so years since H. Furstenberg reproved Szemerédi's theorem using methods from ergodic theory, many striking discoveries have been made in the area now known as *Ergodic Ramsey theory*. Perhaps the most surprising of these is the discovery that recurrence results can be obtained for polynomial sets, meaning sets of values of polynomials. The following pretty theorem, a special case of a more general theorem proved by V. Bergelson, H. Furstenberg, and R. McCutcheon in [1], is a typical result in this direction.

Theorem. *Let $\mathcal{F}$ be the collection of all non-empty finite subsets of $\mathbb{N}$. For any polynomial $p \in \mathbb{Z}[x_1, \ldots, x_k]$ satisfying $p(0, \ldots, 0) = 0$, and for any IP-sets $\{n_\alpha^{(1)}\}_{\alpha \in \mathcal{F}}, \ldots, \{n_\alpha^{(k)}\}_{\alpha \in \mathcal{F}}$, the set*

$$R = \big\{ p(n_\alpha^{(1)}, \ldots, n_\alpha^{(k)}) \,\big|\, \alpha \in \mathcal{F} \big\}$$

is a set of nice recurrence.

To say that $R \subseteq \mathbb{Z}$ is a set of *nice recurrence* means that for any probability space $(X, \mathcal{B}, \mu)$, and any invertible measure-preserving transformation T on X, one has

$$\limsup_{n \in R} \mu(A \cap T^n A) \geq \mu(A)^2$$

for all $A \in \mathcal{B}$. Moreover, an *IP-set* is any set of the form

$$\Big\{ n_\alpha = \sum_{i \in \alpha} n_i \,\Big|\, \alpha \in \mathcal{F} \Big\},$$

for positive integers $n_0, n_1, n_2, \ldots$.

As in Furstenberg's result, this inequality has immediate combinatorial applications. It also turned out that the above theorem was only a first step; much stronger results—combining IP-convergence, multiple recurrence as in Szemerédi's theorem, and polynomial sets—have since been established, for instance in [2].

The purpose of the present paper is to give a different proof for the central result of [1], using *idempotent ultrafilters* instead of IP-limits. While this approach is less constructive, it has the advantage of "making the statements and proofs cleaner

1

and more algebraic," in the words of the survey paper [3]. It also follows the general philosophy that for each result about IP-sets, there should be an analogous result about idempotent ultrafilters.

The main theorem and its proof are presented in Section 5; however, a better point to begin reading is probably Section 2, which treats a special but typical case, and explains the method of proof in some detail. Section 6 contains a small number of applications, of the type mentioned above.

Since ultrafilters on groups and semigroups are used throughout the paper, their basic properties are reviewed in Section 1; readers who are already familiar with $\beta\mathbb{N}$, for instance from [3], will recognize all the material, despite the more general context. To keep the paper self-contained, several generally known results about operators and integer-valued polynomials have also been included; these make up Sections 3 and 4.

Note. Vitaly Bergelson, who advised me during my first two years in graduate school, suggested the problem of reproving the results in [1] using idempotent ultrafilters. I am very grateful to him for his help, as well as for countless pleasant conversations. Unlike wine, the paper has failed to mature during the several years that it has been stored on the hard drive of my computer; nevertheless, I have decided to make it available, since it is in my opinion a nice application of idempotent ultrafilters to recurrence results.

1. THE STONE-ČECH COMPACTIFICATION OF A DISCRETE SEMIGROUP

Ultrafilters. We begin by reviewing the definition and several basic properties of the space of ultrafilters. Let $(S, \circ)$ be a commutative semigroup. An *ultrafilter* on S is a collection p of subsets of S with the following four properties:

(1) $S \in p$ and $\emptyset \notin p$.
(2) If $A \in p$, and $B \supseteq A$, then $B \in p$.
(3) If $A, B \in p$, then $A \cap B \in p$.
(4) For every $A \subseteq S$, either $A \in p$, or $S \setminus A \in p$.

For every $s \in S$, there is a *principal* or *trivial* ultrafilter consisting of all subsets containing s; the construction of other ultrafilters requires the Axiom of Choice.

The space βS of all ultrafilters on S, suitably topologized, is the Stone-Čech compactification of the discrete space S. After briefly stating the basic properties of βS, we will consider two examples: one where S is the group $\mathbb{Z}^n$, and a second one where S equals $\mathcal{F}$, the set of nonempty finite subsets of $\mathbb{N}$. A good and very comprehensive reference for this topic is the book by Hindman and Strauss [4].

Terminology. Since ultrafilters are collections of sets, the following terminology is convenient when dealing with their members. If p is an ultrafilter on S, we call a set p-*big* if it is contained in p; we shall also use the phrase '*for p-many s*' to mean 'for all s in some p-big set.' In the case of several variables, we shall say that $\langle$statement$\rangle$ holds '*for p-many $s_1, \ldots, s_n$*' if

$$\{ s_1 \in S \mid \{ s_2 \in S \mid \cdots \{ s_n \in S \mid \langle\text{statement}\rangle \} \in p \cdots \} \in p \} \in p.$$

In other words, there should be p-many s_1, for which there are p-many s_2, for which $\ldots$, for which there are p-many s_n, for which $\langle$statement$\rangle$ is true. Nested sets of exactly this form will play a role during the proof of the main theorem in Section 5.

Basic properties. As was said above, we let βS be the set of ultrafilters on S, and consider S as a subset of βS, by identifying an element of S with the principal ultrafilter it generates. One can put a topology on βS, in which the sets

$$\bar{A} = \{p \in \beta S \mid A \in p\} \qquad \text{(for } A \subseteq S\text{)}$$

give a basis for the closed sets; each $\bar{A}$ is both closed and open. The result is a compact space (this includes the Hausdorff property) that has S as a discrete and dense subspace.

The semigroup operation $\circ$ extends to βS; given p and q in βS, their product $p \circ q$ may be defined by the property that for any $A \subseteq S$,

$$A \in p \circ q \iff \{s \in S \mid \{t \in S \mid s \circ t \in A\} \in q\} \in p.$$

The new operation is associative and continuous from the left (meaning that for any q, the map $p \mapsto p \circ q$ is continuous), and makes βS into a compact left-topological semigroup.

Idempotent ultrafilters. An ultrafilter $p \in \beta S$ is called *idempotent* if it satisfies the relation $p \circ p = p$. Idempotent ultrafilters are closely related to *IP-sets*, which are sets of the form

$$\left\{\prod_{i \in \alpha} s_i \,\middle|\, \alpha \subset \mathbb{N} \text{ finite, nonempty}\right\},$$

for a given sequence $(s_i)_{i \in \mathbb{N}}$. Any member of an idempotent ultrafilter contains an IP-set, and conversely, every IP-set is contained in some idempotent ultrafilter. This fact is sometimes called Hindman's theorem (see [3, Theorem 3.4] for details); it implies that one can find many idempotent ultrafilters (provided, as usual, that the Axiom of Choice is assumed).

When a finite sequence $s_1, \ldots, s_n$ is used in place of an infinite one, we shall denote the resulting *finite IP-set* by $\mathrm{IP}(s_1, \ldots, s_n)$. The proof that any member of an idempotent has to contain an IP-set allows a much stronger conclusion if we are only looking for finite IP-sets.

Lemma 1. *Let $p \in \beta S$ be an idempotent ultrafilter. If A is a p-big set, then for any $n \in \mathbb{N}$ one has*

$$\{s_1 \in A \mid \cdots \{s_n \in A \mid \mathrm{IP}(s_1, \ldots, s_n) \subseteq A\} \in p \cdots\} \in p.$$

In the terminology introduced above, one can say that there are p-many $s_1, \ldots, s_n$ in A such that $\mathrm{IP}(s_1, \ldots, s_n) \subseteq A$.

Limits along ultrafilters. Another useful notion is that of a *p-limit*, or a limit along some ultrafilter. Let $p \in \beta S$ be an ultrafilter. Given a map $f \colon S \to Y$ into some topological space Y, we say that a point y is a limit of f along p, written

$$y = p\text{-}\lim_s f(s),$$

if for every neighborhood U of y, the set $f^{-1}(U)$ is p-big. When the target space Y is compact, all p-limits exist and are unique.

This notion of limit is related to the Stone-Čech compactification in the following manner. A *compactification* of a Hausdorff space X is a compact space containing X as a dense subspace. The *Stone-Čech compactification βX* is the universal compactification, in the sense that for any compact space Y and any continuous map

$f \colon X \to Y$, there is one and only one continuous extension f_* from βX to Y, as illustrated in the diagram.

$$
\begin{array}{ccc}
X & \hookrightarrow & \beta X \\
{\scriptstyle f}\downarrow & \swarrow{\scriptstyle f_*} & \\
Y & &
\end{array}
$$

Every other compactification is a quotient of βX; furthermore, if $g \colon Y \to Z$ is a second continuous map of compact spaces, one has $(fg)_* = f_* g_*$ because of the uniqueness statement.

Now the space βS, as defined above, is the Stone-Čech compactification of the discrete topological space S; given any map $f \colon S \to Y$ into a compact space Y, the required extension $f_* \colon \beta S \to Y$ is given by

$$
f_*(p) = p\text{-}\lim_s f(s) \qquad\qquad (\text{for } p \in \beta S),
$$

which is continous as a map from βS to Y.

The following lemma is an immediate consequence of the universal property.

Lemma 2. *Any map $\phi \colon S \to T$ between two semigroups S and T induces a continuous map $\phi_* \colon \beta S \to \beta T$, given by $\phi_*(p) = p\text{-}\lim_s \phi(s)$. A set B is $\phi_*(p)$-big if, and only if, its preimage $\phi^{-1}(B)$ is p-big. For any map $f \colon T \to Y$ into a compact space Y, one has*

$$
p\text{-}\lim_s f(\phi(s)) = \phi_*(p)\text{-}\lim_t f(t).
$$

If ϕ is multiplicative, so is ϕ_*; in particular, $\phi_*(p)$ is then always idempotent for idempotent $p \in \beta S$.

There is another important property of p-limits, especially useful for our purposes.

Lemma 3. *Let p and q be two elements of βS. The equality*

$$
p\text{-}\lim_s q\text{-}\lim_t f(s \circ t) = (p \circ q)\text{-}\lim_s f(s)
$$

holds for any map $f \colon S \to Y$ into a compact space Y.

Proof. Let $y = p\text{-}\lim_s q\text{-}\lim_t f(s \circ t)$; for any neighborhood U of y, the set

$$
\left\{ s \,\middle|\, q\text{-}\lim_t f(s \circ t) \in U \right\}
$$

is p-big. Equivalently,

$$
\left\{ s \,\middle|\, \left\{ t \,\middle|\, f(s \circ t) \in U \right\} \in q \right\} \in p,
$$

and this is nothing but the condition $f^{-1}(U) \in p \circ q$. It follows that the right-hand limit $(p \circ q)\text{-}\lim_s f(s)$ also equals y. $\qquad\square$

The lemma explains one useful aspect of idempotent ultrafilters—if p is an idempotent, one has

(1.1) $$ p\text{-}\lim_s p\text{-}\lim_t f(s \circ t) = p\text{-}\lim_s f(s), $$

and this relation is at the base of all applications of ultrafilters to recurrence results.

As an application, let us prove a lemma known as *van der Corput's trick*, for p-limits. It provides a useful sufficient condition for a weak p-limit in a Hilbert space to be zero.

Lemma 4. *Let Y be a closed ball in a Hilbert space $\mathcal{H}$, endowed with the weak topology (and thus compact). Given a map $f : S \to Y$ and an idempotent $p \in \beta S$, let $y = p\text{-}\lim_s f(s)$. If*

$$p\text{-}\lim_s p\text{-}\lim_t \langle f(s \circ t), f(t) \rangle = 0,$$

then $y = 0$.

Proof. One uses (1.1) in a clever way. Notice that for any $N \in \mathbb{N}$,

$$y = p\text{-}\lim_{s_1} \cdots p\text{-}\lim_{s_N} \frac{1}{N} \sum_{n=1}^{N} f(s_n \circ \cdots \circ s_N).$$

Using weak lower semi-continuity of the norm, we obtain

$$\left\| y \right\|^2 \le p\text{-}\lim_{s_1} \cdots p\text{-}\lim_{s_N} \frac{1}{N^2} \left\| \sum_{n=1}^{N} f(s_n \circ \cdots \circ s_N) \right\|^2$$

$$= p\text{-}\lim_{s_1} \cdots p\text{-}\lim_{s_N} \frac{1}{N^2} \left\langle \sum_{m=1}^{N} f(s_m \circ \cdots \circ s_N), \sum_{n=1}^{N} f(s_n \circ \cdots \circ s_N) \right\rangle$$

$$= \frac{1}{N^2} \sum_{m,n} p\text{-}\lim_{s_1} \cdots p\text{-}\lim_{s_N} \langle f(s_m \circ \cdots \circ s_N), f(s_n \circ \cdots \circ s_N) \rangle,$$

and after collapsing the multiple p-limits with the help of (1.1), this becomes

$$= \frac{1}{N^2} \sum_n p\text{-}\lim_s \left\| f(s) \right\|^2 + \frac{2}{N^2} \operatorname{Re} \sum_{m<n} p\text{-}\lim_s p\text{-}\lim_t \langle f(s \circ t), f(t) \rangle$$

$$= \frac{1}{N} p\text{-}\lim_s \left\| f(s) \right\|^2.$$

Since N was arbitrary, we see that $y = 0$. $\qquad\square$

We shall now discuss two concrete examples of semigroups and their Stone-Čech compactifications, namely $\beta \mathbb{Z}^n$ and $\beta \mathcal{F}$.

Abelian groups and $\beta \mathbb{Z}^n$. We are going to use vector notation for elements of $\mathbb{Z}^n$, such as $\vec{a} = (a_1, \ldots, a_n)$. Even though $\mathbb{Z}^n$ is a group, the space of ultrafilters $\beta \mathbb{Z}^n$ is only a semigroup, because there are in general no inverses for elements. Still, we can get information about ultrafilters in $\beta \mathbb{Z}^n$ from the group structure of $\mathbb{Z}^n$; in particular, we shall investigate the relationship between subgroups and idempotent ultrafilters.

Every subgroup of $\mathbb{Z}^n$ is itself free, of rank between 0 and n. The first observation is that subgroups of rank n are contained in every idempotent ultrafilter.

Lemma 5. *For every idempotent $p \in \beta \mathbb{Z}^n$, all rank n subgroups are p-big.*

Proof. A subgroup L of rank n necessarily has finite index. As an ultrafilter, p thus has to contain one of the cosets, say $\vec{z} + L$, and as an idempotent, it then has to contain the set

$$\left\{ \vec{a} \in \vec{z} + L \,\middle|\, \{ \vec{b} \in \vec{z} + L \mid \vec{a} + \vec{b} \in \vec{z} + L \} \in p \right\}$$

as well. In particular, that set is nonempty. The resulting equation $\vec{z} + \vec{z} \equiv \vec{z}$ mod L gives $\vec{z} + L = L$, and we can conclude that L itself is p-big. □

We now define the *dimension* of an ultrafilter p, denoted $\dim p$, to be the smallest possible rank of a p-big subgroup of $\mathbb{Z}^n$. Since we expect p-big sets to be large (especially when p is an idempotent), it would be nice if the dimension of an ultrafilter in $\beta\mathbb{Z}^n$ was always n. This is not true; for instance, the principal ultrafilter generated by 0 is idempotent, and has dimension zero. But as the following lemma shows, in all such examples, the ultrafilter in question really lives on a smaller group.

Lemma 6. *Let $p \in \beta\mathbb{Z}^n$ be an ultrafilter, of dimension $s \in \{0, \ldots, n\}$. If $G \subseteq \mathbb{Z}^n$ is an arbitrary p-big subgroup of rank s, then there is an injective group homomorphism $\phi \colon \mathbb{Z}^s \to \mathbb{Z}^n$ with image G, and an s-dimensional ultrafilter $q \in \beta\mathbb{Z}^s$, such that $p = \phi_*(q)$. If p is idempotent, then any such q is idempotent as well.*

Proof. Let G be a p-big subgroup of rank s in $\mathbb{Z}^n$. Since it is free, it is isomorphic to $\mathbb{Z}^s$, and so there is an injective group homomorphism $\phi \colon \mathbb{Z}^s \to \mathbb{Z}^n$ whose image is exactly G. Define

$$q = \big\{ A \subseteq \mathbb{Z}^s \,\big|\, \phi(A) \in p \big\};$$

since G is p-big, it is easily verified that $q \in \beta\mathbb{Z}^s$, and that $\phi_*(q) = p$. Now q has to have dimension s, for otherwise $\mathbb{Z}^s$, and therefore also G, would contain a p-big subgroup of smaller rank, contradicting the choice of s.

Now assume that p is an idempotent ultrafilter. Since ϕ_* is a homomorphism, we get

$$\phi_*(q \circ q) = \phi_*(q) \circ \phi_*(q) = p \circ p = p;$$

but as q is obviously uniquely determined by the condition that $\phi_*(q) = p$, it follows that $q \circ q = q$, and so q is idempotent as well. □

IP-sets and $\beta\mathcal{F}$. Our second example is the Stone-Čech compactification of $\mathcal{F}$, the set of finite nonempty subsets of $\mathbb{N}$. For any two such finite sets α and β, we may form their union $\alpha \cup \beta$; this operation makes $\mathcal{F}$ into a commutative semigroup. An ultrafilter in this setting is now a set of sets of finite subsets of $\mathbb{N}$; to avoid confusion, we shall reserve the letters α, β, γ for points of $\mathcal{F}$. We also continue to write $\circ$ for the semigroup operation on $\beta\mathcal{F}$. The character of this operation is utterly different from addition on $\beta\mathbb{Z}$; for instance, any principal ultrafilter is now idempotent.

We are mostly going to look at IP-sets in $\mathbb{N}$ and $\mathcal{F}$ from the point of view of $\beta\mathcal{F}$. Since the semigroup operation on $\mathbb{N}$ is addition, an IP-set is now a set of the form

$$\Big\{ \sum_{i \in \alpha} n_i \,\Big|\, \alpha \in \mathcal{F} \Big\},$$

where $(n_i)_{i \in \mathbb{N}}$ is a sequence of positive integers. This can also be considered as a map

$$n_\bullet \colon \mathcal{F} \to \mathbb{N}, \qquad \alpha \mapsto n_\alpha = \sum_{i \in \alpha} n_i$$

that satisfies

$$(1.2) \qquad\qquad\qquad n_{\alpha \cup \beta} = n_\alpha + n_\beta$$

for disjoint $\alpha, \beta \in \mathcal{F}$. It induces a map n_* from $\beta\mathcal{F}$ to $\beta\mathbb{N}$ that we would like to be structure-preserving, in particular with regard to idempotents, but this cannot

be true. The problem is that the map $n_\bullet$ fails to be additive because (1.2) holds for disjoint sets only. One answer is to look at a subclass of ultrafilters in $\beta\mathcal{F}$, excluding—among other things—the principal ones.

One quickly sees that in order to make use of (1.2), it has to be possible, when choosing β inside a member of some ultrafilter, to make it disjoint from a given α. To accomplish this, we introduce the following notion. We let

$$C_n = \big\{\alpha \in \mathcal{F} \,\big|\, n \in \alpha\big\} \qquad\qquad \text{(for } n \in \mathbb{N}\text{)},$$

and call an ultrafilter *congested* if it contains one of the C_n, or *uncongested* if it contains none. Certainly, every principal ultrafilter is congested. We also introduce the notation $\alpha < \beta$ to express that the maximum of the finite set α is less than the minimum of β.

Whenever p is an uncongested ultrafilter and $\alpha \in \mathcal{F}$, the set

$$\big\{\beta \in \mathcal{F} \,\big|\, \alpha < \beta\big\}$$

is obviously p-big, being an intersection of complements of certain C_n. This means that for any set $A \in p$, the set of $\beta \in A$ with $\alpha < \beta$ is still p-big, and so we can impose the even stronger condition $\alpha < \beta$ when choosing an element β from A.

Now let us see what the set $\beta\mathcal{F}^{un}$ of all uncongested ultrafilters looks like. An *IP-ring* is a special type of IP-set in $\mathcal{F}$; it consists of an infinite sequence $(\alpha_i)_{i\in\mathbb{N}}$ of elements of $\mathcal{F}$ satisfying $\alpha_0 < \alpha_1 < \alpha_2 < \cdots$, together with all possible finite unions of these. The notation

$$\mathrm{IP}(\alpha_i)_{i\in\mathbb{N}}$$

will be used for such IP-rings. We then have the following result about $\beta\mathcal{F}^{un}$ and its connection with IP-rings.

Lemma 7. *$\beta\mathcal{F}^{un}$ is a closed (hence compact) sub-semigroup of $\beta\mathcal{F}$. Every idempotent $p \in \beta\mathcal{F}^{un}$ has the property that if a set is p-big, it contains an IP-ring. Conversely, every IP-ring is a member of some uncongested idempotent.*

Proof. By definition,

$$\beta\mathcal{F}^{un} = \bigcap_{n\in\mathbb{N}} \overline{\mathcal{F} \setminus C_n}$$

is an intersection of closed sets, hence closed. Let us show that it is also a semigroup. If a product $p \circ q$ is congested, it has to contain C_n for some n, and so

$$\big\{\alpha \,\big|\, \{\beta \,\big|\, n \in \alpha \cup \beta\} \in q\big\} \in p.$$

If C_n is not in p, one of the α in the outer set does not contain n, in which case the inner q-big set equals C_n. Either way, one of the two factors is congested; products of uncongested ultrafilters are therefore uncongested. It follows that $\beta\mathcal{F}^{un}$ is a compact semigroup.

To verify the second statement—existence of IP-rings in p-big sets for idempotent p—the same proof as for Hindman's theorem will work; when choosing elements, one simply follows the recipe mentioned above.

Third, let us show that every IP-ring is contained in some uncongested idempotent (this also proves the existence of uncongested idempotents). Let $\mathrm{IP}(\alpha_i)_{i\in\mathbb{N}}$ be an IP-ring and set $A_n = \mathrm{IP}(\alpha_i)_{i\geq n}$, with obvious meaning. Following the usual procedure, we will show that the intersection

$$\mathcal{A} = \bigcap_{n\in\mathbb{N}} \overline{A_n}$$

is a closed, nonempty subsemigroup of $\beta\mathcal{F}^{un}$; by Ellis' theorem it then has idempotents, and any such idempotent is uncongested and contains our IP-ring.

$\mathcal{A}$ is certainly closed and nonempty (use the finite intersection property of the compact space $\beta\mathcal{F}$). To verify that it is a semigroup, we need to check that for $p, q \in \mathcal{A}$ and any $n \in \mathbb{N}$, the set A_n is a member of $p \circ q$. Given $\alpha \in A_n$, there is some $k > n$ for which $\alpha \cup \beta \in A_n$ for every $\beta \in A_k$. Hence

$$\left\{ \beta \,\middle|\, \alpha \cup \beta \in A_n \right\}$$

is a q-big set for every $\alpha \in A_n$; as a consequence, we have

$$\left\{ \alpha \,\middle|\, \left\{ \beta \,\middle|\, \alpha \cup \beta \in A_n \right\} \in q \right\} \in p \qquad \text{(for } n \in \mathbb{N}\text{)},$$

and thus $p \circ q \in \mathcal{A}$. Finally, every $p \in \mathcal{A}$ has to be uncongested, for A_{n+1} and C_n are always disjoint. $\qquad\square$

One conclusion is that uncongested idempotents do exist; more importantly, they naturally arise when one is looking at IP-rings in terms of ultrafilters. Indeed, the lemma is the exact analogon of Hindman's theorem for the case of IP-rings.

Another useful property of uncongested ultrafilters is stated in the last lemma of this section; it closely follows our thoughts after (1.2).

Lemma 8. *For any IP-set $n_\bullet \colon \mathcal{F} \to \mathbb{N}$, the induced map $n_* \colon \beta\mathcal{F}^{un} \to \beta\mathbb{N}$ is a homomorphism of semigroups. In particular, $n_*(p)$ is idempotent for each uncongested idempotent $p \in \beta\mathcal{F}^{un}$.*

Proof. Let $p, q \in \beta\mathcal{F}^{un}$ be arbitrary uncongested ultrafilters; we need to show that

$$n_*(p \circ q) = n_*(p) \circ n_*(q).$$

For any subset $A \subseteq \mathbb{N}$, we have $A \in n_*(p)$ exactly when $\left\{ \alpha \,\middle|\, n_\alpha \in A \right\} \in p$; referring back to the definition of the operation $\circ$ on page 3, we then find that

$$A \in n_*(p \circ q) \quad \Longleftrightarrow \quad \left\{ \alpha \,\middle|\, \left\{ \beta \,\middle|\, n_{\alpha \cup \beta} \in A \right\} \in q \right\} \in p,$$

while

$$A \in n_*(p) \circ n_*(q) \quad \Longleftrightarrow \quad \left\{ \alpha \,\middle|\, \left\{ \beta \,\middle|\, n_\alpha + n_\beta \in A \right\} \in q \right\} \in p.$$

But since q is uncongested, these two conditions are actually equivalent. Indeed, given $\alpha \in \mathcal{F}$, we have $n_{\alpha \cup \beta} = n_\alpha + n_\beta$ whenever $\beta > \alpha$, and so

$$\left\{ \beta \,\middle|\, n_{\alpha \cup \beta} \in A \right\} \cap \left\{ \beta \,\middle|\, \beta > \alpha \right\} = \left\{ \beta \,\middle|\, n_{\alpha \cup \beta} \in A \text{ and } \beta > \alpha \right\}$$
$$= \left\{ \beta \,\middle|\, n_\alpha + n_\beta \in A \text{ and } \beta > \alpha \right\}$$
$$= \left\{ \beta \,\middle|\, n_\alpha + n_\beta \in A \right\} \cap \left\{ \beta \,\middle|\, \beta > \alpha \right\}.$$

Now q always contains the set $\left\{ \beta \,\middle|\, \beta > \alpha \right\}$, and so

$$\left\{ \beta \,\middle|\, n_{\alpha \cup \beta} \in A \right\} \in q \quad \Longleftrightarrow \quad \left\{ \beta \,\middle|\, n_\alpha + n_\beta \in A \right\} \in q.$$

This shows that $A \in n_*(p \circ q)$ if, and only if, $A \in n_*(p) \circ n_*(q)$, and thus proves the lemma. $\qquad\square$

This result will later allow us to transfer results from $\beta\mathbb{N}$ or $\beta\mathbb{Z}^n$ to the space $\beta\mathcal{F}$. We shall see applications of this idea in Section 5, after we have proved the main theorem.

2. An extended example

In this section, we want to give an in-depth discussion of a special case of the main results, Theorem 5.1 and Theorem 5.2. We hope that this will help the reader understand the character of the argument—in particular, how the induction used in the proof works. We are going to consider the following theorem.

Theorem 2.1. *Let U be an arbitrary unitary operator on a Hilbert space $\mathcal{H}$, and let $m_\bullet, n_\bullet \colon \mathcal{F} \to \mathbb{N}$ be any two IP-sets. If $p \in \beta\mathcal{F}$ is any uncongested idempotent, the operator P defined by the weak operator limit*

$$P = p\text{-}\lim_\alpha U^{m_\alpha n_\alpha}$$

is an orthogonal projection.

The given (weak operator) limit abbreviates the equality

$$\langle Px, y \rangle = p\text{-}\lim_\alpha \langle U^{m_\alpha n_\alpha} x, y \rangle$$

for all $x, y \in \mathcal{H}$.

We are not going to prove this directly, because the presence of the two IP-sets is inconvenient, in that it obscures part of the underlying structure. Suppose, for example, that the two IP-sets were (more or less) equal; then the essentially two-dimensional situation of the theorem would collapse down to a one-dimensional one, and surely something in the proof will have to change, too. The problem, in other words, is that there appears to be a notion of *dimension* behind the theorem—but it is cumbersome to deal with dimension for IP-sets.

On the other hand, as shown by Lemma 6, there is a good definition of dimension for idempotent ultrafilters in $\beta\mathbb{Z}^2$. Instead of trying to prove Theorem 2.1 in its present form, we should pass instead to the group $\mathbb{Z}^2$, where we can talk about the rank of subgroups and the dimension of ultrafilters.

To this end, define a map $\phi \colon \mathcal{F} \to \mathbb{Z}^2$ by

$$\phi(\alpha) = \big(m_\alpha, n_\alpha\big) \qquad\qquad \text{(for } \alpha \in \mathcal{F}\text{)}.$$

Because p is uncongested, the new ultrafilter $q = \phi_*(p)$ is an idempotent in $\beta\mathbb{Z}^2$ by Lemma 8; moreover, Lemma 2 changes the limit defining P into

$$P = q\text{-}\lim_{\vec{z}} U^{z_1 z_2},$$

where the notation $\vec{z}$ is again used for elements of $\mathbb{Z}^2$. The following more general statement now suggests itself.

Theorem 2.2. *Let U be a unitary operator on a Hilbert space $\mathcal{H}$. If $q \in \beta\mathbb{Z}^2$ is any idempotent, the operator P defined by the weak operator limit*

$$P = q\text{-}\lim_{\vec{z}} U^{z_1 z_2}$$

is an orthogonal projection.

Even in this special case, a proof seems to require two separate steps. We begin by introducing an auxiliary operator

$$Q = q\text{-}\lim_{\vec{a}} q\text{-}\lim_{\vec{z}} U^{a_1 z_2 + a_2 z_1},$$

the polynomial in the exponent arising from the original $z_1 z_2$ as

$$(a_1 + z_1)(a_2 + z_2) - z_1 z_2 - a_1 a_2.$$

Step 1. For the time being, we are going to assume that Q is a projection operator, and use the splitting of the Hilbert space $\mathcal{H} = \ker Q \oplus \operatorname{im} Q$ it induces to prove Theorem 2.2. To show that P is an orthogonal projection, we appeal to Lemma 9: P is clearly normal, being a limit of unitary operators, and so all we need to do is prove the relation $P^2 = P$. To help with that, let us also create, for each $\vec{a} \in \mathbb{Z}^2$, the operator

$$Q_{\vec{a}} = q\text{-}\lim_{\vec{z}} U^{a_1 z_2 + a_2 z_1}.$$

Since the polynomials involved are linear in $\vec{z}$, the reader will prove without much effort that each $Q_{\vec{a}}$ is an orthogonal projection; for example, one can use the identity in (1.1) to show that $Q_{\vec{a}}^2 = Q_{\vec{a}}$, and then apply Lemma 9. Of course, any two of those operators commute, since they are all limits of powers of U.

The point is that under our assumption on Q, the weak operator limits

$$Q = q\text{-}\lim_{\vec{a}} Q_{\vec{a}} = q\text{-}\lim_{\vec{a}} q\text{-}\lim_{\vec{z}} U^{a_1 z_2 + a_2 z_1}$$

are actually strong ones, as we shall see. To prove the identity $P^2 = P$, let us first consider the situation on the space $\ker Q$. If x satisfies $Qx = 0$, we get

$$q\text{-}\lim_{\vec{a}} \left\| Q_{\vec{a}} x \right\|^2 = q\text{-}\lim_{\vec{a}} \left\langle Q_{\vec{a}} x, x \right\rangle = \left\langle Qx, x \right\rangle = 0,$$

and because we have convergence in the norm, we can apply van der Corput's trick to show $Px = 0$. The condition in Lemma 4,

$$q\text{-}\lim_{\vec{a}} q\text{-}\lim_{\vec{z}} \left\langle U^{(a_1 + z_1)(a_2 + z_2)} x, U^{z_1 z_2} x \right\rangle = q\text{-}\lim_{\vec{a}} q\text{-}\lim_{\vec{z}} \left\langle U^{a_1 z_2 + a_2 z_1} x, U^{-a_1 a_2} x \right\rangle$$
$$= q\text{-}\lim_{\vec{a}} \left\langle Q_{\vec{a}} x, U^{-a_1 a_2} x \right\rangle = 0,$$

is satisfied, and we conclude that $Px = 0$, hence $P^2 x = Px$.

Next, let us see what happens if $x \in \operatorname{im} Q$. In this case, $Qx = x$, and we can write

$$q\text{-}\lim_{\vec{a}} q\text{-}\lim_{\vec{z}} \left\| U^{a_2 z_2 + a_2 z_1} x - x \right\|^2$$
$$= 2\left\| x \right\|^2 - 2 \cdot \operatorname{Re} q\text{-}\lim_{\vec{a}} q\text{-}\lim_{\vec{z}} \left\langle U^{a_1 z_2 + a_2 z_1} x, x \right\rangle = 2\left\| x \right\|^2 - 2 \left\langle Qx, x \right\rangle = 0,$$

from which it follows that

$$(2.1) \qquad q\text{-}\lim_{\vec{a}} q\text{-}\lim_{\vec{z}} \left\| U^{(a_1 + z_1)(a_2 + z_2)} x - U^{a_1 a_2} U^{z_1 z_2} x \right\| = 0.$$

To obtain $P^2 x = Px$, we make use of the identity in (1.1) for double q-limits; together with (2.1), we obtain

$$Px = q\text{-}\lim_{\vec{z}} U^{z_1 z_2} x = q\text{-}\lim_{\vec{a}} q\text{-}\lim_{\vec{z}} U^{(a_1 + z_1)(a_2 + z_2)} x$$
$$= q\text{-}\lim_{\vec{a}} q\text{-}\lim_{\vec{z}} U^{a_1 a_2} U^{z_1 z_2} x = P^2 x.$$

We are therefore able to show that P is an orthogonal projection, provided that Q is one. The device of getting strong from weak convergence is frequently useful, by the way; it is formalized in Lemma 10 below.

Step 2. So far, we have been able to reduce Theorem 2.2 to the proof of the following, simpler result.

Theorem 2.3. *Let U be a unitary operator on a Hilbert space $\mathcal{H}$. For any idempotent $q \in \beta\mathbb{Z}^2$, the operator*

$$Q = q\text{-}\lim_{\vec{a}} q\text{-}\lim_{\vec{z}} U^{a_1 z_2 + a_2 z_1}$$

is an orthogonal projection.

Just as in Step 1, everything hinges on having a good splitting of the underlying Hilbert space $\mathcal{H}$. But which splitting one should use depends on the ultrafilter q, more precisely on its dimension—which could be 0, 1, or 2. We will treat these as separate cases here; in the proof of the main theorem, we shall of course want a unified approach.

Dimension 0. If q is 0-dimensional, it contains the set $\{(0,0)\}$, and since we can restrict to a q-big set when taking limits, Q is simply the identity operator. So this case is trivial.

Dimension 1. In case $\dim q = 1$, we can find a subgroup $\mathbb{Z}\vec{c}$ (with $\vec{c} \neq 0$) of rank one in q. Accordingly, we will use the splitting $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_1^\perp$, where

$$\mathcal{H}_1 = \bigcap_{n \neq 0} \ker Q_{n\vec{c}} \qquad \text{and} \qquad \mathcal{H}_1^\perp = \overline{\sum_{n \neq 0} \operatorname{im} Q_{n\vec{c}}}.$$

It is then straightforward to show that Q is orthogonal projection onto $\mathcal{H}_1^\perp$.

Indeed, if x is an element of $\mathcal{H}_1$, then $Q_{n\vec{c}}x = 0$ holds for all nonzero n, and since $\vec{c}$ generates a q-big subgroup and $\{(0,0)\}$ is not q-big, we get

$$Qx = q\text{-}\lim_{\vec{a}} Q_{\vec{a}}x = 0$$

On the other hand, to show that Q restricted to $\mathcal{H}_1^\perp$ is the identity, we need only consider $x \in \operatorname{im} Q_{n\vec{c}}$, as the span of these vectors is dense. For any such x, we have

$$Q_{n\vec{c}}x = q\text{-}\lim_{\vec{z}} U^{n(c_1 z_2 + c_2 z_1)}x = x,$$

which, as before, can be strengthened to

$$(2.2) \qquad q\text{-}\lim_{\vec{z}} \left\| U^{n(c_1 z_2 + c_2 z_1)}x - x \right\| = 0.$$

Now we need to extend this equality, true for only one vector $n\vec{c}$, to some q-big set of vectors. We leave it for the reader to check that (2.2) actually gives

$$q\text{-}\lim_{\vec{z}} \left\| U^{Nn(c_1 z_2 + c_2 z_1)}x - x \right\| = 0.$$

for any $N \in \mathbb{Z}$. (Hint: Use a telescoping sum.) But the set $\mathbb{Z} \cdot n\vec{c}$ is again a q-big subgroup (since q is idempotent), and so

$$q\text{-}\lim_{\vec{a}} q\text{-}\lim_{\vec{z}} \left\| U^{a_1 z_2 + a_2 z_1}x - x \right\| = 0,$$

which implies $Qx = x$. So Q is indeed an orthogonal projection, with image $\mathcal{H}_1^\perp$.

Dimension 2. Finally, let us treat the really interesting case of a two-dimensional q. We use the same argument as before, only the splitting has to be adjusted a bit; instead of focusing on one specific subgroup (like $\mathbb{Z}\vec{c}$), we shall consider all of them. So let

$$\mathcal{H}_2 = \bigcap_{\vec{a},\vec{b}} \ker Q_{\vec{a}} Q_{\vec{b}} \qquad \text{and} \qquad \mathcal{H}_2^\perp = \overline{\sum_{\vec{a},\vec{b}} \operatorname{im} Q_{\vec{a}} Q_{\vec{b}}}$$

be the two complementary subspaces, where both the intersection and the sum are taken over those $\vec{a}, \vec{b} \in \mathbb{Z}^2$ for which the subgroup $\mathbb{Z}\vec{a} + \mathbb{Z}\vec{b}$ has rank two. Again, it will turn out that Q is orthogonal projection onto $\mathcal{H}_2^\perp$.

To prove that Q fixes every vector in $\mathcal{H}_2^\perp$, we may again limit our attention to elements $x \in \operatorname{im} Q_{\vec{a}} Q_{\vec{b}}$ for two vectors $\vec{a}$ and $\vec{b}$ that span a rank two subgroup. The same argument as before shows that

$$q\text{-}\lim_{\vec{z}} \left\| U^{M(a_1 z_2 + a_2 z_1) + N(b_1 z_2 + b_2 z_1)} x - x \right\| = 0$$

for any $M, N \in \mathbb{Z}$; the group generated by $\vec{a}$ and $\vec{b}$ is q-big (remember that it contains some lattice), and so we have

$$q\text{-}\lim_{\vec{z}} \left\| U^{c_1 z_2 + c_2 z_1} x - x \right\| = 0$$

for q-many $\vec{c} \in \mathbb{Z}^2$. Taking the q-limit over $\vec{c}$ then gives the result, namely that $Qx = x$.

To finish the proof, we have to deal with an arbitrary $x \in \mathcal{H}_2$ and show that $Qx = 0$. What we know is that $Q_{\vec{a}} Q_{\vec{b}} x = 0$ for any two vectors $\vec{a}$ and $\vec{b}$ with a two-dimensional span. This is a lot of information, since there are many such pairs—in fact, for any nonzero $\vec{a} \in \mathbb{Z}^2$, a q-big set of $\vec{b}$ has the required property. For suppose, to the contrary, that q-many vectors $\vec{b}$ could span only a subgroup of rank one together with $\vec{a}$. As $\mathbb{Z}\vec{a} + \mathbb{Z}\vec{b}$ is of rank one if and only if $\vec{b}$ is a multiple of $\vec{a}/g$ (here g is the greatest common divisor of the components of $\vec{a}$), it would follow that $\mathbb{Z}\vec{a}/g$ was a q-big subgroup of $\mathbb{Z}^2$, contradicting our assumption on the dimension of q.

In particular, we know $Q_{\vec{a}} Q_{\vec{b}} x = 0$ for sufficiently many $\vec{a}$ and $\vec{b}$ to conclude that

$$Q^2 x = q\text{-}\lim_{\vec{a}} q\text{-}\lim_{\vec{b}} Q_{\vec{a}} Q_{\vec{b}} x = 0;$$

but now the operator Q is very evidently self-adjoint and so $Qx = 0$ as well. This shows that Q is a projection and ends the proof of Theorem 2.3.

Conclusions. Let us end this section with several remarks concerning the nature of the proof. Firstly, the reader will have observed the balance—crude in the case of one-dimensional p, slightly more subtle for two dimensions—between the two spaces of the splitting. In the first space, $\mathcal{H}_1$ or $\mathcal{H}_2$, where we use the null spaces of projections, we need to intersect a large number of them to make up for the weakness of each individual piece; for each operator, we only know that one particular p-limit is zero, and that amounts to nothing by itself. On the other hand, the orthogonal complements, $\mathcal{H}_1^\perp$ or $\mathcal{H}_2^\perp$, involve image spaces of projections; the knowledge that we gain from each piece is far stronger here, and so we can afford to have this knowledge in only one case.

Secondly, it is clear that the dimension of the ultrafilter is important. It was pointed out before that, although the same concept is lurking around in Theorem 2.1, it is less easily quantified and dealt with there. The passage from $\beta\mathcal{F}$ to $\beta\mathbb{Z}^2$ helps to make it visible, by removing the IP-sets. Moreover, it is of course unnecessary to handle the various dimensions by different arguments; the proof is really the same in all cases. Indeed, in Section 5, when proving the main theorem, the first step will be to adjust the dimension of the surrounding group to make it match that of the ultrafilter p. This is where Lemma 6 will play its part.

Finally, the more general result in the main theorem requires more effort to prove; although the proof is, in essence, the same as the one given here, there are several technical points that need to be dealt with. In particular, the presence of polynomials of higher degree needs special care. The following two sections contain a few tools that will be helpful; all necessary results about polynomials are collected in Section 4.

3. Orthogonal projections and limits

In this section, we prove two simple but useful results about orthogonal projections and limits; these are well-known, of course. The first, which has already been used, gives a condition for an operator to be a projection.

Lemma 9. *A normal operator P on a Hilbert space $\mathcal{H}$ is an orthogonal projection if, and only if, it satisfies $P^2 = P$.*

Proof. Necessity is clear. If P meets the condition, the product $Q = P^*P$ of P and its adjoint also does. The latter is self-adjoint in addition, hence satisfies $\langle Qx, x - Qx \rangle = \langle x, Q(x - Qx) \rangle = 0$ and is therefore an orthogonal projection onto the image space of Q. For $x \in \ker Q$, one has $\left\| Px \right\|^2 = \langle x, Qx \rangle = 0$; for $x \in \operatorname{im} Q$, one has $Px = PQx = PP^*Px = P^*Px = x$. Consequently, $P = Q$ is an orthogonal projection. $\qquad\square$

Our second lemma deals with the question of when certain 'weak' limits in a Hilbert space $\mathcal{H}$ are 'strong' limits and is meant to collect the pieces of reasoning used in the previous section. The whole discussion is somewhat vague but the result is useful, though nearly self-evident. Let 'lim' be an abbreviation for some unspecified p-limit, maybe even a multiple one, and let I be the corresponding index set. So for example, lim might equal $p\text{-}\lim_{\vec{a}} p\text{-}\lim_{\vec{b}}$, with both $\vec{a}$ and $\vec{b}$ ranging over $\mathbb{Z}^2$, in which case the index set I would be $\mathbb{Z}^2 \times \mathbb{Z}^2$.

By what we said in Section 1, the limit $\lim x_i$ is defined, in the weak topology, for every bounded family $(x_i)_{i \in I}$ of points, as any closed ball in $\mathcal{H}$ is weakly compact. $\lim x_i = x$ thus means that for any $y \in \mathcal{H}$,

$$\lim\langle x_i, y \rangle = \langle x, y \rangle.$$

On the other hand, the convergence is called *strong* if

$$\lim\left\| x_i - x \right\| = 0.$$

The norm is weakly lower semi-continuous—if $x = \lim x_i$, then

$$\left\| x \right\| \leq \lim\left\| x_i \right\|.$$

One can also define the notions of weak and strong *operator limits*; in fact, we have already been using these. We say, for example, that T is the weak operator

limit of a family $(T_i)_{i \in I}$ of operators—and write $T = \lim T_i$—if

$$\lim \langle T_i x, y \rangle = \langle T x, y \rangle \qquad \text{(for } x, y \in \mathcal{H}\text{)}.$$

A few simple calculations then give the following result.

Lemma 10. *Let* $\lim$ *and* I *be defined as above.*

(1) *If* $T = \lim U_i$ *is the weak operator limit of a family* $(U_i)_{i \in I}$ *of unitary operators, then* T *is normal. For* $x \in \mathcal{H}$, *one has* $\lim \|Tx - U_i x\| = 0$ *if, and only if,* $\|Tx\| = \|x\|$. *In case* T *is an orthogonal projection, this happens exactly when* $Tx = x$.

(2) *If* $T = \lim P_i$ *is the weak operator limit of a family* $(P_i)_{i \in I}$ *of orthogonal projections, then* T *is self-adjoint. For* $x \in \mathcal{H}$, *one has* $\lim \|Tx - P_i x\| = 0$ *if, and only if,* $\langle Tx, x - Tx \rangle = 0$. *In case* T *is itself an orthogonal projection, this condition is always satisfied.*

The result looks innocent enough, but it will be used frequently.

4. POLYNOMIALS

We shall be using polynomials in several variables for which the following notation seems appropriate. Lower-case Roman letters with arrows will usually denote n-dimensional vectors, e.g. $\vec{z} = (z_1, \ldots, z_n)$. We shall be speaking of polynomials *in the variable* $\vec{z}$, meaning really polynomials in the n variables $z_1, \ldots, z_n$. The *degree* of such a polynomial will be its total degree. We shall also consider polynomials in several multi-dimensional variables: $f(\vec{a}, \vec{c})$, say, would be a polynomial in both sets of variables; the *degree in* $\vec{a}$ is the total degree of f as a polynomial in $a_1, \ldots, a_n$, and so on.

If G is any Abelian group, we shall let $G[\vec{z}_1, \ldots, \vec{z}_s]$ stand for the additive group of polynomials in $\vec{z}_1, \ldots, \vec{z}_s$ with coefficients in G; we shall ignore the multiplicative structure. For the subgroup of those polynomials in $\mathbb{Q}[\vec{z}_1, \ldots, \vec{z}_r]$ that produce integer values for integer arguments, we shall write $\mathrm{Int}[\vec{z}_1, \ldots, \vec{z}_s]$.

In the one-dimensional case, $\mathrm{Int}[x]$ consists of all polynomials $f \in \mathbb{Q}[x]$ with $f(\mathbb{Z}) \subseteq \mathbb{Z}$. It is a free group with basis consisting of the polynomials

$$\binom{x}{i} = \frac{x(x-1) \cdots (x - i + 1)}{i!} \qquad \text{(for } i \geq 0\text{)}.$$

Indeed, if f is any polynomial in $\mathrm{Int}[x]$ and m its degree, one may find $m+1$ integers $a_0, \ldots, a_m$ such that

$$f(x) = \sum_{i=0}^{m} a_i \binom{x}{i},$$

by evaluating successively at $x = 0, 1, \ldots, m$, and solving the resulting system of equations. For any number $d \geq 0$, the polynomials of degree at most d form a free subgroup of rank $d + 1$.

The same argument, applied inductively, proves the following.

Lemma 11. $\mathrm{Int}[\vec{z}_1, \ldots, \vec{z}_s]$ *is always a free group; for any integer* $d \geq 0$, *the polynomials of total degree at most* d *constitute a free subgroup of finite rank, and so do the polynomials of degree at most* d *in each variable.*

We now introduce one more useful notion. In the example in Section 2, when dealing with the polynomial $f(x,y) = xy$, we found it useful to form the new polynomial

$$f(a+x, b+y) - f(a,b) - f(x,y),$$

essentially because its degree in (x,y) was lower. An appropriate generalization is as follows. Given a polynomial $f(\vec{z}) \in G[\vec{z}]$ and an integer $s \geq 1$, we recursively define a new polynomial $\Delta^s f(\vec{z}_1, \ldots, \vec{z}_s)$, by letting $\Delta^1 f(\vec{z}_1) = f(\vec{z}_1)$, and

(4.1) $\Delta^{s+1} f(\vec{z}_1, \ldots, \vec{z}_{s+1}) =$

$$\Delta^s f(\vec{z}_1, \ldots, \vec{z}_s + \vec{z}_{s+1}) - \Delta^s f(\vec{z}_1, \ldots, \vec{z}_s) - \Delta^s f(\vec{z}_1, \ldots, \vec{z}_{s+1}).$$

Of course, Δ^s can be described explicitly as

$$\Delta^s f(\vec{z}_1, \ldots, \vec{z}_s) = \sum_{\emptyset \neq \alpha \subseteq \{1, \ldots, s\}} (-1)^{s-|\alpha|} \cdot f\left(\sum_{i \in \alpha} \vec{z}_i\right),$$

and the symmetry in all arguments is more apparent from this description.

Let us investigate some properties of Δ^s. First, we have the following easy lemma.

Lemma 12. *The polynomial $\Delta^2 f(\vec{a}, \vec{z}) = f(\vec{a} + \vec{z}) - f(\vec{a}) - f(\vec{z})$ is of lower degree in each variable than $f(\vec{z})$ itself, whenever $f(\vec{z}) \in G[\vec{z}]$ is nonzero.*

Now let $f(\vec{z})$ be of degree $d \geq 1$. Since $\Delta^s f(\vec{z}_1, \ldots, \vec{z}_s)$ is symmetric in its s arguments, the lemma—together with the relations (4.1)—immediately shows that its degree in any variable can be at most $(d+1-s)$. It follows that $\Delta^{d+1} f(\vec{z}_1, \ldots, \vec{z}_{d+1})$ is a constant, with value

$$\Delta^{d+1} f(0, \ldots, 0) = \sum_{k=1}^{d+1} (-1)^{d+1-k} \binom{d+1}{k} f(0) = (-1)^d f(0).$$

If f happens to satisfy $f(0) = 0$, one has $\Delta^{d+1} f(\vec{z}_1, \ldots, \vec{z}_{d+1}) = 0$. For reasons of symmetry, $\Delta^d f(\vec{z}_1, \ldots, \vec{z}_d)$ is then linear in each of its d arguments.

We have shown the following.

Lemma 13. *For any polynomial $f(\vec{z}) \in G[\vec{z}]$ of degree $d \geq 1$, one has the relation*

$$\Delta^{d+1} f(\vec{z}_1, \ldots, \vec{z}_{d+1}) = (-1)^d f(0).$$

If $f(0) = 0$, then $\Delta^d f(\vec{z}_1, \ldots, \vec{z}_d)$ is a linear function of each argument.

A third lemma deals with the case of homogeneous f.

Lemma 14. *Let $f(\vec{z}) \in G[\vec{z}]$ be a homogeneous polynomial of degree $d \geq 1$. Then*

$$\Delta^d f(\vec{a}, \ldots, \vec{a}) = d! f(\vec{a}).$$

Proof. Using homogeneity, we have

$$\Delta^s f(\vec{a}, \ldots, \vec{a}) = \sum_{k=1}^s (-1)^{s-k} \binom{s}{k} f(k\vec{a}) = \sum_{k=1}^s (-1)^{s-k} \binom{s}{k} k^d \cdot f(\vec{a}).$$

We obviously have to evaluate sums of the form

$$C(s,m) = \sum_{k=1}^s (-1)^{s-k} \binom{s}{k} k^m \qquad \text{(for } s \geq 1, m \geq 0\text{)};$$

in particular, $C(d, d) = d!$ is what we need to show. From the previous lemma, we already know that $C(s, m) = 0$ whenever $s > m$. Now we compute

$$
\begin{aligned}
C(m+1, m+1) &= \sum_{k=1}^{m+1} (-1)^{m+1-k} \binom{m+1}{k} k^{m+1} \\
&= \sum_{k=1}^{m+1} (-1)^{m-(k-1)} (m+1) \binom{m}{k-1} k^m \\
&= (m+1) \sum_{l=0}^{m} (-1)^{m-l} \binom{m}{l} (l+1)^m \\
&= (m+1) \sum_{l=0}^{m} (-1)^{m-l} \binom{m}{l} \left(1 + \sum_{i=1}^{m} \binom{m}{i} l^i \right) \\
&= (m+1) \sum_{l=0}^{m} (-1)^{m-l} \binom{m}{l} + (m+1) \sum_{i=1}^{m} \binom{m}{i} C(m, i) \\
&= (1-1)^m + (m+1) C(m, m) = (m+1) C(m, m),
\end{aligned}
$$

and together with $C(1, 1) = 1$ this proves the lemma by induction. $\qquad \square$

We will now use the previous results to establish an important technical lemma; it is essential for the proof of the main theorem in Section 5. Note that it introduces a feature not present in the example of Section 2, where we had to deal with polynomials of no more than first degree. It does, however, fit in with the general philosophy behind the argument—there is one situation in the proof where one has to make a lot from apparently nothing, meaning where one has to create useful p-big sets from useless ones, and the following lemma does just that.

Lemma 15. *Suppose that $p \in \beta \mathbb{Z}^n$ is an n-dimensional idempotent. Let G be an Abelian group and let $v(\vec{z}) \in G[\vec{z}]$ be a polynomial in $\vec{z} = (z_1, \ldots, z_n)$. Fix a subgroup $V \subseteq G$. If the set*

$$
B = \left\{ \vec{b} \in \mathbb{Z}^n \,\middle|\, N \cdot v(\vec{b}) \in V \text{ for some } N \neq 0 \right\}
$$

is p-big, then so is the set

$$
A = \left\{ \vec{a} \in \mathbb{Z}^n \,\middle|\, v(\vec{a}) - v(0) \in V \right\}.
$$

Proof. The idea of the proof is simple: Whenever B contains $\mathrm{IP}\left(\vec{b}_1, \ldots, \vec{b}_s\right)$, there is some integer $N \neq 0$ such that $N \cdot \Delta^s v(\vec{b}_1, \ldots, \vec{b}_s) \in V$. Using this and the previous results, we can extract from v its homogeneous parts of different degrees, and show that they are each contained in V for p-many $\vec{z}$. Proceeding stepwise, we shall prove two things:

    I. Without loss of generality, it may be assumed that $v(0) = 0$.
    II. The set $A = \left\{ \vec{a} \,\middle|\, v(\vec{a}) \in V \right\}$ is p-big.

The details are as follows.

I. Let d be the degree of $v(\vec{z})$. The set B is a member of the idempotent p, and by Lemma 1, we may select $d+1$ elements $\vec{b}_1, \ldots, \vec{b}_{d+1}$ in B with

$$
F = \mathrm{IP}\left(\vec{b}_1, \ldots, \vec{b}_{d+1}\right) \subseteq B.
$$

For each $\vec{b}$ in this finite IP-set, there exists some $N \neq 0$ such that $N \cdot v(\vec{b}) \in V$; if we let N_1 be the greatest common divisor of these numbers, we guarantee that $N_1 \cdot v(\vec{b}) \in V$ for every $\vec{b} \in F$. By Lemma 13, we now have

$$N_1 \cdot v(0) = N_1 \cdot (-1)^d \Delta^{d+1} v(\vec{b}_1, \ldots, \vec{b}_{d+1}) \in V.$$

If $\vec{b}$ is any element of B and $N \cdot v(\vec{b}) \in V$, the expression $N_1 N \cdot \left(v(\vec{b}) - v(0)\right)$ is also an element of V; this means that the set

$$\left\{\vec{b} \,\middle|\, N \cdot \left(v(\vec{b}) - v(0)\right) \in V \text{ for some } N \neq 0\right\}$$

is equally p-big. We may therefore replace $v(\vec{z})$ by $v(\vec{z}) - v(0)$ and assume $v(0) = 0$.

II. We decompose v into homogenous polynomials,

$$v(\vec{z}) = \sum_{i=1}^{d} h_i(\vec{z}),$$

say, with $h_i(\vec{z})$ homogeneous of degree i. For $1 \leq i \leq d$, let

$$A_i = \left\{\vec{a} \,\middle|\, h_i(\vec{a}) \in V\right\}.$$

We shall argue that $A_d \in p$; once this is known, the same reasoning applies to the polynomial $v(\vec{z}) - h_d(\vec{z})$ where it gives $A_{d-1} \in p$, and so on, until one has $A_i \in p$ for each i. The result follows because A contains the intersection of all A_i.

To show that A_d is p-big, we use a similar—but more careful—approach as before. By Lemma 1, there are actually p-many $\vec{b}_1, \ldots, \vec{b}_d \in B$ with $\mathrm{IP}(\vec{b}_1, \ldots, \vec{b}_d) \subseteq B$, and again, to each choice we may find some $N_1 \neq 0$ such that $N_1 \cdot \Delta^d v(\vec{b}_1, \ldots, \vec{b}_d)$ is an element of V. When selecting only $\vec{b}_1, \ldots, \vec{b}_{d-1}$ from these, the set

$$B_d = \left\{\vec{b} \,\middle|\, N \cdot \Delta^d v(\vec{b}_1, \ldots, \vec{b}_{d-1}, \vec{b}) \in V \text{ for some } N \neq 0\right\}$$

is then always p-big by construction.

B_d is also a subgroup of $\mathbb{Z}^n$, the polynomial $\Delta^d v(\vec{z}_1, \ldots, \vec{z}_d)$ being linear in each variable (see Lemma 13). Since p is n-dimensional, this subgroup has to have rank n and has to contain a set of the form $L \cdot \mathbb{Z}^n$ for some nonzero L. For each $\vec{b} \in \mathbb{Z}^n$, one gets

$$NL \cdot \Delta^d v(\vec{b}_1, \ldots, \vec{b}_{d-1}, \vec{b}) = N \cdot \Delta^d v(\vec{b}_1, \ldots, \vec{b}_{d-1}, L \cdot \vec{b}) \in V$$

for some $N \neq 0$, and so we conclude that $B_d = \mathbb{Z}^n$.

Next, consider the set

$$B_{d-1} = \left\{\vec{b} \,\middle|\, \text{for some } N \neq 0, \ N \cdot \Delta^d v(\vec{b}_1, \ldots, \vec{b}_{d-2}, \vec{b}, \vec{a}_d) \in V \text{ for all } \vec{a}_d\right\};$$

by the above, it is p-big, and a repetition of the argument shows that $B_{d-1} = \mathbb{Z}^n$, too. Continuing in this way, we eventually find a nonzero integer N_2 such that

$$N_2 \cdot \Delta^d v(\vec{a}_1, \ldots, \vec{a}_d) \in V$$

for any choice of $\vec{a}_1, \ldots, \vec{a}_d \in \mathbb{Z}^n$.

By Lemma 12, we have

$$\Delta^d h_d(\vec{a}_1, \ldots, \vec{a}_d) = \Delta^d v(\vec{a}_1, \ldots, \vec{a}_d),$$

because all terms of degree less than d disappear. Finally, using Lemma 14 we get

$$N_2 d! \cdot h_d(\vec{a}) = N_2 \cdot \Delta^d v(\vec{a}, \ldots, \vec{a}) \in V$$

for all $\vec{a}$, and hence $h_d(\vec{a}) \in V$ whenever $\vec{a} \in N_2 d! \cdot \mathbb{Z}^n$. The latter set is p-big and so $A_d \in p$. This ends the proof of the second part, and establishes the lemma. $\square$

5. Statement and proof of the main results

After all the preliminary work in the previous two sections, we are now ready to state and prove the main result. The notation is somewhat heavy, but this generality is needed because of the inductive nature of the proof.

Theorem 5.1. *For $j = 1, \ldots, s$, let $p_j \in \beta\mathbb{Z}^{n_j}$ be an idempotent, and let $U_1, \ldots, U_m$ be commuting unitary operators on a Hilbert space $\mathcal{H}$. Given any m polynomials $f_1, \ldots, f_m \in \mathrm{Int}[\vec{z}_1, \ldots, \vec{z}_s]$—with $\vec{z}_j$ of dimension n_j—satisfying $f_i(0) = 0$ for all $i = 1, \ldots, m$, define an operator P on $\mathcal{H}$ by*

$$P = p_1\text{-}\lim_{\vec{z}_1} \cdots p_s\text{-}\lim_{\vec{z}_s} \prod_{i=1}^{m} U_i^{f_i(\vec{z}_1, \ldots, \vec{z}_s)}.$$

Then P is always an orthogonal projection. Any two operators defined in this way commute.

Proof. We will suppose that all n_j are equal to some n and all p_j are equal to some p, in order to simplify notation. The argument need not be changed in any way to accommodate the more general situation—because of the inductive character of the proof, we find ourselves working on no more than one p-limit at a time anyway.

Let us first observe that the last part of the statement—commutativity of different projections—is obviously true, for all operators generated for various selections of polynomials are certain weak limits of commuting unitary operators. We may therefore assume commutativity wherever needed.

The remainder of the proof is essentially by induction on the number s of p-limits taken, but there are some complications involving the case $s = 1$. In fact, different arguments are needed for $s = 1$ and for $s \geq 2$, since the outmost p-limit is one of unitary operators in the former situation, but one of projections in the latter. To make the induction be more transparent, we shall use the abbreviation (s, d) when referring to the statement of the theorem for a certain value of s and all possible choices of polynomials f_i of degree at most d in any of their variables $\vec{z}_j$.

The proof will be divided into several steps, the second and third inductive in nature:

    I. We argue that the ultrafilter p may be assumed to be n-dimensional, without loss of generality.

    II. We establish the case $(1, d)$, that is, we show that the operator

$$P = p\text{-}\lim_{\vec{z}} \prod_{i=1}^{m} U_i^{f_i(\vec{z})},$$

with f_i of degree at most d, is an orthogonal projection, assuming the statement of the theorem in the two cases $(1, d-1)$ and $(2, d-1)$. Specifically, we need to assume that the operators

$$Q_{\vec{a}} = p\text{-}\lim_{\vec{z}} \prod_{i=1}^{m} U_i^{f_i(\vec{a}+\vec{z})-f_i(\vec{a})-f_i(\vec{z})}$$

and

$$Q = p\text{-}\lim_{\vec{a}} Q_{\vec{a}}$$

are orthogonal projections.

III. For $s \geq 2$, we derive (s, d) from $(s-1, d)$ and $(1, d)$. Essentially, we introduce the new polynomials

$$f_i'(\vec{a}, \vec{z}_2, \ldots, \vec{z}_s) = f_i(\vec{a}, \vec{z}_2, \ldots, \vec{z}_s) - f_i(\vec{a}, 0, \ldots, 0)$$

and assume that for each $\vec{a} \in \mathbb{Z}^n$, the operator

$$P_{\vec{a}} = p\text{-}\lim_{\vec{z}_2} \cdots p\text{-}\lim_{\vec{z}_s} \prod_{i=1}^{m} U_i^{f_i'(\vec{a}, \vec{z}_2, \ldots, \vec{z}_s)}$$

is an orthogonal projection. We then construct a suitable splitting of the underlying Hilbert space $\mathcal{H}$.

IV. Using the splitting introduced in the previous part, we show that P is an orthogonal projection.

Once we have established all of the previous, our work will be done. For the statement of the theorem is definitely true in the case $(1, 0)$—if all f_i equal zero, P is just the identity—and then II and III suffice to prove the entire theorem by induction. Let us now take a detailed look at the four steps.

I. Using Lemma 6, we begin by adjusting the situation to make sure that the dimension of p is equal to the rank of the group. Of course, this will change the polynomials under consideration; but since their degrees are not increased, it does not affect the proof. To write this down precisely is somewhat cumbersome; so let us look at the case of just one operator U and one polynomial $f(\vec{z})$ to see what happens. We shall use $\vec{z}$ for an n-dimensional and $\vec{w}$ for an s-dimensional variable. If $s = \dim p$ and ϕ are as in Lemma 6, then $q = \phi_*(p)$ is idempotent and s-dimensional. By Lemma 3,

$$(5.1) \qquad p\text{-}\lim_{\vec{z}} U^{f(\vec{z})} = q\text{-}\lim_{\vec{w}} U^{f(\phi^{-1}(\vec{w}))}.$$

But $g(\vec{w}) = f(\phi^{-1}(\vec{w})) \in \mathrm{Int}[\vec{w}]$ satisfies $g(0) = 0$, and is of degree no larger than that of f; thus a proof that the right-hand operator in (5.1) is a projection gives the result for the left-hand one, too. The same is true in the general setting of the theorem, though somewhat unpleasant to write down in detail.

In any case, we shall assume from now on that $\dim p = n$. Lemma 15 is then applicable; it will make its entry in the third step of the proof.

II. As stated above, we shall now assume that both Q and all the $Q_{\vec{a}}$ are orthogonal projections; this is permissible because each polynomial

$$f_i(\vec{a} + \vec{z}) - f_i(\vec{a}) - f_i(\vec{z})$$

has degree at most $(d - 1)$ in $\vec{a}$ and $\vec{z}$ (see Lemma 12). The projection Q induces a splitting $\mathcal{H} = \ker Q \oplus \mathrm{im}\, Q$ of the underlying Hilbert space; we shall use it to conclude that $P^2 x = P x$ for all $x \in \mathcal{H}$.

First, consider $x \in \ker Q$. Since $Q x = p\text{-}\lim_{\vec{a}} Q_{\vec{a}} x$, Lemma 10 implies that the convergence is strong,

$$p\text{-}\lim_{\vec{a}} \big\| Q_{\vec{a}} x \big\| = 0.$$

We now use van der Corput's trick to get $Px = 0$, the condition

$$p\text{-}\lim_{\vec{a}} p\text{-}\lim_{\vec{z}} \Big\langle \prod_{i=1}^{m} U_i^{f_i(\vec{a}+\vec{z})} x, \prod_{i=1}^{m} U_i^{f_i(\vec{z})} x \Big\rangle$$

$$= p\text{-}\lim_{\vec{a}} p\text{-}\lim_{\vec{z}} \Big\langle \prod_{i=1}^{m} U_i^{f_i(\vec{a}+\vec{z}) - f_i(\vec{a}) - f_i(\vec{z})} x, \prod_{i=1}^{m} U_i^{-f_i(\vec{a})} x \Big\rangle$$

$$= p\text{-}\lim_{\vec{a}} \Big\langle Q_{\vec{a}} x, \prod_{i=1}^{m} U_i^{-f_i(\vec{a})} x \Big\rangle = 0$$

in Lemma 4 being fulfilled. A fortiori, $P^2 x = Px$.

Second, consider an arbitrary $x \in \operatorname{im} Q$, which then satisfies $Qx = x$. We again get strong convergence from Lemma 10, so that

$$p\text{-}\lim_{\vec{a}} p\text{-}\lim_{\vec{z}} \Big\| \prod_{i=1}^{m} U_i^{f_i(\vec{a}+\vec{z})} x - \prod_{i=1}^{m} U_i^{f_i(\vec{a}) + f_i(\vec{z})} x \Big\| = 0.$$

But then

$$P^2 x = p\text{-}\lim_{\vec{a}} p\text{-}\lim_{\vec{z}} \prod_{i=1}^{m} U_i^{f_i(\vec{a}) + f_i(\vec{z})} x = p\text{-}\lim_{\vec{a}} p\text{-}\lim_{\vec{z}} \prod_{i=1}^{m} U_i^{f_i(\vec{a}+\vec{z})} x$$

$$= p\text{-}\lim_{\vec{z}} \prod_{i=1}^{m} U_i^{f_i(\vec{z})} x = Px.$$

We now have $P^2 = P$; obviously, P is normal, and the result—that P is an orthogonal projection—follows from Lemma 9.

III. This is the most interesting part of the argument. We start from the inductive assumption that each $P_{\vec{a}}$ is an orthogonal projection, and aim for a useful splitting of the space $\mathcal{H}$, depending on the projections $P_{\vec{a}}$ and the polynomials

$$f_i'(\vec{a}, \vec{z}_2, \ldots, \vec{z}_s) = f_i(\vec{a}, \vec{z}_2, \ldots, \vec{z}_s) - f_i(\vec{a}, 0, \ldots, 0).$$

Note that $f_i'(\vec{a}, 0, \ldots, 0) = 0$ for any $\vec{a}$, which means that the new polynomials f_i' still satisfy the conditions of the theorem, while having fewer variables.

After the introduction of $g_i(\vec{a}) = f_i(\vec{a}, 0, \ldots, 0)$, the operator P is then given by the limit

$$P = p\text{-}\lim_{\vec{a}} \Big(\prod_{i=1}^{m} U_i^{g_i(\vec{a})} \Big) P_{\vec{a}}.$$

We shall let $F \subseteq \operatorname{Int}[\vec{z}_2, \ldots, \vec{z}_s]$ denote the set of polynomials of degree at most d in each variable; F is a free group of finite rank by Lemma 11. The product F^m is also free, as are all of its subgroups, and for any $\vec{a} \in \mathbb{Z}^n$, the vector

$$v(\vec{a}) = \big(\tilde{f}_1(\vec{a}, \vec{z}_2, \ldots, \vec{z}_s), \ldots, \tilde{f}_m(\vec{a}, \vec{z}_2, \ldots, \vec{z}_s) \big)$$

is an element of F^m. We introduce the notation $V(\vec{a}_1, \ldots, \vec{a}_r)$ for the subgroup of F^m generated by the vectors $v(\vec{a}_1), \ldots, v(\vec{a}_r)$.

The *crucial idea* is to let $r \geq 0$ be the maximal integer for which

$$\big\{ \vec{a}_1 \,\big|\, \cdots \big\{ \vec{a}_r \,\big|\, V(\vec{a}_1, \ldots, \vec{a}_r) \text{ has rank } r \big\} \in p \cdots \big\} \in p.$$

Such an r has to exist, because we are working inside a fixed group of finite rank; if not even $\big\{ \vec{a}_1 \,\big|\, V(\vec{a}_1) \text{ has rank one} \big\}$ is in p, we set $r = 0$ to keep the notation

consistent. Whenever $\vec{a}_1, \ldots, \vec{a}_r$ are taken, in the correct order, from these nested sets, the rank of the group $V(\vec{a}_1, \ldots, \vec{a}_r)$ is r.

With r being defined in that manner, one also has

$$\left\{ \vec{a}_1 \,\middle|\, \cdots \left\{ \vec{a}_{r+1} \,\middle|\, V(\vec{a}_1, \ldots, \vec{a}_{r+1}) \text{ has rank less than } r \right\} \in p \cdots \right\} \in p.$$

Intersecting with the previous set and using that p is an ultrafilter, we obtain

$$\left\{ \vec{a}_1 \,\middle|\, \cdots \left\{ \vec{a}_r \,\middle|\, V(\vec{a}_1, \ldots, \vec{a}_r) \text{ has rank } r \text{ and} \right.\right.$$
$$\left.\left. \left\{ \vec{b} \,\middle|\, V(\vec{a}_1, \ldots, \vec{a}_r, \vec{b}) \text{ also has rank } r \right\} \in p \right\} \in p \cdots \right\} \in p.$$

But if $V(\vec{a}_1, \ldots, \vec{a}_r)$ and $V(\vec{a}_1, \ldots, \vec{a}_r, \vec{b})$ both have rank r, it means that some nonzero multiple of $\vec{b}$ has to lie in the first group. We can therefore conclude from the previous line that

(5.2)
$$\left\{ \vec{a}_1 \,\middle|\, \cdots \left\{ \vec{a}_r \,\middle|\, V = V(\vec{a}_1, \ldots, \vec{a}_r) \text{ has rank } r \text{ and} \right.\right.$$
$$\left.\left. \left\{ \vec{b} \,\middle|\, N \cdot v(\vec{b}) \in V \text{ for some } N \neq 0 \right\} \in p \right\} \in p \cdots \right\} \in p.$$

Finally, let $\mathcal{A}$ denote the set of r-tuples $(\vec{a}_1, \ldots, \vec{a}_r)$, taken in the right order from the nested sets in (5.2); for any one of them, the group $V = V(\vec{a}_1, \ldots, \vec{a}_r)$ has rank r and the set $\left\{ \vec{b} \,\middle|\, N \cdot v(\vec{b}) \in V \text{ for some } N \neq 0 \right\}$ is p-big.

We have now arrived at our destination—we shall use the splitting $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_1^{\perp}$, where

$$\mathcal{H}_1 = \bigcap_{\mathcal{A}} \ker P_{\vec{a}_1} \cdots P_{\vec{a}_r} \qquad \text{and} \qquad \mathcal{H}_1^{\perp} = \overline{\sum_{\mathcal{A}} \operatorname{im} P_{\vec{a}_1} \cdots P_{\vec{a}_r}}.$$

This ends the third step; the proof, based on this splitting, that P is an orthogonal projection is contained in the remaining part of the proof.

IV. It remains to prove that the operator P really is a projection. Because of the splitting from III, we have two subspaces to consider. Let us begin with the one that is easier to handle, and show that P is zero on $\mathcal{H}_1$. If $x \in \mathcal{H}_1$, we have $P_{\vec{a}_1} \cdots P_{\vec{a}_r} x = 0$ for p-many $\vec{a}_1, \ldots, \vec{a}_r$, thus

$$P^r x = p\text{-}\lim_{\vec{a}_1} \cdots p\text{-}\lim_{\vec{a}_r} \left(\prod_{i=1}^{m} U_i^{g_i(\vec{a}_1) + \cdots + g_i(\vec{a}_r)} \right) P_{\vec{a}_1} \cdots P_{\vec{a}_r} x = 0.$$

As P is self-adjoint, $P^r x = 0$ quickly leads to $Px = 0$.

The complementary subspace $\mathcal{H}_1^{\perp}$, on the other hand, requires more attention. Here, we shall show that P is equal to another projection P', to be defined below, and constructed with the help of the inductive assumptions. So suppose that $x \in \operatorname{im} P_{\vec{a}_1} \cdots P_{\vec{a}_r}$ for a certain tuple $(\vec{a}_1, \ldots, \vec{a}_r) \in \mathcal{A}$; we shall reason that $Px = P'x$, which, by the usual density argument, is sufficient for equality on all of $\mathcal{H}_1^{\perp}$.

Since x lies in the image of the product $P_{\vec{a}_1} \cdots P_{\vec{a}_r}$, we get $P_{\vec{a}_k} x = x$ for each $k = 1, \ldots, r$. Apply Lemma 10 to get strong convergence, in the form

$$p\text{-}\lim_{\vec{z}_2} \cdots p\text{-}\lim_{\vec{z}_s} \left\| \prod_{i=1}^{m} U_i^{f_i'(\vec{a}_k, \vec{z}_2, \ldots, \vec{z}_s)} x - x \right\| = 0.$$

One easily derives that for any r integers $N_1, \ldots, N_r$,

$$p\text{-}\lim_{\vec{z}_2} \cdots p\text{-}\lim_{\vec{z}_s} \left\| \prod_{i=1}^{m} U_i^{N_1 f_i'(\vec{a}_1, \vec{z}_2, \ldots, \vec{z}_s) + \cdots + N_r f_i'(\vec{a}_r, \vec{z}_2, \ldots, \vec{z}_s)} x - x \right\| = 0.$$

The vectors $v(\vec{a}_k)$ span the group $V = V(\vec{a}_1, \ldots, \vec{a}_r)$, and so we can conclude that the equality

$$(5.3) \qquad p\text{-}\lim_{\vec{z}_2} \cdots p\text{-}\lim_{\vec{z}_s} \left\| \prod_{i=1}^{m} U_i^{h_i(\vec{z}_2, \ldots, \vec{z}_s)} x - x \right\| = 0$$

is true for any element $(h_1, \ldots, h_m) \in V$.

We already saw, after (5.2), that the set

$$\left\{ \vec{b} \,\middle|\, N \cdot v(\vec{b}) \in V \text{ for some } N \neq 0 \right\};$$

is p-big, because of how the integer r was chosen. As a consequence of Lemma 15, which applies because $\dim p = n$, the set

$$\left\{ \vec{c} \,\middle|\, v(\vec{c}) - v(0) \in V \right\}$$

is now also p-big. Together with (5.3), this gives

$$p\text{-}\lim_{\vec{z}_2} \cdots p\text{-}\lim_{\vec{z}_s} \left\| \prod_{i=1}^{m} U_i^{f_i'(\vec{c}, \vec{z}_2, \ldots, \vec{z}_s) - f_i'(0, \vec{z}_2, \ldots, \vec{z}_s)} x - x \right\| = 0$$

for p-many $\vec{c} \in \mathbb{Z}^n$, hence

$$p\text{-}\lim_{\vec{c}} p\text{-}\lim_{\vec{z}_2} \cdots p\text{-}\lim_{\vec{z}_s} \left\| \prod_{i=1}^{m} U_i^{f_i'(\vec{c}, \vec{z}_2, \ldots, \vec{z}_s)} x - \prod_{i=1}^{m} U_i^{f_i(0, \vec{z}_2, \ldots, \vec{z}_s)} x \right\| = 0.$$

The following computation now ends the proof of the fourth and last step:

$$Px = p\text{-}\lim_{\vec{c}} \left(\prod_{i=1}^{m} U_i^{g_i(\vec{c})} \right) P_{\vec{c}} x =$$

$$= p\text{-}\lim_{\vec{c}} \left(\prod_{i=1}^{m} U_i^{g_i(\vec{c})} \right) p\text{-}\lim_{\vec{z}_2} \cdots p\text{-}\lim_{\vec{z}_s} \prod_{i=1}^{m} U_i^{f_i'(\vec{c}, \vec{z}_2, \ldots, \vec{z}_s)} x$$

$$= \left(p\text{-}\lim_{\vec{c}} \prod_{i=1}^{m} U_i^{g_i(\vec{c})} \right) \left(p\text{-}\lim_{\vec{z}_2} \cdots p\text{-}\lim_{\vec{z}_s} \prod_{i=1}^{m} U_i^{f_i(0, \vec{z}_2, \ldots, \vec{z}_s)} \right) x$$

By the hypotheses $(1, d)$ and $(s-1, d)$, both bracketed expressions define orthogonal projections, which moreover commute with each other. Let P' be the projection operator defined as their product; then we have shown that $Px = P'x$, for all $x \in \operatorname{im} P_{\vec{a}_1} \cdots P_{\vec{a}_r}$. This proves that P is also an orthogonal projection when restricted to $\mathcal{H}_1^{\perp}$, and thus completes the proof. $\qquad \square$

Comparison with the original. Since the main purpose of this paper is to re-prove the result of [1] using ultrafilters, it may be worthwhile to compare the proof there with the one just given. The overall argument is the same—a proof by induction, relying on a splitting defined in terms of certain groups, which depend on the data (the given IP-sets in [1], the idempotent ultrafilter here). To make the inductive step work out, we have to allow for multiple p-limits, necessitating a more intricate argument; our proof is consequently slightly longer than the original one.

We should point out that the third step of the proof uses the same splitting as that in [1]. The notion of dimension, mentioned before, is also apparent in the original paper, and has to be dealt with in much the same way. While we use Lemma 6 for this purpose, Bergelson, Furstenberg, and McCutcheon rely on the Milliken-Taylor theorem to handle the different possible dimensions in a unified

manner. Several other auxiliary results, proved or quoted in the other paper, also occur at some point in our proof.

Lastly, IP-limits have been replaced by limits along ultrafilters, which means that no subsequences (or more strictly sub-IP-rings) have to be chosen to get convergence. This adds much convenience to the argument.

An IP-version. As in the example in Section 2, we can derive from the previous theorem a version with IP-sets. It will be seen from the statement that it may sometimes be advantageous not to give the most general version of a result.

Theorem 5.2. *For $j = 1, \ldots, s$, let $q_j \in \beta\mathcal{F}$ be an uncongested idempotent, and let $U_1, \ldots, U_m$ be commuting unitary operators on a Hilbert space $\mathcal{H}$. Given m polynomials $f_1, \ldots, f_m \in \mathrm{Int}[\vec{z}_1, \ldots, \vec{z}_s]$—with $\vec{z}_j$ of dimension n_j—satisfying $f_i(0) = 0$ for all $i = 1, \ldots, m$, and given additionally IP-sets $n_\bullet^{i,j}$, with indices $j = 1, \ldots, s$ and $i = 1, \ldots, n_j$, define an operator P on $\mathcal{H}$ by*

$$P = q_1\text{-}\lim_{\vec{\alpha}_1} \cdots q_s\text{-}\lim_{\vec{\alpha}_s} \prod_{i=1}^{m} U_i^{f_i(n_{\vec{\alpha}_1}^{i,1}, \ldots, n_{\vec{\alpha}_1}^{i,n_1}, \ldots, n_{\vec{\alpha}_s}^{i,1}, \ldots, n_{\vec{\alpha}_s}^{i,n_s})}$$

Then P is always an orthogonal projection. Any two of the operators defined in this way commute.

Proof. We use Lemma 2 and define the following maps. For each $j = 1, \ldots, s$, let

$$\phi_j \colon \mathcal{F} \to \mathbb{Z}[\vec{z}_j], \qquad \phi_j(\alpha) = \left(n_\alpha^{i,1}, \ldots, n_\alpha^{i,n_j}\right),$$

and introduce new ultrafilters

$$p_j = \phi_{j*}(q_j) \in \beta\mathbb{Z}^{n_j}.$$

Since the original ultrafilters were uncongested, all p_j are idempotents by virtue of Lemma 8, and we obtain

$$P = p_1\text{-}\lim_{\vec{z}_1} \cdots p_s\text{-}\lim_{\vec{z}_s} \prod_{i=1}^{m} U_i^{f_i(\vec{z}_1, \ldots, \vec{z}_s)}$$

from Lemma 2. The result now follows from the previous theorem. $\qquad\square$

6. Consequences

From the two theorems in the previous section, we can now derive several other results. In order to simplify the statements, we shall only consider single p-limits. Let us begin by showing why it is useful that the weak operator limits we considered are orthogonal projections.

Theorem 6.1. *Let $(X, \mathcal{B}, \mu)$ be a probability measure space, and let $A \subseteq X$ be a measurable set. Let $T_1, \ldots, T_m$ be commuting invertible measure-preserving transformations on X. Furthermore, assume that polynomials $f_1, \ldots, f_m \in \mathrm{Int}[\vec{z}]$ are given, where $\vec{z} = (z_1, \ldots, z_n)$, such that $f_i(0) = 0$ for all $i = 1, \ldots, m$.*

(1) For any idempotent $p \in \beta\mathbb{Z}^n$, one has

$$p\text{-}\lim_{\vec{z}} \mu\left(A \cap \left(\prod_{i=1}^{m} T_i^{f_i(\vec{z})}\right)^{-1} A\right) \geq \mu(A)^2.$$

(2) *For any uncongested idempotent $q \in \beta\mathcal{F}$ and for any collection of IP-sets $n_{\bullet}^{i,j}$, with indices $i = 1, \ldots, m$ and $j = 1, \ldots, n$, one has*

$$q\text{-}\lim_{\alpha} \mu\left(A \cap \left(\prod_{i=1}^{m} T_i^{f_i(n_{\alpha}^{i,1}, \ldots, n_{\alpha}^{i,n})}\right)^{-1} A \right) \geq \mu(A)^2.$$

Proof. We shall only prove the first statement; the argument will likely be familiar to the reader anyway. On the Hilbert space $\mathcal{H} = L^2(X, \mu)$, introduce m commuting unitary operators $U_1, \ldots, U_m$, defining U_i by the rule $U_i g = g \circ T_i$ for $g \in \mathcal{H}$. By virtue of Theorem 5.1, the operator

$$P = p\text{-}\lim_{\vec{z}} \prod_{i=1}^{m} U_i^{f_i(\vec{z})}.$$

is an orthogonal projection onto some closed subspace of $\mathcal{H}$. Write g for the characteristic function of the set A, and introduce the abbreviation

$$T = \prod_{i=1}^{m} T_i^{f_i(\vec{z})}.$$

Then we have

$$p\text{-}\lim_{\vec{z}} \mu\left(A \cap T^{-1} A\right) = p\text{-}\lim_{\vec{z}} \left\langle g, \prod_{i=1}^{m} U_i^{f_i(\vec{z})} g \right\rangle = \langle g, Pg \rangle.$$

Now P is a projection; therefore, if $e \equiv 1$ denotes the function identically equal to 1, of norm $\|e\| = 1$,

$$\langle g, Pg \rangle = \|Pg\|^2 = \|Pg\|^2 \|e\|^2 \geq \langle Pg, e \rangle^2 = \langle g, Pe \rangle^2.$$

Finally, $Pe = e$, since e is invariant under the action of the unitary operators U_i, and so

$$\langle g, Pe \rangle^2 = \langle g, e \rangle^2 = \mu(A)^2.$$

Combining the three displayed (in)equalities gives the desired result. $\square$

The consequences of the preceding theorem are twofold. First, when applied to the case of a single measure-preserving transformation, the two inequalities in Theorem 6.1 show precisely that the sets

$$\left\{ f(\vec{z}) \,\middle|\, \vec{z} \in \mathbb{Z}^n \right\}$$

and

$$\left\{ f(n_{\alpha}^1, \ldots, n_{\alpha}^n) \,\middle|\, \alpha \in \mathcal{F} \right\}$$

are sets of nice recurrence; here f may be any polynomial in $\mathrm{Int}[\vec{z}]$ satisfying $f(0) = 0$, and $n_{\bullet}^j$, with $j = 1, \ldots, n$, can be arbitrary IP-sets.

Secondly, we can exploit the fact that the idempotent ultrafilters in Theorem 6.1 may be chosen arbitrarily. Under the assumptions made above, for any $\epsilon > 0$, the set

$$R_{\epsilon} = \left\{ \vec{z} \in \mathbb{Z}^n \,\middle|\, \mu\left(A \cap T^{-1} A\right) \geq \mu(A)^2 - \epsilon \right\}$$

has to be contained in every idempotent ultrafilter in $\beta\mathbb{Z}^n$. The reader will remember that this is equivalent to saying that R_{ϵ} is IP*, that is, intersects every IP-set of $\mathbb{Z}^n$. In particular, every R_{ϵ} is a syndetic set, because the IP* property implies syndeticity. A special case of this result is Khintchine's recurrence theorem,

which states that for a single measure-preserving transformation T, the sets of nice returns

$$\left\{ n \in \mathbb{N} \,\middle|\, \mu\!\left(A \cap T^{-1}A\right) \geq \mu(A)^2 - \epsilon \right\}$$

are syndetic. Several other applications may be found in the original paper [1].

References

[1] V. Bergelson, H. Furstenberg, and R. McCutcheon, *IP-sets and polynomial recurrence*, Ergodic Theory and Dynamical Systems **16** (1996), no. 5, 963–974.

[2] V. Bergelson and R. McCutcheon, *An ergodic IP polynomial Szemerédi theorem*, Memoirs of the American Mathematical Society **146** (2000), no. 695.

[3] V. Bergelson, *Ergodic Ramsey theory—an update*, Ergodic theory of $\mathbb{Z}^d$ actions (Warwick, 1993–1994), 1996, pp. 1–61.

[4] N. Hindman and D. Strauss, *Algebra in the Stone-Čech compactification*, de Gruyter Expositions in Mathematics, vol. 27, Walter der Gruyter & Co., Berlin, 1998.

The Ohio State University, 231 West 18th Avenue, Columbus, OH 43210

E-mail address: schnell@math.ohio-state.edu