Ramsey properties of nonlinear Diophantine equations

Mauro Di Nasso*

Department of Mathematics, University of Pisa,
Largo Pontecorvo 5, 56127 Pisa, ITALY
mauro.di.nasso@unipi.it

Lorenzo Luperi Baglini†

University of Vienna, Faculty of Mathematics,
Oskar-Morgenstern-Platz 1, 1090 Vienna, AUSTRIA,
lorenzo.luperi.baglini@univie.ac.at

Abstract

We present general sufficient and necessary conditions for the partition regularity of Diophantine equations, which extend the classic Rado's Theorem by covering large classes of nonlinear equations. The goal is to contribute to an overall theory of Ramsey properties of (nonlinear) Diophantine equations that encompasses the known results in this area under a unified framework.

Sufficient conditions are obtained by exploiting algebraic properties in the space of ultrafilters $\beta\mathbb{N}$, grounding on combinatorial properties of positive density sets and IP sets. Necessary conditions are proved by a new technique in nonstandard analysis, based on the use of the relation of u-equivalence for the hypernatural numbers $^*\mathbb{N}$.

Introduction

Ramsey theory studies structural combinatorial properties that are preserved under finite partitions. An active area of research in this framework has overlaps with additive number theory, and it focuses on partition properties of the natural numbers related to their semiring structure. Historically, the first result of this kind dates back to 1916; it is a combinatorial lemma that I. Schur [49] used to prove the existence of non-trivial solutions to Fermat equations $x^n + y^n = z^n$ modulo p for all sufficiently large primes p. Precisely, Schur's Lemma states that in every finite coloring (partition) of the natural numbers, one finds a monochromatic triple of the form $a, b, a + b$. Such a property can be phrased by saying that the equation $x + y = z$ is partition regular on $\mathbb{N}$. Another simple equation that is partition regular is $x + y = 2z$; indeed, this amounts to saying that in every finite coloring of $\mathbb{N}$ one finds a 3-term monochromatic arithmetic progression $a, a+d, a+2d$. (We recall that by van der Waerden's Theorem [51], another classic result in Ramsey theory that was proved in 1927, in every finite coloring of $\mathbb{N}$ one finds arbitrarily long monochromatic arithmetic progressions.) However, simple examples of equations that are not partition regular are easily found: *e.g.*, $x + y = 3z$.

In 1933, R. Rado [47] completely characterized partition regular systems of linear Diophantine equations on $\mathbb{N}$, by isolating a simple sufficient and necessary condition on the coefficients, the so-called *column property*. Here is the formulation for a single equation.[1]

Rado's Theorem. *A linear Diophantine equation with no constant term*

$$c_1 x_1 + \cdots + c_n x_n = 0$$

is partition regular on $\mathbb{N}$ if and only if the following condition is satisfied:

- *"There exists a nonempty set $J \subseteq \{1, \ldots, n\}$ such that $\sum_{j \in J} c_j = 0$."*

An active research focused on possible extensions of Rado's Theorem in several directions; in particular, a large amount of interesting results have been obtained during the last twenty years about the various aspects of partition regularity of finite and infinite systems of linear equations (see, *e.g.*, [1, 4, 10, 12, 25, 29, 30, 31, 33, 34, 36, 40, 50]). However, progress on the non-linear case has been sporadic, and structural theorems that provide an overall

---

[1] For a full treatment of Rado's Theorem, see §3.2 and §3.3 of [22].

understanding of Ramsey properties of nonlinear Diophantine equations are still missing.

Let us briefly recall all the relevant results on this last topic that we are aware of. The simplest result is the multiplicative formulation of Rado's Theorem.

Multiplicative Rado's Theorem. *A nonlinear Diophantine equation*

$$\prod_{i=1}^{n} x_i^{c_i} = 1$$

is partition regular on $\mathbb{N}$ if and only if the following condition is satisfied:

- *"There exists a nonempty set $J \subseteq \{1, \ldots, n\}$ such that $\sum_{j \in J} c_j = 0$."*

The first attempt for a systematic study of the nonlinear case is found in the paper [41] of 1991, where H. Lefmann proved several results about the partition regularity of systems of homogeneous polynomials. In particular, he gave a simple characterization in the case where every monomial contains a single variable raised to the same exponent $1/k$. Here is the formulation for a single equation.

Lefmann's Theorem. *Let $k \in \mathbb{N}$. A Diophantine equation of the form*

$$c_1 x_1^{1/k} + \cdots + c_n x_n^{1/k} = 0$$

is partition regular on $\mathbb{N}$ if and only if "Rado's condition" is satisfied:

- *"There exists a nonempty set $J \subseteq \{1, \ldots, n\}$ such that $\sum_{j \in J} c_j = 0$."*

H. Lefmann also proved a similar characterization when all exponents of variables x_i are -1. For instance, the analog of Schur's Lemma for reciprocals is valid, *i.e.* the equation $1/x + 1/y = 1/z$ is partition regular. More generally, T.C. Brown and V. Rődl [8] showed that the family of partition regular homogeneous functions is closed under the operation of taking reciprocals of variables: $f(x_1, \ldots, x_n) = 0 \mapsto f(1/x_1, \ldots, 1/x_n) = 0$ (see Theorem 2.6 in the next section).

In 1996, grounding on a classic density result by A. Sárkőzy and H. Furstenberg, V. Bergelson [2] showed the partition regularity of all Diophantine equations of the form $x - y = P(z)$ where the polynomial $P(z) \in \mathbb{Z}[z]$ has no constant term (see the remarks preceding Question 11).[2]

---

[2] Precisely, A. Sárkőzy [48] and H. Furstenberg [20] proved independently that for every $P(z) \in \mathbb{Z}[z]$ with no constant term and for every set $A \subseteq \mathbb{N}$ of positive upper density there exist pairs $x, y \in A$ having distance $x - y = P(n)$ for some $n \in \mathbb{N}$.

With the few exceptions mentioned above, all research on the partition regularity of nonlinear Diophantine equations has been developed in the past 10 years, and the great part of it appeared in the last two years.

In 2006, A. Khalfalah and E. Szemerédi [39] proved that if $P(z) \in \mathbb{Z}[z]$ takes even values on some integer, then the equation $x + y = P(z)$ is "partially" partition regular in the variables x and y, *i.e.*, for every finite coloring of $\mathbb{N}$ one finds a solution x, y, z where x and y are monochromatic.

In 2010, P. Csikvári, K. Gyarmati and A. Sárközy [9] proved density results involving nonlinear problems over $\mathbb{N}$ and over finite fields. In particular, they proved that the equation $x + y = z^2$ is not partition regular.[3] At the foot of the paper, they left as an open problem the partition regularity of $x + y = tz$, which is particularly relevant as the most basic equation that mixes additive and multiplicative structure on $\mathbb{N}$.

In 2010, by using algebra in the space of ultrafilters $\beta\mathbb{N}$, V. Bergelson [3] solved that problem in the positive. Independently, also N. Hindman [32] proved that property and, more generally, the partition regularity of all equations of the form $\sum_{i=1}^{n} x_i = \prod_{i=1}^{n} y_i$. In 2014, the second named author [44] extended Hindman's result, and by nonstandard methods he proved the following: For every choice of sets $F_i \subseteq \{1, \ldots, m\}$, the equation $\sum_{=1}^{n} c_i \, x_i (\prod_{j \in F_i} y_j) = 0$ is partition regular whenever $\sum_{i \in J} c_j = 0$ for some nonempty $J \subseteq \{1, \ldots, m\}$. (It is agreed that $\prod_{j \in \emptyset} y_j = 1$.)

An important contribution in the case of quadratic equations was given by N. Frantzikinakis and B. Host [19] (the first version of their paper was made available on the web in 2014). As a consequence of their structural theorem for multiplicative functions, they proved that the equations $16x^2 + 9y^2 = z^2$ and $x^2 - xy + y^2 = z^2$ are partially partition regular in the variables x and y.

In [18], M. Riggio and the first named author used nonstandard analysis to identify a large class of Fermat-like equations that are not partition regular, the simplest cases being $x^m + y^n = z^k$ where $k \notin \{n, m\}$.[4]

The partition regularity of nonlinear Diophantine equations can be seen as a particular case of the study of monochromatic polynomial configurations. Arguably, the most interesting open problem in this field, asked on numerous occasions by several researchers (see, *e.g.*, [2, Question 11 (ii)], or [35,

---

[3] The paper [9] appeared only in 2012, but its first draft circulated since 2010. There, one finds also a proof of the partition regularity of $x(y + z) = yz$, which is the same as the "reciprocal Schur-equation" mentioned above, as well as a proof of the partition regularity of $xy = z^2$, which is a particular case of the multiplicative Rado's Theorem.

[4] Here we do not count the constant solution $z = y = z = 2$ of $x^n + y^n = z^{n+1}$.

Question 3]) is the partition regularity of the configuration $\{x, y, x+y, xy\}$. As reported in [28], already in the last years of the 1970s, R. Graham proved with computer assistance that one finds monochromatic $a, b, a+b, ab$ in every 2-coloring of $\{1, \ldots, 252\}$, and that the same property fails in $\{1, \ldots, 251\}$. Notice that this gives no answer to the general problem where one considers partitions with an arbitrary finite number of pieces.

Several papers over the last few months have investigated around that problem on additive-multiplicative structure of the integers.

V. Bergelson and J. Moreira [6, 7] showed that patterns $\{x+y, xy\}$ are partition regular on infinite fields. In January 2015, V. Bergelson, J.H. Johnson jr. and J. Moreira [4] proved the partition regularity of many polynomial configurations, including $\{x, y+x^2, z, z+y^2\}$, by using a generalization of the notion of (m, p, c)-set, originally introduced by W. Deuber [11].

In February of this year 2016, B. Green and T. Sanders [24] solved positively the partition regularity problem of $\{x, y, x+y, xy\}$ on every finite field $\mathbb{F}_p$.

In May 2016, when this paper was already completed and under a final revision, J. Moreira [46] proved the partition regularity of a large class of configurations, including $\{x, x+y, xy\}$. As a corollary of his main theorem, one gets the partition regularity of all Diophantine equations of the form $c_1 x_1^2 + \ldots + c_n x_n^2 = y$ where the sum of coefficients $c_1 + \ldots + c_n = 0$. (So, $e.g.$, $x^2 + y = z^2$ is partition regular.) We remark that the above equations where $\sum_{i \in J} c_i \neq 0$ for every nonempty $J \subseteq \{1, \ldots, n\}$ are not partition regular, as we will show in Section 3 (see Corollary 3.13).

Finally, in the last days of May 2016, it was breaking news that M. J. H. Heule, O. Kullmann and V. W. Marek [26] solved an old problem that was posed by P. Erdős and R. Graham in 1975 (see, $e.g.$ [23]), namely the *Boolean Pythagorean triples problem*, that asked whether the equation $x^2 + y^2 = z^2$ is partition regular for 2-colorings of $\mathbb{N}$. By means of a computer-assisted proof, they proved that any 2-coloring of $\{1, 2, \ldots, 7825\}$ contains a monochromatic Pythagorean triple, and that 7825 is the least number with such a property.[5] Notice that the huge proof, contained in a file of 200 terabytes, still does not solve the full problem of partition regularity of the Pythagorean equation $x^2 + y^2 = z^2$, where a finite (but *arbitrary*) number of colors is allowed in partitions.

---

[5] See also the article of E. Lamb appeared online in the journal *Nature* on May 26, 2016.

The several contributions appeared over the very last months give evidence on the rising interest of researchers in Ramsey properties of nonlinear Diophantine equations.

In this paper we consider Diophantine equations in their full generality, aiming at finding simple conditions on coefficients and exponents that imply either partition regularity or non-partition regularity. The ultimate goal is to extend Rado's Theorem and develop a general Ramsey theory of Diophantine equations.

The techniques that are used here are twofold. On the side of sufficient conditions for partition regularity (Section 2), we use the algebraic structure of the space of ultrafilters $\beta\mathbb{N}$, combined with properties of difference sets of sets of positive asymptotic density. On the side of necessary conditions (Section 3), we work in the setting of hypernatural numbers $^*\mathbb{N}$ of nonstandard analysis, the instrumental tool being the relation of u-equivalence and its properties. Basically, u-equivalence formalizes the well-known characterization of partition regularity in terms of ultrafilters within a nonstandard framework. However, whilst this technique is based on nonstandard analysis, the arguments used are of a purely combinatorial nature.

1 Preliminary definitions and results

1.1 Asymptotic density

Following a common practice in number theory, with $\mathbb{N}$ we denote the set of *positive* integers; and with $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ we denote the set of *non-negative* integers. Recall that the *upper asymptotic density* of a set $A \subseteq \mathbb{N}$ is defined as follows:

$$\overline{d}(A) = \limsup_{n \to \infty} \frac{|A \cap [1, n]|}{n}.$$

By replacing initial intervals $[1, n]$ with arbitrary intervals, one obtains the following generalization.

Definition 1.1. *The* Banach density *$BD(E)$ of a set $E \subseteq \mathbb{Z}^t$ is the greatest of the following superior limits of relative densities*

$$\limsup_{n \to \infty} \frac{|E \cap R_n|}{|R_n|}$$

6

where $(R_n = \prod_{i=1}^{t}[a_{ni}, b_{ni}])_{n\in\mathbb{N}}$ are sequences of rectangles whose size in every direction approaches infinity, i.e. $\lim_{n\to\infty}(b_{ni} - a_{ni}) = +\infty$ *for $i = 1, \ldots, t$.*

It can be checked that such a greatest value is actually attained. In the one-dimensional case, equivalently one can define $\mathrm{BD}(E) = \lim_n e_n/n = \inf e_n/n$, where e_n is the greatest cardinality of an intersection $E \cap I$ where I is an interval of length n. Clearly, $\overline{d}(A) \leq \mathrm{BD}(A)$ for every $A \subseteq \mathbb{N}$.

1.2 IP-sets

A relevant notion in combinatorial number theory is that of *IP-set*.

Definition 1.2. *Let $G = (g_i)_{i\in\mathbb{N}}$ be an increasing sequence of natural numbers. The* IP-set generated *by G is the set of finite sums*

$$FS(G) = FS(g_i)_{i\in\mathbb{N}} = \left\{ \sum_{j=1}^{k} g_{i_j} \;\middle|\; i_1 < i_2 < \cdots < i_k \right\}.$$

A set $A \subseteq \mathbb{N}$ is called IP-large *if it contains an IP-set.* Multiplicative IP-sets *and multiplicative* IP-large sets *are defined similarly.*

By the celebrated *Hindman's Theorem* [27], in every finite partition of the natural numbers $\mathbb{N} = C_1 \cup \ldots \cup C_r$, one of the pieces is IP-large; this result can be improved to obtain the existence of a single piece that is both additively and multiplicatively IP-large (see §5.3 of [37]).

An instrumental tool for the main result of the next section is a theorem proved by V. Bergelson, H. Furstenberg and R. McCutcheon [5, Theorem C], that we now recall.

Let us first fix a convenient notation. Let *Fin* denote the family of all nonempty finite subsets of $\mathbb{N}$. Given an increasing sequence $G = (g_i)_{i\in\mathbb{N}}$ of natural numbers, for $\alpha \in Fin$ denote by $n_\alpha = \sum_{i\in\alpha} g_i$. Clearly, $n_\alpha + n_\beta = n_{\alpha\cup\beta}$ whenever $\alpha \cap \beta = \emptyset$, and the IP-set $FS(G)$ is obtained as the range of the sequence $(n_\alpha)_{\alpha\in Fin}$. Conversely, if $(n_\alpha)_{\alpha\in Fin}$ is a sequence such that $n_\alpha + n_\beta = n_{\alpha\cup\beta}$ whenever $\alpha \cap \beta = \emptyset$, then its range is an IP-set, namely $\{n_\alpha \mid \alpha \in Fin\} = FS(G)$ where $G = (n_{\{i\}})_{i\in\mathbb{N}}$. So, in a precise sense, the two notions are equivalent.

Theorem 1.3. ([5, Theorem C]) Let $E \subseteq \mathbb{Z}^t$ have positive Banach density, and let

- $P_1, \ldots, P_t \in \mathbb{Z}[x_1, \ldots, x_k]$ be polynomials with no constant terms;

- $(n_\alpha^{(1)})_{\alpha \in Fin}, \ldots, (n_\alpha^{(k)})_{\alpha \in Fin}$ be IP-sets.

Then there exist $e_1, e_2 \in E$ and $\alpha \in Fin$ such that

$$e_1 - e_2 = \left(P_1(n_\alpha^{(1)}, \ldots, n_\alpha^{(k)}), \ldots, P_t(n_\alpha^{(1)}, \ldots, n_\alpha^{(k)}) \right).$$

1.3 Algebra in the space of ultrafilters $\beta\mathbb{N}$

In this paper we assume the reader to be familiar with the fundamental properties of the space $\beta\mathbb{N}$ of ultrafilters on $\mathbb{N}$ endowed with the operations of pseudo-sum $\oplus$ and pseudo-product $\odot$:

- $A \in \mathcal{U} \oplus \mathcal{V} \Leftrightarrow \{n \mid A - n \in \mathcal{V}\} \in \mathcal{U}$, where $A - n = \{m \in \mathbb{N} \mid m + n \in A\}$;

- $A \in \mathcal{U} \odot \mathcal{V} \Leftrightarrow \{n \mid A/n \in \mathcal{V}\} \in \mathcal{U}$, where $A/n = \{m \in \mathbb{N} \mid mn \in A\}$.

In particular, we assume some knowledge of *idempotent ultrafilters* and left and right *ideals* in the compact *topological right semigroups* $(\beta\mathbb{N}, \oplus)$ and $(\beta\mathbb{N}, \odot)$. For simplicity, we will use the adjective "additive" when referring to the former, and "multiplicative" when referring to the latter. So, for instance, the ultrafilter $\mathcal{U}$ is additively idempotent if $\mathcal{U} \oplus \mathcal{U} = \mathcal{U}$, and $\mathcal{U}$ is multiplicatively idempotent if $\mathcal{U} \odot \mathcal{U} = \mathcal{U}$. We will use the following notation.

- $K(\oplus)$ is the minimal additive two sided ideal;

- $K(\odot)$ is the minimal multiplicative two sided ideal;

- $\mathbb{I}(\oplus)$ is the set of additively idempotent ultrafilters;

- $\mathbb{I}(\odot)$ is the set of multiplicatively idempotent ultrafilters;

- $\mathbb{M}(\oplus) = \mathbb{I}(\oplus) \cap K(\oplus)$ is the set of minimal additive idempotents;

- $\mathbb{M}(\odot) = \mathbb{I}(\odot) \cap K(\odot)$ is the set of minimal multiplicative idempotents;

- $\mathcal{BD} = \{\mathcal{U} \in \beta\mathbb{N} \mid \forall A \in \mathcal{U} \ \mathrm{BD}(A) > 0\}$;

- $\mathcal{D} = \{\mathcal{U} \in \beta\mathbb{N} \mid \forall A \in \mathcal{U} \ \bar{d}(A) > 0\} \subseteq \mathcal{BD}$.

For convenience, we itemize the known results about algebra in $\beta\mathbb{N}$ that we will use in this paper. A comprehensive reference is Hindman and Strauss' book [37], where all proofs can be found.[6]

(B1) The closure $\overline{\mathbb{I}(\oplus)} = \{\mathcal{U} \in \beta\mathbb{N} \mid \forall A \in \mathcal{U} \ A \text{ is IP-large}\}$;

(B2) $\overline{\mathbb{I}(\oplus)}$ is a multiplicative left ideal;

(B3) The closure $\overline{\mathbb{I}(\odot)} = \{\mathcal{U} \in \beta\mathbb{N} \mid \forall A \in \mathcal{U} \ A \text{ is multiplicative IP-large}\}$;

(B4) $\overline{\mathbb{M}(\oplus)}$ is a multiplicative left ideal;

(B5) $\mathcal{BD}$ is a closed additive two sided ideal;

(B6) $\mathcal{BD}$ is a closed multiplicative left ideal;

(B7) $\mathcal{D}$ is a closed additive left ideal;

(B8) $\mathcal{D}$ is a closed multiplicative left ideal;

(B9) $\mathcal{D} \cap \overline{\mathbb{M}(\oplus)} \cap \mathbb{M}(\odot) \neq \emptyset$.

Ultrafilters in $\mathcal{D} \cap \overline{\mathbb{M}(\oplus)} \cap \mathbb{M}(\odot)$ are particularly relevant. They were first isolated and studied by N. Hindman and D. Strauss, who named them *combinatorially rich* ([37, Definition 17.1]).

1.4 Partition regularity of functions

By *finite coloring* we mean a finite partition of the natural numbers. Elements $a_1, \ldots, a_k$ are called *monochromatic* with respect to a given finite coloring $\mathbb{N} = C_1 \cup \ldots \cup C_r$ if there exists C_i such that $a_1, \ldots, a_k \in C_i$.

Definition 1.4. A function $f(x_1, \ldots, x_n)$ is called *partition regular* on $\mathbb{N}$ (or simply PR) if in every finite coloring of $\mathbb{N}$ one finds a monochromatic root (or solution), *i.e.* monochromatic elements $a_1, \ldots, a_n$ with $f(a_1, \ldots, a_n) = 0$.

---

[6] Precisely, property (B1) and (B3) are particular cases of Lemma 5.11; property (B2) is Theorem 5.20; property (B4) is Theorem 5.20. A proof of properties (B5) and (B6) is found in §20.1, where $\mathcal{BD}$ is denoted $\Delta^*(\mathbb{N}, +)$; and properties (B7) and (B8) are in Theorem 6.79, where $\mathcal{D}$ is denoted $\Delta(\mathbb{N}, +)$. Finally, property (B9) is Lemma 17.2 where our set $\overline{\mathbb{M}(\oplus)}$ is denoted $\mathbb{M}$.

When it is possible to find such elements a_i that are pairwise different, the function is called *injectively* PR.[7]

More generally, if $\{x_{i_1}, \ldots, x_{i_p}\} \subseteq \{x_1, \ldots, x_n\}$, the function $f(x_1, \ldots, x_n)$ is called partition regular with *injectivity* $|\{x_{i_1}, \ldots, x_{i_p}\}| \geq s$ if in every finite coloring one finds a monochromatic root $a_1, \ldots, a_n$ with $|\{a_{i_1}, \ldots, a_{i_p}\}| \geq s$.

A function $f(x_1, \ldots, x_n)$ is called *non-trivially* PR if it is partition regular with injectivity $|\{x_1, \ldots, x_n\}| \geq 2$.

The above definitions of partition regularity are extended to equations $f(x_1, \ldots, x_n) = g(y_1, \ldots, y_k)$ in the obvious way, by considering the corresponding notions for the function $f - g$. So, for instance, the classic *Schur's Theorem* [49] of 1916 can be equivalently formulated as: "The function $f(x, y, z) = x + y - z$ is injectively PR", or as: "The equation $x + y = z$ is injectively PR".

A fundamental result about partition regularity that dramatically generalizes the result of Shur's mentioned above, was proved in 1933 by R. Rado [47], who completely solved the linear Diophantine case.

Theorem 1.5 (Rado). *A linear Diophantine homogeneous equation*

$$c_1 x_1 + \cdots + c_n x_n \ = \ 0$$

is PR on $\mathbb{N}$ if and only if the following "Rado's condition" is satisfied:

- *"There exists a nonempty set $J \subseteq \{1, \ldots, n\}$ such that $\sum_{j \in J} c_j = 0$."*

Moreover, a linear Diophantine inhomogeneous equation

$$c_1 x_1 + \cdots + c_n x_n = d$$

is PR on $\mathbb{N}$ if and only if

- *either there exists a natural number k such that $\sum_{i=1}^{n} c_i k = d$,*

- *or there exists an integer z such that $\sum_{i=1}^{n} c_i z = d$ and there exists a nonempty subset $J \subseteq \{1, ..., n\}$ such that $\sum_{j \in J} c_j = 0$.*

---

[7] The expression "injectively PR" is commonly used in the literature (see, *e.g.*, [31]). Let us remark that the word "injectively" in this definition is not related to the injectivity of the function $f(x_1, \ldots, x_n)$.

Notice that one cannot have *injective* PR when the number of variables $n = 2$ because, in this case, Rado's condition implies that $c_1 = -c_2$ and the equation reduces to the trivial equality $x_1 = x_2$. On the other hand, as recently shown by N. Hindman and I. Leader in a more general setting, the following holds:

Theorem 1.6. ([31, Theorem 3.1]) A linear Diophantine equation in more than two variables is PR on $\mathbb{N}$ if and only if it is injectively PR on $\mathbb{N}$.

E.g., for $n \geq 2$ the following equations are injectively PR:

$$x_1 = x_2 + a_1 y_1 + \ldots + a_n y_n.$$

It is well-known that partition regularity can be equivalently expressed in terms of ultrafilters (see, *e.g.*, [37, Theorem 5.7]). Here we refine the equivalence by also considering injectivity conditions.

Definition 1.7. An ultrafilter $\mathcal{U}$ on $\mathbb{N}$ is called a PR-*witness* of the function $f(x_1, \ldots, x_n)$ with injectivity $|\{x_{i_1}, \ldots, x_{i_p}\}| \geq s$ if for every $A \in \mathcal{U}$ there exist $a_1, \ldots, a_n \in A$ such that $f(a_1, \ldots, a_n) = 0$ and $|\{a_{i_1}, \ldots, a_{i_p}\}| \geq s$.

Proposition 1.8. *A function $f(x_1, \ldots, x_n)$ is partition regular with injectivity $|\{x_{i_1}, \ldots, x_{i_p}\}| \geq s$ if and only if there exists a PR-witness of $f(x_1, \ldots, x_n)$ with injectivity $|\{x_{i_1}, \ldots, x_{i_p}\}| \geq s$.*

Proof. One direction is trivial because in every finite coloring, one and only one of the colors belongs to $\mathcal{U}$, by the property of ultrafilter. Conversely, notice that the following family

$$\mathcal{F} = \{A \subseteq \mathbb{N} \mid \forall a_1, \ldots, a_n \in A^c \ |\{a_{i_1}, \ldots, a_{i_p}\}| \geq s \Rightarrow f(a_1, \ldots, a_n) \neq 0\}$$

has the finite intersection property. Indeed, if by contradiction $A_1, \ldots, A_n \in \mathcal{F}$ were such that $\bigcap_{i=1}^n A_i = \emptyset$, then the finite coloring $\mathbb{N} = A_1^c \cup \ldots \cup A_n^c$ would provide a counter-example to the PR of $f(x_1, \ldots, x_n)$ with injectivity $|\{x_{i_1}, \ldots, x_{i_p}\}| \geq s$. Finally, notice that any ultrafilter $\mathcal{U} \supseteq \mathcal{F}$ is the desired PR-witness; indeed, if $B \in \mathcal{U}$ was a counter-example, then its complement $B^c \in \mathcal{F} \subseteq \mathcal{U}$, and hence $\emptyset = B \cap B^c \in \mathcal{U}$, a contradiction. $\qquad\square$

2   Sufficient conditions for PR

Let us first prove a useful property about ultrafilters that simultaneously witness several equations.

Lemma 2.1. *Assume that for $i = 1, \ldots, k$, the same ultrafilter $\mathcal{U}$ is a PR-witness of $f_i(x_{i,1}, \ldots, x_{i,n_i}) = 0$ with injectivity $|\{x_{i,j_1}, \ldots, x_{i,j_{p_i}}\}| \geq s_i$. If the functions f_i have pairwise disjoint sets of variables[8] then $\mathcal{U}$ is also a PR-witness of the following system of equations:*

$$\begin{cases} f_i(x_{i,1}, \ldots, x_{i,n_i}) = 0 \quad i = 1, \ldots, k; \\ x_{1,1} = \ldots = x_{k,1} \end{cases}$$

with injectivity $|\{x_{i,j_1}, \ldots, x_{i,j_{p_i}}\}| \geq s_i$ for $i = 1, \ldots, k$.

Proof. Let $A \in \mathcal{U}$ be fixed. For every $i = 1, \ldots, k$ let

$$\Lambda_i \ = \ \{ a \in A \mid \exists \, a_{i,2}, \ldots, a_{i,n_i} \in A \text{ s.t. } |\{a_{i,j_1}, \ldots, a_{i,j_{p_i}}\}| \geq s_i$$
$$\& \ f_i(a, a_{i,2}, \ldots, a_{i,n_i}) = 0 \}.$$

(If $j_1 = 1$, we agree that $a_{i,1} = a$.) Notice that $\Lambda_i \in \mathcal{U}$, as otherwise

$$\Lambda_i^c \cap A \ = \ \{ a \in A \mid \forall \, a_{i,2}, \ldots, a_{i,n_i} \in A \ \ |\{a_{i,j_1}, \ldots, a_{i,j_{p_i}}\}| \geq s_i \Rightarrow$$
$$f_i(a, a_{i,2}, \ldots, a_{i,n_i}) \neq 0 \}$$

would belong to $\mathcal{U}$, contradicting the hypothesis that $\mathcal{U}$ is a PR-witness of f_i with injectivity $|\{x_{i,j_1}, \ldots, x_{i,j_{p_i}}\}| \geq s_i$. Then the intersection $\Lambda = \bigcap_{i=1}^k \Lambda_i \in \mathcal{U}$ is nonempty and we can pick an element $a_{i,1} \in \Lambda \subseteq A$. It directly follows from the definitions that there are elements $a_{i,2}, \ldots, a_{i,n_i} \in A$ with $|\{a_{i,j_1}, \ldots, a_{i,j_{p_i}}\}| \geq s_i$ and such that $f_i(a_{i,1}, a_{i,2}, \ldots, a_{i,n_i}) = 0$ for $i = 1, \ldots, k$. This shows the existence of solutions in A to the considered system, with the desired injectivity properties. $\square$

Example 2.2. As shown by V. Bergelson [2], the equation $u - v = t^2$ is PR. By Lemma 2.1, if $\mathcal{U}$ is a witness of its PR, then $\mathcal{U}$ is a witness also of the PR of the system

$$\begin{cases} u_1 - y = x^2; \\ u_2 - z = t^2; \\ y = t. \end{cases}$$

---

[8] That is, $\{x_{i,1}, \ldots, x_{i,k_i}\} \cap \{x_{j,1}, \ldots, x_{j,k_j}\} = \emptyset$ for $j \neq i$.

It is readily seen that this is equivalent to the PR of the configuration $\{x, y, z, y + x^2, z + y^2\}$. Notice that this improves on a result contained in [4], namely the PR of $\{x, y + x^2, z, z + y^2\}$.

Example 2.3. Let $P_1, \ldots, P_n \in \mathbb{Z}[x]$ be polynomials with no constant terms. By Theorem 1.3, it follows that there exists an ultrafilter $\mathcal{U}$ that is a common PR-witness of all equations $x_i - y_i = P_i(z_i)$.[9] Then one can apply Lemma 2.1 and obtain that $\mathcal{U}$ also witnesses the PR of the system

$$\begin{cases} x_i - y_i = P_i(z_i) & i = 1, \ldots, n \,; \\ y_i = z_{i+1} & i = 1, \ldots, n-1. \end{cases}$$

As a consequence, one obtains the PR of the system

$$\begin{cases} x_1 - y_1 = P_1(z_1)\,; \\ x_i - y_i = P_i(y_{i-1}) & i = 2, \ldots, n. \end{cases}$$

Notice that this is precisely Corollary 1.11 of [4].

Recall that a function $f(x_1, \ldots, x_n)$ is *homogeneous* if there exists ℓ such that for every $\lambda, x_1, \ldots, x_n$ one has $f(\lambda x_1, \ldots, \lambda x_n) = \lambda^\ell f(x_1, \ldots, x_n)$. In this case ℓ is called the *degree of homogeneity* of f.

The following ultrafilter property was first proved by the second named author by nonstandard analysis; the proof given below uses an essentially equivalent ultrafilter argument.

Theorem 2.4. ([44, Theorem 3.1]) Assume that the equation $f(x_1, \ldots, x_n)$ is PR with injectivity $|\{x_{i_1}, \ldots, x_{i_p}\}| \geq s$. If f is homogeneous then the set of PR-witnesses

$$\mathfrak{W}_f = \{\mathcal{U} \in \beta\mathbb{N} \mid \mathcal{U} \text{ is a PR-witness of } f \text{ with injectivity } |\{x_{i_1}, \ldots, x_{i_p}\}| \geq s\}$$

is a closed multiplicative two sided ideal.

Proof. Let $\mathcal{U} \in \mathfrak{W}_f$ and $\mathcal{V} \in \beta\mathbb{N}$. By definition, $B \in \mathcal{U} \odot \mathcal{V}$ if and only if $\widehat{B} = \{m \in \mathbb{N} \mid B/m \in \mathcal{V}\} \in \mathcal{U}$. Now let $b_1, \ldots, b_n \in \widehat{B}$ be such that $f(b_1, \ldots, b_n) = 0$ and $|\{b_{i_1}, \ldots, b_{i_p}\}| \geq s$, pick any $\lambda \in \bigcap_{j=1}^p B/b_{i_j} \in \mathcal{V}$, and consider the elements $\lambda b_1, \ldots, \lambda b_n \in B$. By homogeneity, $f(\lambda b_1, \ldots, \lambda b_n) =$

---

[9] A strenghtening of this fact will be proved in Theorem 2.9.

$\lambda^\ell f(b_1, \ldots, b_n) = 0$; moreover, $|\{\lambda b_{i_1}, \ldots, \lambda b_{i_p}\}| = |\{b_{i_1}, \ldots, b_{i_p}\}| \geq s$. This shows that $\mathcal{U} \odot \mathcal{V} \in \mathfrak{W}_f$, and hence we can conclude that $\mathfrak{W}_f$ is a multiplicative right ideal. Moreover, it is verified in a straightforward manner that $\mathfrak{W}_f$ is (topologically) closed in $\beta\mathbb{N}$. Finally, recall that every closed right ideal is also a left ideal (see [37, Theorem 2.19]). $\qquad\square$

The intersection of all closed two sided ideals equals the closure of the minimal ideal, and so:

Corollary 2.5. *Let $f(x_1, \ldots, x_n)$ be a homogeneous function that is PR with injectivity $|\{x_{i_1}, \ldots, x_{i_p}\}| \geq s$. Then every $\mathcal{U} \in \overline{K(\odot)}$ is a PR-witness of f with injectivity $|\{x_{i_1}, \ldots, x_{i_p}\}| \geq s$.*

Next, we give an ultrafilter proof of a result by T.C. Brown and V. Rődl [8], showing that the class of PR homogeneous functions is stable under the operation of "inverting variables".

Theorem 2.6. [8, Theorem 2.1]) *If a homogeneous function $f(x_1, \ldots, x_n)$ is PR with injectivity $|\{x_{i_1}, \ldots, x_{i_p}\}| \geq s$ then also $f(1/x_1, \ldots, 1/x_n)$ is PR with injectivity $|\{x_{i_1}, \ldots, x_{i_p}\}| \geq s$.*

Proof. Pick a non-principal ultrafilter $\mathcal{U}$ on $\mathbb{N}$ that is a PR-witness of f. Let $\rho : \mathbb{N} \to \mathbb{Q}$ be the "reciprocal map" $\rho(n) = 1/n$, let $\varphi : \mathbb{N} \to \mathbb{Q}$ be "factorial map" $\varphi(n) = n!$, and consider the *image ultrafilters* $\mathcal{U}_1 = \rho(\mathcal{U})$ and $\mathcal{U}_2 = \varphi(\mathcal{U})$.[10] Since $\mathcal{U}_1, \mathcal{U}_2$ are ultrafilters on $\mathbb{Q}$, it makes sense to consider their pseudo-product $\mathcal{V} = \mathcal{U}_1 \odot \mathcal{U}_2$ in $\beta\mathbb{Q}$, which is defined similarly as pseudo-products in $\beta\mathbb{N}$. We want to show that $\mathbb{N} \in \mathcal{V}$, and that the ultrafilter $\mathcal{V}_\mathbb{N} = \mathcal{V} \cap \mathcal{P}(\mathbb{N})$ is a PR-witness of $f(1/x_1, \ldots, 1/x_n)$. By the definitions,

$$\mathbb{N} \in \mathcal{V} \iff \{u \in \mathbb{Q} \mid \mathbb{N}/u \in \mathcal{U}_2\} \in \mathcal{U}_1 \iff \Lambda = \{n \in \mathbb{N} \mid \mathbb{N}/_{1/n} \in \mathcal{U}_2\} \in \mathcal{U}.$$

For every $n \in \mathbb{N}$, we have that

$$\mathbb{N}/_{1/n} \in \mathcal{U}_2 \iff \Gamma_n = \{m \in \mathbb{N} \mid m! \in \mathbb{N}/_{1/n}\} = \{m \in \mathbb{N} \mid m!/n \in \mathbb{N}\} \in \mathcal{U}.$$

Notice that $\Gamma_n \in \mathcal{U}$ because it contains all $m \geq n$ and $\mathcal{U}$ is non-principal. So, $\Lambda = \mathbb{N} \in \mathcal{U}$, and this proves that $\mathbb{N} \in \mathcal{V}$. Now let $B \in \mathcal{V}_\mathbb{N}$. Since $B \in \mathcal{V}$, the set $\Lambda(B) = \{n \in \mathbb{N} \mid B/_{1/n} \in \mathcal{U}_2\} \in \mathcal{U}$, and hence there exist

---

[10] Recall that if $\mathcal{U}$ is an ultrafilter on a set I and $f : I \to J$ is a function, the image $f(\mathcal{U})$ is the ultrafilter on J where $A \in f(\mathcal{U}) \iff \{i \in I \mid f(i) \in A\} \in \mathcal{U}$ for every $A \subseteq J$.

$a_1, \ldots, a_n \in \Lambda(B)$ such that $|\{a_{i_1}, \ldots, a_{i_p}\}| \geq s$ and $f(a_1, \ldots, a_n) = 0$. Now recall that $B/_{1/a_i} \in \mathcal{U}_2 \Leftrightarrow \Gamma_i(B) = \{m \in \mathbb{N} \mid m!/a_i \in B\} \in \mathcal{U}$; in particular we can pick $k \in \bigcap_{i=1}^n \Gamma_i(B) \in \mathcal{U}$. Finally, notice that elements $b_i = k!/a_i \in B$ are such that $|\{b_{i_1}, \ldots, b_{i_p}\}| \geq s$, since $|\{a_{i_1}, \ldots, a_{i_p}\}| \geq s$; moreover,

$$f(1/b_1, \ldots, 1/b_n) \;=\; f\left(\frac{a_1}{k!}, \ldots, \frac{a_n}{k!}\right) \;=\; \left(\frac{1}{k!}\right)^\ell \cdot f(a_1, \ldots, a_n) \;=\; 0\,,$$

where ℓ is the degree of homogeneity of f. $\qquad\square$

We are now ready to extend Rado's Theorem on the side of sufficient conditions for PR. Let us start with the following consequence of Theorem 1.3, which is particularly relevant to our purposes.

Theorem 2.7. *Let $c\,(x_1 - x_2) = P(y_1, \ldots, y_k)$ be a Diophantine equation where the polynomial P has no constant term and $c \neq 0$. If the set $A \subseteq \mathbb{N}$ is IP-large and has positive Banach density then there exist $\xi_1, \xi_2 \in A$ and mutually distinct $\eta_1, \ldots, \eta_k \in A$ such that $c\,(\xi_1 - \xi_2) = P(\eta_1, \ldots, \eta_k)$. Moreover, if $k = 1$ then one can take $\xi_1 \neq \xi_2$.*

Proof. First of all, notice that we can pick an IP-set $(n_\alpha)_{\alpha \in Fin} \subseteq A$ such that $n_\alpha \neq n_\beta$ for $\alpha \neq \beta$. Indeed, given any IP-set $\mathrm{FS}(g_i)_{i \in \mathbb{N}}$, one can inductively define a sub-IP-set $\mathrm{FS}(g_i') \subseteq \mathrm{FS}(g_i)$ with the desired property, by setting $g_1' = g_1$ and $g_{i+1}' = \min\{g_j \mid g_j > g_1' + \ldots + g_i'\}$.

Now fix a permutation $\sigma : \mathbb{N} \to \mathbb{N}$ with no finite cycles, *i.e.* such that for every $s \in \mathbb{N}$, the iterated composition $\sigma^s(n) \neq n$ for all $n \in \mathbb{N}$. This ensures that for every $\alpha \in Fin$, the sets $\alpha_s = \{\sigma^s(i) \mid i \in \alpha\}$ are pairwise distinct. Indeed, if $\alpha_{s+\ell} = \alpha_s$ for some $s, \ell \in \mathbb{N}$, then $\alpha_\ell = \alpha$ and σ^ℓ would have a finite cycle, contradicting our assumption on σ. In consequence, by our choice of the IP-set, we have $n_{\alpha_s} \neq n_{\alpha_{s'}}$ for $s \neq s'$. Moreover, for every s, the set $\{n_{\alpha_s} \mid \alpha \in Fin\}$ is an IP-set, because for every $\alpha, \beta \in Fin$ one has that $n_{(\alpha \cup \beta)_s} = n_{\alpha_s} + n_{\beta_s}$ whenever $\alpha \cap \beta = \emptyset$. (Notice that $(\alpha \cup \beta)_s = \sigma^s(\alpha \cup \beta) = \sigma^s(\alpha) \cup \sigma^s(\beta) = \alpha_s \cup \beta_s$, where $\alpha_s \cap \beta_s = \emptyset$ because $\alpha \cap \beta = \emptyset$.)

Now consider the set $cA = \{ca \mid a \in A\}$. As $\mathrm{BD}(cA) = \frac{\mathrm{BD}(A)}{|c|} > 0$, we can apply Theorem 1.3 with $t = 1$, $E = cA$, $P_1 = P(y_1, \ldots, y_k)$, and the IP-sets $(n_\alpha^{(s)})$ where $n_\alpha^{(s)} = n_{\alpha_s}$ for $s = 1, \ldots, k$. We obtain the existence of elements $x_1 = c\,\xi_1$, $x_2 = c\,\xi_2$ where $\xi_1, \xi_2 \in A$, and of numbers $\eta_1 = n_\alpha^{(1)}, \ldots, \eta_k = n_\alpha^{(k)}$ such that

$$x_1 - x_2 = c \cdot (\xi_1 - \xi_2) = P(\eta_1, \ldots, \eta_k).$$

By our definition of the IP-sets $(n_\alpha^{(s)})$, the elements $\eta_1, \ldots, \eta_k$ are mutually distinct. Finally, notice that when $k = 1$ the polynomial $P(y_1)$ can only have finitely many roots, and so the above arguments also apply to $A' = A \setminus \{\text{roots of P}\}$, which is still an IP-large set with positive Banach density.[11] Clearly, in this case $\xi_1 \neq \xi_2$ because $c(\xi_1 - \xi_2) = P(\eta_1) \neq 0$. □

We can now isolate a simple sufficient condition for a Diophantine nonlinear equation to be PR.

Definition 2.8. *A polynomial with integer coefficients is called a* Rado polynomial *if it can be written in the form*

$$c_1 x_1 + \cdots + c_n x_n + P(y_1, \ldots, y_k)$$

where $n \geq 2$, P has no constant term, and there exists a nonempty subset $J \subseteq \{1, \ldots, n\}$ such that $\sum_{j \in J} c_j = 0$.[12]

Notice that, by Rado's Theorem, a *linear* polynomial with integer coefficients is PR if and only if it is a Rado polynomial. We now show that one implication in Rado's theorem (namely, that every Rado polynomial is PR with certain injectivity conditions) can be extended to *all* Rado polynomials.

Theorem 2.9. *Let*

$$R(x_1, \ldots, x_n, y_1, \ldots, y_k) = c_1 x_1 + \ldots + c_n x_n + P(y_1, \ldots, y_k)$$

be a Rado polynomial. Then every ultrafilter $\mathcal{U} \in \overline{K(\odot)} \cap \overline{\mathbb{I}(\oplus)} \cap \mathcal{BD}$ is a PR-witness of R with injectivity $|\{x_1, \ldots, x_n\}| \geq n-1$ and $|\{y_1, \ldots, y_k\}| = k$.

When $n = 2$, every $\mathcal{U} \in \overline{\mathbb{I}(\oplus)} \cap \mathcal{BD}$ satisfies the above property, and one has injectivity $|\{x_1, x_2\}| = 2$ if $k = 1$. Moreover, if $P \neq 0$ is linear then every $\mathcal{U} \in \overline{K(\odot)}$ is a PR-witness of R with full injectivity $|\{x_1, \ldots, x_n, y_1, \ldots, y_k\}| = n + k$.

Notice that the set $\overline{K(\odot)} \cap \overline{\mathbb{I}(\oplus)} \cap \mathcal{BD}$ is nonempty; indeed, it contains all *combinatorially rich ultrafilters* (see [37, Definition 17.1]).

---

[11] This argument does not apply to the general case $k > 1$; indeed, while A' still has positive Banach density, it may no longer be additively IP-large.

[12] It is assumed that the sets of variables $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_k\}$ are disjoint.

Proof. Assume first that $n \geq 3$, and consider the following system:

$$\begin{cases} c_1 z + c_2 x_2 + \ldots + c_n x_n = 0; \\ c_1(w - x_1) = P(y_1, \ldots, y_k); \\ z = w. \end{cases}$$

The first equation is injectively PR by Theorem 1.6 and, since it is homogeneous, it is witnessed by any $\mathcal{U} \in \overline{K(\odot)}$, by Corollary 2.5. Moreover, if $\mathcal{U} \in \overline{\mathbb{I}(\oplus)} \cap \mathcal{BD}$, every $A \in \mathcal{U}$ is additively IP-large and has positive Banach density and so, by Theorem 2.7, the second equation is witnessed by $\mathcal{U}$ with injectivity $|\{y_1, \ldots, y_k\}| = k$. Then, by Lemma 2.1, every $\mathcal{U} \in \overline{K(\odot)} \cap \overline{\mathbb{I}(\oplus)} \cap \mathcal{BD}$ is a witness of the above system with injectivity $|\{z, x_2, \ldots, x_n\}| = n$ and $|\{y_1, \ldots, y_k\}| = k$. By combining, we finally obtain that $\mathcal{U}$ is a witness of the equation

$$c_1 x_1 + c_2 x_2 + \ldots + c_n x_n + P(y_1, \ldots, y_k) = 0$$

with the desired injectivity properties.

When $n = 2$, by the hypothesis of Rado polynomial, one has that $c_1 = -c_2 = c$. In this case, the equation $R = 0$ reduces to the equation in Theorem 2.7, that applies to every set $A \in \mathcal{U} \in \overline{\mathbb{I}(\oplus)} \cap \mathcal{BD}$.

If $P \neq 0$ is linear, then the injective PR of R is given by Theorem 1.6. In this linear case, R is trivially homogeneous and so every $\mathcal{U} \in \overline{K(\odot)}$ is a witness, by Corollary 2.5. $\qquad\square$

For completeness, we now present an ultrafilter proof of a theorem proved by the second named author in [44] by using nonstandard analysis.[13] It is a generalization of a previous result by N. Hindman [32], namely the PR of equations $\sum_{i=1}^{n} x_i = \prod_{j=1}^{k} y_j$ with injectivity $|\{x_1, \ldots, x_n\}| = n$ and $|\{y_1, \ldots, y_k\}| = k$ (for $n > 2$). The particular case $n = 2, k = 2$ of this last result was independently proved by V. Bergelson [3, Theorem 6.1].

Theorem 2.10. *Let* $R(x_1, \ldots, x_n) = c_1 x_1 + \ldots + c_n x_n$ *be a linear Rado polynomial, and let* $F_1, \ldots, F_n \subseteq \{1, \ldots, m\}$. *Then every* $\mathcal{U} \in \mathbb{I}(\odot) \cap \overline{K(\odot)}$

---

[13] Precisely, this is [44, Theorem 3.3]. In that paper it is also proved a (rather technical) generalization that shows also the PR of certain polynomials where some of the variables may have a degree larger than one (see [44, Theorem 4.2]).

is a witness of the partition regularity of the equation[14]

$$\sum_{i=1}^{n} c_i \, x_i \left(\prod_{j \in F_i} y_j \right) = 0$$

with full injectivity when $n > 2$, and with injectivity $|\{x_1, x_2, y_1, \ldots, y_m\}| \geq m + 1$ when $n = 2$.

Proof. Since $\mathcal{U} \in \overline{K(\odot)}$ and the equation $R(x_1, \ldots, x_n) = 0$ is homogeneous, $\mathcal{U}$ witnesses its PR, by Corollary 2.5. Given $A \in \mathcal{U}$, set $B_0 = A$ and inductively define

$$B_k = \{x \in B_{k-1} \mid \{y \in B_{k-1} \mid x \cdot y \in B_{k-1}\} \in \mathcal{U}\}.$$

Clearly, $B_m \subseteq B_{m-1} \subseteq \cdots \subseteq B_1 \subseteq B_0 = A$. Moreover, since $\mathcal{U}$ is multiplicatively idempotent and $B_0 \in \mathcal{U}$, it directly follows by induction that all sets $B_1, \ldots, B_m \in \mathcal{U}$. Since $\mathcal{U}$ is a witness of R, we can pick $a_1, \ldots, a_n \in B_m$ such that $c_1 a_1 + \ldots + c_n a_n = 0$ and where $|\{a_1, \ldots, a_n\}| = n$ if $n > 2$. Now let $\bar{a} = \max\{a_1, \ldots, a_n\}$.

Claim: There exist natural numbers $b_1, \ldots, b_m$ such that:

1. $b_k > \bar{a} \cdot \prod_{j<k} b_j$ for every $k = 1, \ldots, m$,

2. $b_k \in B_{m-k}$ for every $k = 1, \ldots, m$,

3. $a_i \cdot \prod_{j \in G} b_j \in B_{m-\max G}$ for every i and for every set $G \subseteq \{1, \ldots, m\}$.[15]

Let us see first how the thesis follows from the claim. For $i = 1, \ldots, n$, let

$$d_i := a_i \cdot \prod_{j \in F_i^c} b_j.$$

The d_i are pairwise distinct when $n > 2$; indeed, let us assume by contradiction that there are indexes $s \neq t$ such that $d_s = d_t$. If $F_s = F_t$ then $\prod_{j \in F_s^c} b_j = \prod_{j \in F_t^c} b_j$, and this is impossible because it would imply that $a_s = a_t$. If $F_s \neq F_t$, let $\bar{j} = \max F_s^c \triangle F_t^c$, and let $H_1 = \{j \in F_s^c \mid j < \bar{j}\}$,

---

[14] We agree that $\prod_{j \in F} y_j = 1$ when $F = \emptyset$.

[15] We agree that $\prod_{j \in G} b_j = 1$ and $\max G = 0$ when $G = \emptyset$.

18

$H_2 = \{j \in F_t^c \mid j < \bar{j}\}$, and $H_3 = \{j \in F_s^c \mid j > \bar{j}\} = \{j \in F_t^c \mid j > \bar{j}\}$. Without loss of generality, let us assume that $\bar{j} \in F_s^c$. Then

$$d_s = a_s \cdot b_{\bar{j}} \cdot \left(\prod_{j \in H_1} b_j\right)\left(\prod_{j \in H_3} b_j\right) \quad \text{and} \quad d_t = a_t \cdot \left(\prod_{j \in H_2} b_j\right)\left(\prod_{j \in H_3} b_j\right).$$

Since $d_s = d_t$, it follows that

$$a_s \cdot b_{\bar{j}} \cdot \prod_{j \in H_1} b_j = a_t \cdot \prod_{j \in H_2} b_j,$$

which is absurd because, by the definition of the b_j, one has

$$a_s \cdot b_{\bar{j}} \cdot \prod_{j \in H_1} b_j \geq b_{\bar{j}} > \bar{a} \cdot \prod_{j < \bar{j}} b_j \geq a_t \cdot \prod_{j \in H_2} b_j.$$

Elements $d_i, b_j \in A$ are a solution of $\sum_{i=1}^n c_i\, x_i\left(\prod_{j \in F_i} y_j\right) = 0$. Indeed, by the claim, we have that $d_i \in B_{m - \max F_i^c} \subseteq A$ and $b_j \in B_{m-j} \subseteq A$. Moreover,

$$\sum_{i=1}^n c_i\, d_i \left(\prod_{j \in F_i} b_j\right) = \sum_{i=1}^n c_i\, a_i \left(\prod_{j \in F_i^c} b_j\right)\left(\prod_{j \in F_i} b_j\right) = \left(\prod_{j=1}^m b_j\right)\left(\sum_{i=1}^n c_i a_i\right) = 0.$$

Since $|\{b_1, \ldots, b_m\}| = m$ and $|\{d_1, \ldots, d_n\}| = n$ when $n > 2$, the desired injectivity properties follows by noticing that $b_k \neq d_h$ for all k, h. Let us show that no equality $d_h = a_h \cdot \prod_{j \in F_h^c} b_j = b_k$ is possible. If $F_h^c = \emptyset$ then $d_h = a_h < b_k$. Otherwise, let $h_* = \max F_h^c$. If $k > h_*$, then $b_k \geq b_{h_*+1} > d_h$; if $k < h_*$ then $d_h \geq b_{h_*} > b_k$; and if $k = h_*$ then $b_k < a_h b_k \leq d_h$ (here we assumed, without loss of generality, that all $a_h > 1$).

We are left to prove the claim. We define b_k recursively for $k \leq m$.

Let $k = 1$. For every $i = 1, \ldots, n$, the element $a_i \in B_m$, and so

$$C_i = \{y \in B_{m-1} \mid a_i \cdot y \in B_{m-1}\} \in \mathcal{U}.$$

Pick $b_1 \in C_1 \cap \cdots \cap C_n \in \mathcal{U}$ with $b_1 > \bar{a}$. Trivially, $b_1 \in B_{m-1}$. Moreover, for every i, one has $a_i \cdot \prod_{j \in \{1\}} b_j = a_i \cdot b_1 \in B_{m-1} = B_{m-\max\{1\}}$, and $a_i \cdot \prod_{j \in \emptyset} b_j = a_i \in B_m = B_{m-\max \emptyset}$.

At the inductive step, assume that numbers $b_1, \ldots, b_k$ where $k \leq m-1$ have been defined which fulfill the properties of the claim. We want to

19

define b_{k+1}. For every set $G \subseteq \{1, \ldots, k\}$ and for every i, by the inductive hypothesis $a_i \cdot \prod_{j \in G} b_j \in B_{m - \max G}$, and hence

$$C_{G,i} = \left\{ y \in B_{m - \max G - 1} \,\middle|\, a_i \cdot \prod_{j \in G} b_j \cdot y \in B_{m - \max G - 1} \right\} \in \mathcal{U}.$$

Now pick a number $b_{k+1} > \bar{a} \cdot \prod_{j=1}^{k} b_j$ with

$$b_{k+1} \in \bigcap_{i=1}^{n} \left(\bigcap_{G \subseteq \{1, \ldots, k\}} C_{G,i} \right) \in \mathcal{U}.$$

Notice that every $C_{G,i} \subseteq B_{m - \max G - 1} \subseteq B_{m - (k+1)}$, and so $b_{k+1} \in B_{m - (k+1)}$. Now let $G \subseteq \{1, \ldots, k+1\}$. If $G \subseteq \{1, \ldots, k\}$ then, by the inductive hypothesis, $a_i \cdot \prod_{j \in G} b_j \in B_{m - \max G}$ for every i. Now assume $k + 1 \in G$ and let $G' = G \setminus \{k + 1\}$. For every i, by the inductive hypothesis on G', we know that

$$a_i \cdot \prod_{j \in G'} b_j \in B_{m - \max G'} \subseteq B_{m-k},$$

and since $b_{k+1} \in C_{G',i}$, we conclude that

$$a_i \cdot \prod_{j \in G} b_j = a_i \cdot \prod_{j \in G'} b_j \cdot b_{k+1} \in B_{m - \max G' - 1} \subseteq B_{m - k - 1} \subseteq B_{m - \max G}. \qquad \square$$

We are finally ready to state the following theorem, that puts together all that we have proved so far, and further extends the class of nonlinear polynomials proved to be PR.

Theorem 2.11. *Let $\mathfrak{F}$ be the family of functions whose PR on $\mathbb{N}$ is witnessed by at least an ultrafilter $\mathcal{U} \in \mathbb{I}(\odot) \cap \overline{K(\odot)} \cap \overline{\mathbb{I}(\oplus)} \cap \mathcal{BD}$. Then $\mathfrak{F}$ includes:*

1. *Every Rado polynomial*

$$c_1 x_1 + \ldots + c_n x_n + P(y_1, \ldots, y_k)$$

   with injectivity $|\{x_1, \ldots, x_n\}| \geq n - 1$ and $|\{y_1, \ldots, y_k\}| = k$, and with injectivity $|\{x_1, x_2\}| = 2$ when $n = 2$ and $k = 1$, and with full injectivity $|\{x_1, \ldots, x_n, y_1, \ldots, y_k\}| = n + k$ when $P \neq 0$ is linear;

20

2. *Every polynomial of the form*

$$\sum_{i=1}^{n} c_i\, x_i \left(\prod_{j \in F_i} y_j \right)$$

where $\sum_{i=1}^{n} c_i\, x_i$ is a Rado polynomial and sets $F_i \subseteq \{1, \ldots, m\}$, with full injectivity when $n > 2$, and with injectivity $|\{x_1, x_2, y_1, \ldots, y_m\}| \geq m + 1$ when $n = 2$;

3. *Every function f of the form*

$$f(x, y_1, \ldots, y_k) = x - \prod_{i=1}^{k} y_i$$

with full injectivity $|\{x, y_1, \ldots, y_k\}| = k + 1$;

4. *Every function f of the form*

$$f(x, y_1, \ldots, y_k) = x - \prod_{i=1}^{k} y_i^{a_i}$$

with full injectivity $|\{x, y_1, \ldots, y_k\}| = k + 1$, whenever the exponents $a_i \in \mathbb{Z}$ satisfy $\sum_{i=1}^{n} a_i = 1$.

Moreover, the family $\mathfrak{F}$ satisfies the following closure properties:

(i) *Assume that $f(z, y_1, \ldots, y_k) \in \mathfrak{F}$ with injectivity $|\{y_{i_1}, \ldots, y_{i_p}\}| \geq s$ and that $z - g(x_1, \ldots, x_n) \in \mathfrak{F}$ with injectivity $|\{x_{j_1}, \ldots, x_{j_q}\}| \geq t$. Then $f(g(x_1, \ldots, x_n), y_1, \ldots, y_k) \in \mathfrak{F}$ with injectivity $|\{x_{j_1}, \ldots, x_{j_q}\}| \geq t - 1$ and $|\{y_{i_1}, \ldots, y_{i_p}\}| \geq s - 1$.*

(ii) *Assume that the homogeneous function $f(x_1, \ldots, x_n) \in \mathfrak{F}$ with injectivity $|\{x_{i_1}, \ldots, x_{i_p}\}| \geq s$. Then $f(1/x_1, \ldots, 1/x_n) \in \mathfrak{F}$ with injectivity $|\{x_{i_1}, \ldots, x_{i_p}\}| \geq s$.*

Proof. (1) and (2) are given by Theorems 2.9 and 2.10, respectively.

(3). Since $\mathcal{U} \in \overline{\mathbb{I}(\odot)}$, every $A \in \mathcal{U}$ is multiplicatively IP-large, and so it contains injective solutions to the equation $x = \prod_{i=1}^{n} y_i$.

(4). Notice first that $\prod_{i=1}^{n} y_i^{a_i} = x$ is injectively PR. Indeed, given a finite coloring $\mathbb{N} = C_1 \cup \ldots \cup C_r$, one considers the partition as given by the sets

$D_s = \{n \mid 2^n \in C_s\}$. By Theorem 1.6, one finds pairwise distinct monochromatic $\eta_1, \ldots, \eta_n, \xi \in D_s$ such that $\sum_{i=1}^{n} a_i \eta_i = \xi$. Then $2^{\eta_1}, \ldots, 2^{\eta_n}, 2^{\xi} \in C_s$ are an injective monochromatic solution of $\prod_{i=1}^{n} y_i^{a_i} = x$. Now, the function $f(x, y_1, \ldots, y_k) = x - \prod_{i=1}^{k} y_i^{a_i}$ is homogeneous since $\sum_{i=1}^{n} a_i = 1$ and so, by Corollary 2.5, every ultrafilter $\mathcal{U} \in \overline{\mathbb{M}(\odot)} \subseteq \overline{K(\odot)}$ is an injective witness.

(i). It directly follows from Lemma 2.1.

(ii). By Theorem 2.6, we have that $f(1/x_1, \ldots, 1/x_n)$ is PR with injectivity $|\{x_{j_1}, \ldots, x_{j_q}\}| \geq t$. Since the function is homogeneous, such a PR is witnessed by all ultrafilters $\mathcal{U} \in \overline{K(\odot)}$, by Corollary 2.5. $\qquad \square$

Let us now give some examples of equations whose PR is obtained by applying Theorem 2.11.

Example 2.12. Consider the injectively PR polynomials $x_1 x_2 = z^2$, $y_1 + y_2 = y_3$ and $t_1 - t_2 = t_3$, which are in $\mathfrak{F}$. Then, by the closure property (i), it follows that the following equations are PR with full injectivity.

- $x(y_1 + y_2) = z^2$,

- $x(t_1 - t_2) = z^2$,

- $x_1 x_2 = (y_1 + y_2)^2$,

- $x_1 x_2 = (t_1 - t_2)^2$,

- $x(t_1 - t_2) = (y_1 + y_2)^2$,

- $x(y_1 + y_2) = (t_1 - t_2)^2$,

- $(y_1 + y_2)(t_1 - t_2) = z^2$,

Example 2.13. The example above generalizes as follows: Let $n, m \in \mathbb{N}$ and assume that, for every $i \leq n$, $j \leq m$, the equations

$$x_{i,1} = \sum_{h=1}^{r_i} c_{i,h} x_{i,h}, \quad y_{j,1} = \sum_{k=1}^{s_j} d_{j,k} y_{j,k}$$

are PR. Let $a_1, \ldots, a_n, b_1, \ldots, b_m$ be such that $\sum_{i=1}^{n} a_i = \sum_{j=1}^{m} b_j$ and consider the homogeneous PR equation $\prod_{i=1}^{n} t_i^{a_i} = \prod_{j=1}^{m} z^{b_j}$. All these equations

are PR and homogeneous and therefore, by the closure property (i), also

$$\prod_{i=1}^{n} \left(\sum_{h=1}^{r_i} c_{i,h} x_{i,h} \right)^{a_i} = \prod_{j=1}^{m} \left(\sum_{k=1}^{s_j} d_{j,k} y_{j,k} \right)^{b_j}$$

is PR with full injectivity.

Example 2.14. Notice that all the equations considered in the previous examples are homogeneous. Therefore by the closure property (ii) applied to some of the equations of Example 2.12 we obtain, *e.g.*, that the following equations are PR with full injectivity.

- $x^2 y_1 y_2 = z^2 (y_1 + y_2)$;

- $x^2 t_1 t_2 = z^2 (t_2 - t_1)$;

- $(y_1 + y_2)(t_2 - t_1) z^2 = y_1 y_2 t_1 t_2$.

Example 2.15. Let $c_1 x_1 + \ldots + c_n x_n$ be a Rado polynomial with $n \geq 3$, and assume also that for $j = 1, \ldots, m$ the following is a Rado polynomial:

$$b_j z_j - \sum_{h=1}^{n_j} a_{h,j} z_{h,j}$$

For every $F_1, \ldots, F_n \subseteq \{1, \ldots, m\}$, the following polynomial is injectively PR:

$$\sum_{i=1}^{n} c_i x_i \left(\prod_{j \in F_i} y_j \right).$$

All the above polynomials are in $\mathfrak{F}$, so we can apply the closure property (i) with the conditions $y_j = z_j$ for $j = 1, \ldots, m$, and deduce that the function

$$\sum_{i=1}^{n} c_i x_i \left(\prod_{j \in F_i} \frac{1}{b_j} \left(\sum_{h=1}^{n_j} a_{h,j} z_{h,j} \right) \right) = 0$$

is in $\mathfrak{F}$, and it is PR with full injectivity. For example, in this way one obtains the PR of

$$x_1 (2y_1 + y_2) + x_2 y_3 - x_3 (2y_1 + y_2) y_3 = 0.$$

23

Example 2.16. For every $n \in \mathbb{N}$, the function $u - v - z^n$ is in $\mathfrak{F}$ with full injectivity; moreover, for every $k \geq 2$ the function $x = \prod_{j=1}^{k} x_j$ is in $\mathfrak{F}$ with full injectivity. Therefore, for every $h, k \geq 2$ we can apply the closure property (i) of $\mathfrak{F}$ to the system

$$
\begin{cases}
u - v = z^n; \\
x = \prod_{j=1}^{h} x_j; \\
y = \prod_{j=1}^{k} y_j; \\
x = t; \\
y = v.
\end{cases}
$$

This shows that the equation

$$
\prod_{j=1}^{h} x_j - \prod_{j=1}^{k} y_j = z^n
$$

is in $\mathfrak{F}$ with full injectivity.[16]

Let us notice that for $n = h = k = 2$, Example 2.16 reduces to

$$
x_1 x_2 = y_1 y_2 + z^2.
$$

Such an equation can be considered as a modified version of the Pythagorean equation $x^2 = y^2 + z^2$.

The range of Theorem 2.11 includes most but not all of the known PR polynomials.

Example 2.17. The polynomial $P(x_1, x_2, x_3) = x_1 x_2 - 2x_3$ is PR but it does not belong to the family $\mathfrak{F}$ of Theorem 2.11.[17]

Given a finite coloring $\mathbb{N} = C_1 \cup \ldots \cup C_r$, consider the coloring $\mathbb{N} = C_1' \cup \ldots \cup C_r'$ where $C_i' = \{n \in \mathbb{N} \mid 2^n \in C_i\}$. By the non-homogenous part of Rado's Theorem, the polynomial $y_1 + y_2 - y_3 - 1$ is PR. Let $a, b, c \in C_i'$ be monochromatic numbers such that $a + b - c - 1 = 0$. Then $2^a, 2^b, 2^c \in C_i$ are monochromatic solutions $P(2^a, 2^b, 2^c) = 0$. This shows that P is PR.

---

[16] However, as it will be shown in Section 3, the equation $x^n - y^m = z^k$ is not PR if $n \notin \{m, k\}$.

[17] Other examples of PR polynomials not included in the family $\mathfrak{F}$ of Theorem 2.11, can be found in [44].

Now assume by contradiction that $P(x_1, x_2, x_3) \in \mathfrak{F}$. Notice that the polynomial $x_3 - y_1 y_2 \in \mathfrak{F}$, and so, by the closure property (4), also $P(x_1, x_2, y_1 y_2) := x_1 x_2 - 2y_1 y_2$ would belong to $\mathfrak{F}$. This is not possible because $x_1 x_2 - 2y_1 y_2$ is not PR. To see this, consider the partition $\mathbb{N} = C_1 \cup C_2$ where C_1 is the set of natural numbers n such that the greatest exponent k with $2^k \,|\, n$ is even. It is easily verified that if $a_1, a_2, b_1, b_2 \in C_i$ then $a_1 a_2 \neq 2b_1 b_2$ for $i = 1, 2$.

The previous example shows that being Rado is not a necessary condition for a polynomial to be PR.

3 Necessary conditions for PR

In this section we isolate necessary conditions for a Diophantine equation to be PR. Instead of working in the space of ultrafilters $\beta \mathbb{N}$ as done in the previous section, here we will use a related but different non-elementary technique, namely *nonstandard analysis* on the *hypernatural numbers* $^*\mathbb{N}$.

It is worth noticing that in the last years nonstandard methods have fruitfully applied to prove combinatorial properties of integers, both in the direction of density results and of Ramsey results (see, *e.g.*, [38, 13, 44, 16, 17]).

We will assume the reader to be familiar with the fundamental notions and results of nonstandard analysis, namely *hyper-extensions* (or *nonstandard extensions*) of sets and functions, the *transfer* principle, the *overspill* principle, the κ-*enlargement* and the κ-*saturation* properties. All such topics can be found in any of the monographies on nonstandard analysis (see, *e.g.*, the books [21, 42]).

3.1 u-equivalence

We will work in a $\mathfrak{c}^+$-saturated extension of $\mathbb{N}$. In addition to the fundamental principles of nonstandard analysis, our proofs will also use properties of the relation of u-*equivalence* on hypernatural numbers, as introduced by the first named author in [14]. (See also [15], where u-equivalent pairs are named *indiscernible*, and [13, 43], where algebraic properties of u-equivalence are proved by means of iterated hyper-extensions.)

Definition 3.1. *Two hypernatural numbers* $\xi, \zeta \in {}^*\mathbb{N}$ *are* u-equivalent *if they cannot be distinguished by any hyper-extension,* i.e. *if for every* $A \subseteq \mathbb{N}$

one has either $\xi, \xi' \in {}^*A$ or $\xi, \xi' \notin {}^*A$.

The "u" in u-equivalence stands for "ultrafilter". Indeed, to every $\xi \in {}^*\mathbb{N}$ is associated the ultrafilter $\mathfrak{U}_\xi = \{A \subseteq \mathbb{N} \mid \xi \in {}^*A\}$, and $\xi \underset{u}{\sim} \zeta$ means that the associated ultrafilters coincide: $\mathfrak{U}_\xi = \mathfrak{U}_\zeta$.

We will use the following properties (see [14]).

(U1) If $k \in \mathbb{N}$ is finite and $\xi \underset{u}{\sim} k$ then $\xi = k$;

(U2) For every $f : \mathbb{N} \to \mathbb{N}$, if $\xi \underset{u}{\sim} \zeta$ then ${}^*f(\xi) \underset{u}{\sim} {}^*f(\zeta)$;

(U3) For every $f : \mathbb{N} \to \mathbb{N}$, if ${}^*f(\xi) \underset{u}{\sim} \xi$ then ${}^*f(\xi) = \xi$.

(U4) If $\xi \underset{u}{\sim} \zeta$ and $\xi < \zeta$ then $\zeta - \xi$ is infinite.

In the language of nonstandard analysis, we have the following counterpart of Proposition 1.8 (see also [43, Theorem 2.2.9], which gives a more general version of this result).

Proposition 3.2. *A function $f(x_1, \ldots, x_n)$ is partition regular with injectivity $|\{x_{i_1}, \ldots, x_{i_p}\}| \geq s$ if and only if there exist hypernatural numbers $\xi_1 \underset{u}{\sim} \ldots \underset{u}{\sim} \xi_n$ with ${}^*f(\xi_1, \ldots, \xi_n) = 0$ and $|\{\xi_{i_1}, \ldots, \xi_{i_p}\}| \geq s$.*

Proof. Assume first that there exist $\xi_1 \underset{u}{\sim} \ldots \underset{u}{\sim} \xi_n$ with the above properties, and consider the ultrafilter $\mathcal{U} = \mathfrak{U}_{\xi_1} = \ldots = \mathfrak{U}_{\xi_n}$. For every $A \in \mathcal{U}$, the elements ξ_i witness that the following is true:

$$\exists\, y_1, \ldots, y_n \in {}^*A \text{ s.t. } {}^*f(y_1, \ldots, y_n) = 0 \,\&\, |\{y_{i_1}, \ldots, y_{i_p}\}| \geq s.$$

By *transfer*, we obtain the existence of elements $a_1, \ldots, a_n \in A$ such that $f(a_1, \ldots, a_n) = 0$ and $|\{a_{i_1}, \ldots, a_{i_p}\}| \geq s$, thus showing that $\mathcal{U}$ is a PR-witness of f with injectivity $|\{x_{i_1}, \ldots, x_{i_p}\}| \geq s$.

Conversely, pick a PR-witness $\mathcal{U}$ of f with injectivity $|\{x_{i_1}, \ldots, x_{i_p}\}| \geq s$. Then for every $A \in \mathcal{U}$ the following set is nonempty:

$$\Gamma(A) \;=\; \left\{ (a_1, \ldots, a_n) \in A^n \mid f(a_1, \ldots, a_n) = 0 \,\&\, |\{a_{i_1}, \ldots, a_{i_p}\}| \geq s \right\}.$$

Since $\Gamma(A_1) \cap \ldots \cap \Gamma(A_k) = \Gamma(A_1 \cap \ldots \cap A_k)$, the family $\{\Gamma(A) \mid A \in \mathcal{U}\}$ has the *finite intersection property*, and hence, by $\mathfrak{c}^+$-*enlargement* (which is implied by $\mathfrak{c}^+$-saturation), we can pick $(\xi_1, \ldots, \xi_n) \in \bigcap_{A \in \mathcal{U}} {}^*\Gamma(A)$. It is readily verified that $\mathfrak{U}_{\xi_1} = \ldots = \mathfrak{U}_{\xi_n} = \mathcal{U}$, and that $\xi_1, \ldots, \xi_n$ satisfy the desired properties. $\qquad\square$

In particular, we will use the following characterization.

Corollary 3.3. *A function $f(x_1, \ldots, x_n)$ is non-trivially PR if and only if there exist infinite hypernatural numbers $\xi_1 \underset{u}{\sim} \ldots \underset{u}{\sim} \xi_n$ with ${}^*f(\xi_1, \ldots, \xi_n) = 0$.*

Proof. By the previous proposition, we can pick elements $\xi_1 \underset{u}{\sim} \ldots \underset{u}{\sim} \xi_n$ with ${}^*f(\xi_1, \ldots, \xi_n) = 0$ and $|\{\xi_1, \ldots, \xi_n\}| > 1$. If one of the ξ_i equals a finite $k \in \mathbb{N}$, then by property (U1) we would have $\xi_j = k$ for all $j = 1, \ldots, n$, a contradiction. $\square$

As a first easy example of application of u-equivalence, let us prove the following fact.[18]

Proposition 3.4. *Let $f : \mathbb{N}^n \to \mathbb{N}$ and let $\varphi : \mathbb{N} \to \mathbb{N}$. If $f(\varphi(x_1), \ldots, \varphi(x_n))$ is PR then $f(x_1, \ldots, x_n)$ is PR. Moreover, if φ is onto, then one has the equivalence: $f(x_1, \ldots, x_n)$ is PR $\Leftrightarrow$ $f(\varphi(x_1), \ldots, \varphi(x_n))$ is PR.*

Proof. Pick $\xi_1 \underset{u}{\sim} \ldots \underset{u}{\sim} \xi_n$ such that ${}^*f({}^*\varphi(\xi_1), \ldots, {}^*\varphi(\xi_n)) = 0$, and let $\eta_i = {}^*\varphi(\xi_i)$. Then $\eta_1 \underset{u}{\sim} \ldots \underset{u}{\sim} \eta_n$ and trivially ${}^*f(\eta_1, \ldots, \eta_n) = 0$. If φ is onto, pick $\psi : \mathbb{N} \to \mathbb{N}$ such that $\varphi \circ \psi$ is the identity, and consider $\eta_i = {}^*\psi(\xi_i)$. Then $\eta_1 \underset{u}{\sim} \ldots \underset{u}{\sim} \eta_n$ are such that ${}^*f({}^*\varphi(\eta_1), \ldots, {}^*\varphi(\eta_n)) = {}^*f(\xi_1, \ldots, \xi_n) = 0$. $\square$

When dealing with polynomials in several variables, it is convenient to use the multi-index notation. Let us fix the terminology.

- An n-dimensional *multi-index* is an n-tuple $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$;

- $\alpha \leq \beta$ means that $\alpha_i \leq \beta_i$ for all $i = 1, \ldots, n$;

- $\alpha < \beta$ means that $\alpha \leq \beta$ and $\alpha \neq \beta$;

- If $\mathbf{x} = (x_1, \ldots, x_n)$ is vector and $\alpha = (\alpha_1, \ldots, \alpha_n)$ is a multi-index, the product $\prod_{i=1}^n x_i^{\alpha_i}$ is denoted by $\mathbf{x}^\alpha$;

- The *length* of a multi-index $\alpha = (\alpha_1, \ldots, \alpha_n)$ is $|\alpha| = \sum_{i=1}^n \alpha_i$;

- A set I of n-dimensional multi-indexes having all the same length is called *homogeneous*;

---

[18] This basic property was first pointed out by H. Lefmann [41] for bijective functions f, in the context of rings. Indeed, Proposition 3.4 also holds if one replaces $\mathbb{N}$ with an arbitrary ring R (of course in this case "PR on $\mathbb{N}$" is replaced by "PR on R").

- Polynomials $P \in \mathbb{Z}[x_1, \ldots, x_n]$ are written in the form $P(\mathbf{x}) = \sum_\alpha c_\alpha \mathbf{x}^\alpha$ where α are multi-indexes;

- The *support* of P is the finite set $\mathrm{supp}(P) = \{\alpha \mid c_\alpha \neq 0\}$;

- A polynomial $P(\mathbf{x}) = \sum_\alpha c_\alpha \mathbf{x}^\alpha$ is *homogeneous* if $\mathrm{supp}(P)$ is a homogeneous set of indexes.

Definition 3.5. Let $P(\mathbf{x}) = \sum_\alpha c_\alpha \mathbf{x}^\alpha \in \mathbb{Z}[x_1, \ldots, x_n]$. We say that a multi-index $\alpha \in \mathrm{supp}(P)$ is *minimal* if there are no $\beta \in \mathrm{supp}(P)$ with $\beta < \alpha$. The notion of *maximal* multi-index is defined similarly.

A nonempty set $J \subseteq \mathrm{supp}(P)$ is called a *Rado set of indexes* if for every $\alpha, \beta \in J$ there exists a nonempty $\Lambda \subseteq \{1, \ldots, n\}$ with $\sum_{i \in \Lambda} \alpha_i = \sum_{i \in \Lambda} \beta_i$.

Notice that every singleton $\{\alpha\} \subseteq \mathrm{supp}(P)$ is trivially a Rado set. When $P(x_1, \ldots, x_n) = c_1 x_1 + \ldots + c_n x_n$ is a linear polynomial with no constant term, then we can write $P = \sum_{s=1}^n c_s \mathbf{x}^{\alpha(s)}$ where $\alpha(s)$ is the multi-index where the s-th entry is 1, and all other entries are 0. In this case, every nonempty $J \subseteq \mathrm{Supp}(P) = \{\alpha(1), \ldots, \alpha(n)\}$ is a Rado set of both minimal and maximal indexes.

Theorem 3.6. *Let $P(\mathbf{x}) = \sum_\alpha c_\alpha \mathbf{x}^\alpha \in \mathbb{Z}[x_1, \ldots, x_n]$ be a polynomial with no constant term. Suppose there exists a prime p such that:*

1. *$\sum_\alpha c_\alpha z^{|\alpha|} \equiv 0 \mod p$ has no solutions $z \not\equiv 0$;*

2. *For every Rado set J of minimal indexes, $\sum_{\alpha \in J} c_\alpha z^{|\alpha|} \equiv 0 \mod p$ has no solutions $z \not\equiv 0$.*

Then $P(\mathbf{x})$ is not PR, except possibly for constant solutions $x_1 = \ldots = x_n$.

Proof. By contradiction, let us suppose that the polynomial $P(\mathbf{x})$ is non-trivially PR, and pick infinite $\xi_1 \underset{u}{\sim} \ldots \underset{u}{\sim} \xi_n$ such that

$$P(\boldsymbol{\xi}) = \sum_\alpha c_\alpha \boldsymbol{\xi}^\alpha = 0.$$

Pick a prime p as given by the hypothesis, and write the numbers ξ_i in the following form:

$$\xi_i = a_i + \zeta_i p^{\tau_i}$$

28

where $0 \le a_i \le p-1$, where ζ_i is not divisible by p, and where $\tau_i \ge 1$. Denote by $b_i \in \{1, \ldots, p-1\}$ the number such that $\zeta_i \equiv b_i \mod p$.

Let $f : \mathbb{N} \to \{0, 1, \ldots, p-1\}$ be the function where $f(m) \equiv m \mod p$; let $g : \mathbb{N} \to \mathbb{N}$ be the function where $g(m)$ is the greatest exponent of p that divides $m - f(m)$; and let $h : \mathbb{N} \to \{1, \ldots, p-1\}$ be the function where $h(m) \equiv (m - f(m))/p^{g(m)} \mod p$. Notice that $^*f(\xi_i) = a_i$, $^*g(\xi_i) = \tau_i$ and $^*h(\xi_i) = b_i$. So, the u-equivalences $\xi_1 \underset{u}{\sim} \ldots \underset{u}{\sim} \xi_n$ imply that $a_1 \underset{u}{\sim} \ldots \underset{u}{\sim} a_n$, $\tau_1 \underset{u}{\sim} \ldots \underset{u}{\sim} \tau_n$, and $\zeta_1 \underset{u}{\sim} \ldots \underset{u}{\sim} \zeta_n$. Since finite u-equivalent numbers are necessarily equal, there exist $0 \le a \le p-1$ and $1 \le b \le p-1$ such that $a_i = a$ and $b_i = b$ for all i. Now,

$$0 \;=\; P(\boldsymbol{\xi}) \;\equiv\; \sum_\alpha c_\alpha a^{|\alpha|} \mod p,$$

and hence, by the hypothesis (1), it must be $a = 0$. In consequence, with the multi-index notation,

$$\boldsymbol{\xi}^\alpha \;=\; \boldsymbol{\zeta}^\alpha \cdot p^{\sum_{i=1}^n \alpha_i \tau_i}$$

where $\boldsymbol{\zeta}^\alpha \equiv b^{|\alpha|} \not\equiv 0 \mod p$. Now let $\sigma = \min \{\sum_{i=1}^n \alpha_i \tau_i \mid \alpha \in \mathrm{supp}(P)\}$, and let $J = \{\alpha \mid \sum_{i=1}^n \alpha_i \tau_i = \sigma\}$. We have that

$$0 \;=\; \sum_\alpha c_\alpha \boldsymbol{\xi}^\alpha \;=\; p^\sigma \cdot \left(\sum_{\alpha \in J} c_\alpha \boldsymbol{\zeta}^\alpha + \sum_{\beta \notin J} c_\beta \, \boldsymbol{\zeta}^\beta \, p^{(\sum_{i=1}^n \beta_i \tau_i) - \sigma} \right).$$

Then

$$\sum_\alpha c_\alpha \boldsymbol{\zeta}^\alpha \;\equiv\; \sum_{\alpha \in J} c_\alpha b^{|\alpha|} \;\equiv\; 0 \mod p.$$

This shows that the equation

$$\sum_{\alpha \in J} c_\alpha z^{|\alpha|} \;\equiv\; 0 \mod p$$

has the solution $b \not\equiv 0 \mod p$. We will reach a contradiction with hypothesis (2), by showing that J is a Rado set of minimal indexes. Notice first that J only contains minimal indexes; indeed, if $\beta < \alpha \in J$ then $\sigma - \sum_{i=1}^n \beta_i \tau_i = \sum_{i=1}^n (\alpha_i - \beta_i) \tau_i > 0$ since all $\tau_i \ge 1$, and so $\beta \notin \mathrm{supp}(P)$. Let us now prove that J is a Rado set. Take any two distinct indexes $\alpha, \beta \in J$. (If J is a singleton, the thesis is trivial.) Then $\sum_{i=1}^n \beta_i \tau_i - \sum_{i=1}^n \alpha_i \tau_i = \sigma - \sigma = 0$. Since $\tau_1 \underset{u}{\sim} \ldots \underset{u}{\sim} \tau_n$, by the nonstandard characterization, the equation $\sum_{i=1}^n (\beta_i - \alpha_i) y_i = 0$ is PR. In consequence, by Rado's theorem, there exists a nonempty $\Lambda \subseteq \{1, \ldots, n\}$ such that $\sum_{i \in \Lambda} (\beta_i - \alpha_i) = 0$, as desired. $\qquad \square$

The range of Diophantine equations covered by Theorem 3.6 is quite large. Two easy examples are the following.

Example 3.7. Let $P(x_1, x_2, x_3) = x_1^2 x_2 - 2x_3$. Pick any prime number p with $p \equiv 3$ or $p \equiv 5 \mod 8$, so that 2 is not a quadratic residue modulo p. Then condition (1) of Theorem 3.6 is satisfied because $z^3 - 2z \equiv 0$ iff $z \equiv 0$, and also condition (2) is easily verified. Since it has no constant solutions $x_1 = x_2 = x_3$, we can conclude that $P(x_1, x_2, x_3)$ is not PR.

Example 3.8. Let $P(x_1, \ldots, x_p, y, z) = \prod_{i=1}^{p} x_i - py^2 + z$, where $p > 2$ is a prime number. Conditions (1) and (2) of Theorem 3.6 are satisfied as immediate consequences of Fermat's Little Theorem, and hence P is not PR.

As a particular case of Theorem 3.6, we obtain a result about homogeneous equations, first proved by the second named author in [43].

Corollary 3.9. *Let* $P(\mathbf{x}) = \sum_\alpha c_\alpha \mathbf{x}^\alpha \in \mathbb{Z}[x_1, \ldots, x_n]$ *be an homogeneous polynomial. If for every nonempty* $J \subseteq supp(P)$ *one has* $\sum_{\alpha \in J} c_\alpha \neq 0$, *then* $P(\mathbf{x})$ *is not PR.*

Proof. If d is the degree of $P(\mathbf{x})$, then for every prime number $p > \sum_\alpha |c_\alpha|$, we have that $\sum_{\alpha \in J} c_\alpha z^{|\alpha|} = \left(\sum_{\alpha \in J} c_\alpha \right) z^d \equiv 0 \mod p$ if and only if $z \equiv 0 \mod p$, and so condition (1) of Theorem 3.6 is satisfied. Notice that the hypothesis directly implies that also condition (2) holds, and so we can conclude that $P(\mathbf{x})$ is not PR. $\square$

While the above corollary provides a necessary condition for homogeneous Diophantine equations to be PR, let us mention that H. Leifmann [41, Fact 2.8] isolated a sufficient condition for a special class of homogeneous quadratic equations to be PR.[19]

Another necessary condition for PR applies when every monomial of $P(\mathbf{x})$ contains a single variable, *i.e.* when P has the form $P_1(x_1) + \ldots + P_n(x_n)$.

As the multi-index notation would make the statement of the following theorem less transparent, we switch back the the usual notation for one-variable polynomials.

---

[19] Precisely, $\sum_{i=1}^{n} c_i x_i^2$ is PR if there exists a nonempty $J \subseteq \{1, \ldots, n\}$ and there exist numbers $a \in \mathbb{N}$ and $b_i \in \mathbb{Z}$ such that: (1) $\sum_{i \in J} c_i = 0$; (2) $\sum_{i \in J} b_i c_i = 0$; (3) $\sum_{i \in J} b_i^2 c_i + a^2 \sum_{i \notin J} c_i = 0$.

Theorem 3.10. *For every $i = 1, \ldots, n$ let $P_i(x_i) = \sum_{s=1}^{d_i} c_{i,s} x_i^s$ be a polynomial of degree d_i in the variable x_i with no constant term. If the Diophantine equation*

$$\sum_{i=1}^{n} P_i(x_i) = 0$$

is PR then the following "Rado's condition" is satisfied:

- *"There exists a nonempty set $J \subseteq \{1, \ldots, n\}$ such that $d_i = d_j$ for every $i, j \in J$, and $\sum_{j \in J} c_{j,d_j} = 0$."*

Proof. For every i, let $\Lambda(i) = \{s \mid c_{i,s} \neq 0\}$ be the support of $P_i(x_i)$, and for every s, let $\Gamma(s) = \{i \mid c_{i,s} \neq 0\}$. If we denote by

$$P(\mathbf{x}) \;=\; \sum_{i=1}^{n} P_i(x_i) \;=\; \sum_{i=1}^{n} \sum_{s \in \Lambda(i)} c_{i,s} x_i^s,$$

by the nonstandard characterization of non-trivial PR, we can pick infinite $\xi_1 \underset{u}{\sim} \ldots \underset{u}{\sim} \xi_n$ such that $P(\boldsymbol{\xi}) = 0$. Now fix any finite number $p \geq 2$, and write the numbers ξ_i in base p:

$$\xi_i \;=\; \sum_{t=0}^{\tau_i} a_{i,t} \, p^{\tau_i - t}$$

where $0 \leq a_{i,t} \leq p - 1$ and $a_{i,0} \neq 0$. In particular, $p^{\tau_i} \leq \xi_i < p^{\tau_i + 1}$.

Let $s_* \tau_* = \max\{s\, \tau_i \mid i \in \Gamma(s)\}$. Let us observe that the values of s_* (and, henceforth, of τ_*) are uniquely determined: if there exist $s_1 \neq s_2, \tau_{i_1} \neq \tau_{i_2}$ such that $s_1 \tau_{i_1} = s_2 \tau_{i_2}$ then, as $\tau_{i_1} \underset{u}{\sim} \tau_{i_2}$, we deduce that the equation

$$s_1 x - s_2 y = 0$$

is injectively PR, and this is false by Rado's theorem. Notice also that $d_i = s_*$ for every $i \in \Gamma(s_*)$, by the maximality of $s_* \tau_*$.

Now let $I_* = \{i \in \Gamma(s_*) \mid \tau_i = \tau_*\}$, and decompose $P(\boldsymbol{\xi}) = \Theta + \Psi + \Phi$, where:

- $\Theta = \sum_{i \in I_*} c_{i,s_*} \, \xi_i^{s_*}$;

- $\Psi = \sum_{i \in \Gamma(s_*) \setminus I_*} c_{i,s_*} \, \xi_i^{s_*}$;

- $\Phi = \sum_{s \neq s_*} \sum_{i \in \Gamma(s)} c_{i,s} \, \xi_i^s$.

31

For numbers $\xi, \xi' \in {}^*\mathbb{N}$, in the sequel we will write $\xi \lll \xi'$ to mean that $\xi' - \xi$ is infinite.

Lemma 3.11.

1. $\Theta = \left(\sum_{i \in I_*} c_{i,s_*} \right) \zeta + \Theta'$ for suitable $\zeta \geq p^{s_* \tau_*}$ and $|\Theta'| \lll p^{s_* \tau_*}$.

2. $|\Psi| \lll p^{s_* \tau_*}$.

3. $|\Phi| \lll p^{s_* \tau_*}$.

Since $P(\boldsymbol{\xi}) = \Theta + \Psi + \Phi = 0$, the above inequalities imply that the sum of coefficients $\sum_{i \in I_*} c_{i,s_*} = 0$. We claim that $J = I_*$ is the desired set of indexes. In fact, I_* is trivially nonempty; moreover, $d_i = d_j = s_*$ for all $i, j \in J$; and $\sum_{j \in J} c_{j,d_j} = \sum_{j \in J} c_{j,s_*} = 0$.

We are left to prove the Lemma; let us start with some preparatory work.

Let $\varphi : \mathbb{N} \to \mathbb{N}_0$ be the function where $p^{\varphi(m)} \leq m < p^{\varphi(m)+1}$; and for every $t \in \mathbb{N}_0$, let $\psi_t(m) : \mathbb{N} \to \{0, 1, \ldots, p-1\}$ be the function where $\psi_t(m)$ is the $(t+1)$-th digit from the left when m is written in base p. Then ${}^*\varphi(\xi_i) = \tau_i$ and ${}^*\psi_t(\xi_i) = a_{i,t}$, and the u-equivalences $\xi_1 \underset{u}{\sim} \ldots \underset{u}{\sim} \xi_n$ imply that $\tau_1 \underset{u}{\sim} \ldots \underset{u}{\sim} \tau_n$ and $a_{1,t} \underset{u}{\sim} \ldots \underset{u}{\sim} a_{n,t}$. Since finite u-equivalent numbers are necessarily equal, it must be $a_{1,t} = \ldots = a_{n,t}$. Then, by *overspill*, there exists an infinite $\nu \in {}^*\mathbb{N}$ and numbers $b_t \in \{0, \ldots, p-1\}$ for $t \leq \nu$ such that $a_{i,t} = b_t$ for every $i = 1, \ldots, n$ and for every $t \leq \nu$. Let us denote by

$$\zeta_i = \sum_{t=0}^{\nu} b_t \, p^{\tau_i - t}.$$

We will use the following decomposition:

- *For every $a \in \mathbb{N}$ one has $\xi_i^a = \zeta_i^a + \vartheta_{i,a}$ where $p^{a\tau_i} \leq \zeta_i^a \leq \xi_i^a < p^{a\tau_i + a}$ and $\vartheta_{i,a} \lll p^{a\tau_i}$.*

Since $p^{\tau_i} \leq \zeta_i \leq \xi_i < p^{\tau_i + 1}$, it directly follows that $p^{a\tau_i} \leq \zeta_i^a \leq \xi_i^a < p^{a\tau_i + a}$; besides, the difference $\eta_i = \xi_i - \zeta_i = \sum_{t=\nu+1}^{\tau_i} a_{i,t} p^{\tau_i - t} < p^{\tau_i - \nu}$. Now, $\xi_i^a = (\zeta_i + \eta_i)^a = \zeta_i^a + \vartheta_{i,a}$ where $\vartheta_{i,a} = \sum_{j=1}^{a} \binom{a}{j} \zeta_i^{a-j} \eta_i^j$. Pick a large enough

$\ell_i \in \mathbb{N}$ so that $\binom{a}{j} < p^{\ell_i}$ for all j. Then

$$\vartheta_{i,a} < p^{\ell_i} \sum_{j=1}^{a} \zeta_i^{a-j} \eta_i^j < p^{\ell_i} \sum_{j=1}^{a} \left(p^{\tau_i+1}\right)^{a-j} \cdot \left(p^{\tau_i-\nu}\right)^j =$$

$$= p^{\ell_i} \sum_{j=1}^{a} p^{a\tau_i+a-j(\nu+1)} < p^{2\ell_i} p^{a\tau_i+a-\nu-1} = p^{a\tau_i-(\nu-2\ell_i-a+1)} \lll p^{a\tau_i}.$$

Indeed, since ν is infinite, also $\nu - 2\ell_i - a + 1$ is infinite.

We are now ready to prove points (1), (2) and (3) of Lemma 3.11.

1. With the notation introduced above,

$$\Theta = \sum_{i \in I_*} c_{i,s_*} \zeta_i^{s_*} + \sum_{i \in I_*} c_{i,s_*} \vartheta_{i,s_*}.$$

For every $i \in I_*$, $\zeta_i = \sum_{t=0}^{\nu} b_t p^{\tau_* - t} = \zeta_*$. By the above estimates we know that $\zeta_* \geq p^{\tau_*}$ and $\vartheta_{i,s_*} \lll p^{s_* \tau_i} = p^{s_* \tau_*}$. Then $\sum_{i \in I_*} c_{i,s_*} \zeta_i^{s_*} = \left(\sum_{i \in I_*} c_{i,s_*}\right) \zeta$ where $\zeta = \zeta_*^{s_*} \geq p^{s_* \tau_*}$, and $|\Theta'| = |\sum_{i \in I_*} c_{i,s_*} \vartheta_{i,s_*}| \leq \sum_{i \in I_*} |c_{i,s_*}| \vartheta_{i,s_*} \lll p^{s_* \tau_*}$.

2. If $i \in \Gamma(s_*) \setminus I_*$ then $s_* \tau_i < s_* \tau_*$, and since $s_* \tau_i \underset{u}{\sim} s_* \tau_*$, it follows that $s_* \tau_i \lll s_* \tau_*$. Then

$$|\Psi| \leq \sum_{i \in \Gamma(S_*) \setminus I_*} |c_{i,s_*}| \xi_i^{s_*} < \sum_{i \in \Gamma(S_*) \setminus I_*} |c_{i,s_*}| p^{s_*(\tau_i+1)} \lll p^{s_* \tau_*}.$$

3. Let us show that for every $s \neq s_*$ and for every $i \in \Gamma(s)$, one has $s\tau_i \lll s_* \tau_*$. By the definition of $s_* \tau_*$, clearly $s\tau_i \leq s_* \tau_*$. If by contradiction $s_* \tau_* - s\tau_i = h \in \mathbb{Z}$, then we would have ${}^* f(\tau_i) = \tau_*$ where $f : \mathbb{N} \to \mathbb{N}$ is the function $f(m) = \lfloor (sm + h)/s_* \rfloor$.[20] Since $\tau_* \underset{u}{\sim} \tau_i$, it would follow that ${}^* f(\tau_i) = \tau_i$, and hence $(s_* - s)\tau_i = h$. But τ_i is infinite while $h \in \mathbb{Z}$, and so we must conclude that $s_* = s$, against our hypothesis. The thesis is directly obtained by the following inequalities:

$$|\Phi| \leq \sum_{s \neq s_*} \sum_{i \in \Gamma(s)} |c_{i,s}| \xi_i^s < \sum_{s \neq s_*} \sum_{i \in \Gamma(s)} |c_{i,s}| p^{s(\tau_i+1)} \lll p^{s_* \tau_*}. \quad \square$$

---

[20] By $\lfloor x \rfloor$ we denote the *integer part* of x.

Example 3.12. The polynomial

$$P(x, y) = x^3 + 2x + y^3 - 2y$$

is not PR (even if it contains a partial sum of coefficients that equals zero).

A particular case of Theorem 3.10 is the following.

Corollary 3.13. *Let us consider a Diophantine equation of the form*

$$\sum_{i=1}^{n} c_i x_i^k = P(y)$$

where $P(y)$ is a polynomial of degree $d \neq k$ with no constant term. If for every nonempty set $J \subseteq \{1, \ldots, n\}$ one has $\sum_{i \in J} c_i \neq 0$, then the above equation is not PR.

Finally, by combining Theorems 2.9 and 3.10 one obtains an extension of Rado's Theorem to a large family of nonlinear polynomials.

Corollary 3.14. *Let $n \geq 3$. A polynomial of the form*

$$Q(x_1, \ldots, x_n, y) := c_1 x_1 + \cdots + c_n x_n + P(y)$$

where P is nonlinear is non-trivially PR if and only if there exists a nonempty set $J \subseteq \{1, \ldots, n\}$ such that $\sum_{j \in J} c_j = 0$.

Proof. The sufficient condition is Theorem 2.9; and the necessary condition follows from Corollary 3.13 where $k = 1$ and $d \neq k$ because P is nonlinear. □

Let us itemize some explicit examples of polynomials whose non-PR is proved by our results.

Example 3.15. The equation $x - 2y = P(z)$ is *not* partition regular for any nonlinear polynomial $P(z) \in \mathbb{Z}[z]$. This gives a negative answer to Question 11 (iii) posed by V. Bergelson in [2].

Example 3.16. The equation $x + y = z^2$ is not PR, except for the constant solution $x = y = z = 2$. (This was first proved by P. Csikvári, K. Gyarmati and A. Sárközy in [9].)

Example 3.17. A. Khalfah and E. Szemerédi [39] proved that if $P(z) \in \mathbb{Z}[z]$ takes even values on some integer, then for every finite coloring the equation $x + y = P(z)$ has a solution where x and y are monochromatic. However, as a consequence of Corollary 3.14, it is never the case that $x + y = P(z)$ is partition regular when P is nonlinear.

Example 3.18. In [18], it is proved that the polynomials $x^n + y^m = z^k$ are not PR for $k \notin \{n, m\}$. This result is a consequence of Theorem 3.10.

4 Final remarks and open questions

Over the last ten years, the interest in problems related to the partition regularity of nonlinear Diophantine equations has been rising constantly (see, *e.g.*, [39, 32, 9, 50, 44, 45, 4, 18, 19, 46, 26]). We hope that this paper will contribute to a general Ramsey theory of nonlinear Diophantine equations. In this direction, we think that at least four distinct directions of research are worth pursuing.

The first one is trying to extend our results so as to fully characterize the class of nonlinear PR Diophantine equations in "Rado's style", *i.e.* by means of decidable simple conditions on coefficients and exponents.[21] As the general problem seems highly complicated, it would surely be helpful to start by isolating other classes of PR and non-PR equations. For example, we think that it would be really interesting to find a solution to the following.

Open Problem 1. *Under the additional assumption that the given equation admits infinitely many solutions in $\mathbb{N}$, can the implication in Theorem 3.10 or, at least, in Corollary 3.13, be reversed?* [22]

Notice that a positive answer to this question Theorem 3.10 would imply the PR of the Pythagorean equation for $x^2 + y^2 = z^2$, which is probably the most investigated open problem in this field. It is our opinion that nonstandard analysis may play an important role in this research, also in the positive direction of PR results. Indeed, techniques based on u-equivalence

---

[21] Here the word "decidable" has the precise sense as defined in computability theory to formalize the idea of an "effective method" .

[22] The hypothesis on the existence of solutions is needed, as otherwise the conjecture would be false, as shown, *e.g.*, by Fermat equations $x^n + y^n = z^n$ with $n \geq 3$.

have already been used by the second named author in [44] to prove the PR of several classes of nonlinear equations.

A second possible direction of research is the study the PR of nonlinear Diophantine equations on sets of numbers different from the natural numbers. In this respect, let us point out a few facts.

1. A *homogeneous* Diophantine equation is PR on $\mathbb{N}$ if and only if it is PR on $\mathbb{Z}$ if and only if it is PR on $\mathbb{Q}$.

2. There are homogeneous Diophantine equations that are PR on the positive reals $\mathbb{R}_{>0}$ but not on $\mathbb{N}$.

3. For non-*homogeneous* equations, the equivalences in (1) do not hold.

The "only if" implication in (1) is trivial. Conversely, let us observe that if $P(x_1, \ldots, x_n) = 0$ is a homogeneous Diophantine equation that is PR on $\mathbb{Q}$, then every ultrafilter $\mathcal{U} \in \overline{K(\beta\mathbb{Q}, \odot)}$ is a witness. (This follows from the analogues of Theorem 2.4 and Corollary 2.5 for $\mathbb{Q}$.) Since the set $\mathbb{N}$ is thick in the group $(\mathbb{Q}, \cdot)$, and hence piecewise syndetic, we can pick an ultrafilter $\mathcal{U} \in \overline{K(\beta\mathbb{Q}, \odot)}$ such that $\mathbb{N} \in \mathcal{U}$.[23] Then $\mathcal{U}_{\mathbb{N}} = \{B \cap \mathbb{N} \mid B \in \mathcal{U}\}$ is an ultrafilter on $\mathbb{N}$ that witnesses the PR on $\mathbb{N}$ of the equation $P(x_1, \ldots, x_n) = 0$.

Easy examples to show (2) are given by all Fermat equations $x^n + y^n = z^n$ with $n \geq 3$, which do not admit solutions in $\mathbb{N}$ but are PR on $\mathbb{R}_{>0}$. Indeed, the Schur equation $x + y = z$ is PR on $\mathbb{N}$, and hence on $\mathbb{R}_{>0}$. By taking the function $\varphi(x) = x^n$, which is onto $\mathbb{R}_{>0}$, one can apply the analogue of Proposition 3.4 to the semigroup $\mathbb{R}_{>0}$.

As for (3), consider, *e.g.*, the equation $x_1 y_1 - x_2 = 0$. By the multiplicative Rado's Theorem, that equation is PR on $\mathbb{N}$, and hence it is PR on $\mathbb{Z}$. Then, by Proposition 3.4 for the group $\mathbb{Z}$ applied to the function $f(x) = -x$, we obtain that also $x_1 y_1 + x_2 = 0$ is PR on $\mathbb{Z}$, whilst it is has no solutions in $\mathbb{N}$.

A general question that arises naturally is the following.

Open Problem 2. *Are there simple decidable conditions under which a given (non-homogeneous) Diophantine equation with no constant term is PR on $\mathbb{N}$ if and only if it is PR on $\mathbb{Z}$ if and only if it is PR on $\mathbb{Q}$?*

A problem that seems to have its own peculiarities is the study of PR of Diophantine equations on finite fields. About this, a relevant result has

---

[23] Recall that a subset $A \subseteq S$ of a semigroup $(S, \cdot)$ is piecewise syndetic if and only if it belongs to some ultrafilter in $\overline{K(\beta S, \cdot)}$ (see Corollary 4.41 in [37].)

been recently obtained by P. Csikvári, K. Gyarmati and A. Sárkőzy [9], who proved the PR of every Fermat equation $x^n + y^n = z^n$ on sufficiently large finite fields $\mathbb{F}_p$ (with $xyz \neq 0$).

It seems natural to ask whether the techniques used in this paper may help towards the following.

Open Problem 3. *Are there simple "Rado-like" necessary and sufficient conditions under which a given Diophantine equation with no constant term is PR on sufficiently large finite fields $\mathbb{F}_p$?*

Finally, another really wide direction of research is investigating the PR of finite and infinite systems of nonlinear Diophantine equations. Whilst certain particular results are already known, such as the multiplicative version of Hindman's Theorem, general results in this area are still missing. It is worth remarking that although extensively studied in the recent literature (see, *e.g.*, [30, 12, 33, 34, 31, 29, 40, 10, 25, 50, 36, 1, 4]), infinite linear systems are not fully understood yet. In order to adapt our nonstandard techniques to infinite systems, one would need a characterization of PR systems in terms of u-equivalence. The characterization given in Corollary 3.2 is easily generalized to finite systems, but we do not see how to extend it to infinite systems.

Open Problem 4 *Is there a characterization of PR infinite systems of Diophantine equations in terms of u-equivalence? (Or, equivalently, by means of ultrafilters?)*

Acknowledgement. We thank the referees for their careful reading of the paper and for their helpful comments.

References

[1] B. Barber, N. Hindman, I. Leader, and D. Strauss, *Partition regularity without the columns property*, Proc. Amer. Math. Soc. 143 (2015), 3387–3399.

[2] V. Bergelson, *Ergodic Ramsey Theory - un update*, in "Ergodic Theory of $\mathbb{Z}^d$-actions" (M. Pollicott and K. Scmidt, eds.), London Math. Soc. Lecture Notes Series 228 (1996), 1–61.

[3] V. Bergelson, *Ultrafilters, IP sets, dynamics, and combinatorial number theory*, in "Ultrafilters Across Mathematics", (V. Bergelson, A. Blass, M. Di Nasso and R. Jin, eds.), Contemp. Math. 530, AMS (2010), 23–47.

[4] V. Bergelson, J.H. Johnson jr., and J. Moreira, *New polynomial and multidimensional extensions of classical partition results*, arXiv:1501.02408 (2015).

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

[6] V. Bergelson and J. Moreira, *Ergodic theorem involving additive and multiplicative groups of a field and $\{x + y, xy\}$ patterns*, Ergodic Theory Dynam. Systems, to appear.

[7] V. Bergelson and J. Moreira, Measure preserving actions of affine semigroups and $\{x + y, xy\}$ patterns, arXiv: 1509.07574 (2015)

[8] T.C. Brown and V. Rődl, *Monochromatic solutions to equations with unit fractions*, Bull. Aust. Math. Soc. 43 (1991), 387–392.

[9] P. Csikvári, K. Gyarmati and A. Sárkőzy, *Density and Ramsey type results on algebraic equations with restricted solution sets*, Combinatorica 32, Issue 4 (2012), 425–449.

[10] D. Davenport, N. Hindman, I. Leader, and D. Strauss, *Multiply partition regular matrices*, Discrete Math. 322 (2014), 61–68.

[11] W. Deuber, *Partitionen und lineare Gleichungssysteme*, Math. Z. 133 (1973), 109–123.

[12] W. Deuber, N. Hindman, I. Leader and H. Lefmann, *Infinite partition regular matrices*, Combinatorica 15, Issue 3 (1995), 333–355.

[13] M. Di Nasso, *Iterated hyper-extensions and an idempotent ultrafilter proof of Rado's Theorem*, Proc. Amer. Math. Soc. 143 (2015), 1749–1761.

[14] M. Di Nasso, *Hypernatural numbers as ultrafilters*, Chapter 11 of [42] (2015), 443–474.

[15] M. Di Nasso, *A taste of nonstandard methods in combinatorics of numbers*, in "Geometry, Structure and Randomness in Combinatorics" (J. Matousek, J. Nešetřil, M. Pellegrini, eds.), CRM Series, Scuola Normale Superiore, Pisa, 2015.

[16] M. Di Nasso, I. Goldbring, R. Jin, S. Leth, M.Lupini, and K. Mahlburg, *High density piecewise syndeticity of sumsets*, Adv. Math. 278 (2015), pp. 1-33.

[17] M. Di Nasso, I. Goldbring, R. Jin, S. Leth, M.Lupini, and K. Mahlburg, *Approximate polynomial structure in additively large sets*, Integers 16 (2016), Article A49.

[18] M. Di Nasso and M. Riggio, *Fermat-like equations that are not partitions regular*, arXiv:1605.07527 (2016).

[19] N. Frantzikinakis and B. Host, *Higher order Fourier analysis of multiplicative functions and applications*, J. Amer. Math. Soc., to appear. (Published electronically: March 1, 2016.)

[20] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton University Press, 1981.

[21] R. Goldblatt, *Lectures on the Hyperreals – An Introduction to Nonstandard Analysis*, Graduate Texts in Mathematics **188**, Springer, New York, 1998.

[22] R. Graham, B. Rothschild and J. Spencer, *Ramsey Theory* (2nd edition), Wiley, New York, 1990.

[23] R. Graham, *Some of my favorite problems in Ramsey theory*, Comb. Number Theory 7(2) (2007), 229–236.

[24] B. Green and T. Sanders, *Monochromatic sums and products*, Discrete Analysis (2016).

[25] D. Gunderson, N. Hindman, and H. Lefmann, *Some partition theorems for infinite and finite matrices*, Integers 14 (2014), Article A12.

[26] M.J.H. Heule, O. Kullmann and V.W. Marek, *Solving and verifying the boolean Pythagorean Triples problem via Cube-and-Conquer*, arXiv: 1605.00723 (2016).

[27] N. Hindman, *Finite sums from sequences within cells of a partition of* $\mathbb{N}$, J. Combin. Theory Ser. A, 17 (1974), 1–11.

[28] N. Hindman, *Partitions and sums and products of integers*, Trans. Amer. Math. Soc. 247 (1979), 227–245.

[29] N. Hindman, *Partition regularity of matrices*, Integers 7 (2007), Article A18.

[30] N. Hindman and I. Leader, *Image partition regularity of matrices*, Comb. Prob. and Comp. 2 (1993), 437–463.

[31] N. Hindman and I. Leader, *Nonconstant Monochromatic Solutions to Systems of Linear Equations*, in *Topics in Discrete Mathematics*, Springer, Berlin (2006), 145–154.

[32] N. Hindman, *Monochromatic Sums Equal to Products in* $\mathbb{N}$, Integers 11A (2011), Article 10, 1–10.

[33] N. Hindman, I. Leader, and D. Strauss, *Image partition regular matrices – bounded solutions and preservation of largeness*, Discrete Mathematics 242 (2002), 115–144.

[34] N. Hindman, I. Leader, and D. Strauss, *Infinite partition regular matrices – solutions in central sets*, Trans. Amer. Math. Soc. 355 (2003), 1213–1235.

[35] N. Hindman, I. Leader, and D. Strauss, *Open problems in partition regularity*, Combin. Probab. Comput. 12 (2003), 571–583.

[36] N. Hindman, I. Leader and D. Strauss, *Extensions of infinite partition regular systems*, Electron. J. Combin., 22 (2015), Paper # P2.29.

[37] N. Hindman and D. Strauss, *Algebra in the Stone-Čech Compactification: Theory and Applications*, 2nd edition, de Gruyter, Berlin, 2011.

[38] R. Jin, *The sumset phenomenon*, Proc. Amer. Math. Soc. 130 (2002), 855–861.

[39] A. Khalfalah and E. Szemerédi, *On the number of monochromatic solutions of* $x + y = z^2$, Comb. Prob. and Comp. 15 (2006), 213–227

[40] I. Leader and P.A. Russell, *Independence for partition regular equations*, J. Combin. Theory Ser. A, 114 (2007), 825–839.

[41] H. Lefmann, *On partition regular systems of equations*, J. Combin. Theory Ser. A, 58 (1991), 35–53.

[42] P.A. Loeb and M. Wolff (eds.), *Nonstandard Analysis for the Working Mathematician*, 2nd edition, Springer, 2015.

[43] L. Luperi Baglini, *Hyperintegers and Nonstandard Techniques in Combinatorics of Numbers*, Ph.D. Dissertation (2012), Università di Siena, arXiv: 1212.2049.

[44] L. Luperi Baglini, *Partition regularity of nonlinear polynomials: a nonstandard approach*, Integers 14 (2014), Article 30.

[45] L. Luperi Baglini, *A nonstandard technique in combinatorial number theory*, European J. Combin. 48 (2015), 71–80.

[46] J. Moreira, *Monochromatic sums and products in* $\mathbb{N}$, arXiv:1605.01469 (2016).

[47] R. Rado, *Studien zur Kombinatorik*, Math. Z. 36 (1933), 242–280.

[48] A. Sárkőzy, *On difference sets of sequences of integers I*, Acta Math. Acad. Sci. Hungar. 31 (1978), 125–149.

[49] I. Schur, *Uber die Kongruenz* $x^m + y^m = z^m \,(\bmod p)$, Jahresber. Dtsch. Math.-Ver. 25 (1916), 114–117.

[50] O. Serra and L. Vena, *On the number of monochromatic solutions of integer linear systems on Abelian groups*, European J. Combin. 35 (2014), 459–473.

[51] B.L. van der Waerden, *Beweis einer baudetschen vermutung*, Nieuw. Arch. Wisk. 15 (1927), 212–216.