Nonstandard Methods for Additive and Combinatorial Number Theory—A Survey

Renling Jin
Department of Mathematics
College of Charleston
Charleston, SC 29424
jinr@cofc.edu

0 The beginning

In this article my research on the subject described in the title is summarized. I am not the only person who has worked on this subject. For example, several interesting articles by Steve Leth [21, 22, 23] were published around 1988. I would like to apologize to the reader that no efforts have been made by the author to include other people's research.

My research on nonstandard analysis started when I was a graduate student in the University of Wisconsin. A large part of my thesis was devoted towards solving the problems posed in [19]. By the time when my thesis was finished, many of the problems had been solved. However, some of them were still open including [19, Problem 9.13]. It took me another three years to find a solution to [19, Problem 9.13]. Before this my research on nonstandard analysis was mainly focused on foundational issues concerning the structures of nonstandard universes. After I told Steve about my solution to [19, Probelm 9.13], he immediately informed me how it could be applied to obtain interesting results in combinatorial number theory. This opened a stargate in front of me and lead me into a new and interesting field.

For nonstandard analysis we use a superstructure approach. We fix an $\aleph_1$-saturated nonstandard universe ${}^*\mathbb{V}$. For each standard set A we write ${}^*A$ for the nonstandard version of A in ${}^*\mathbb{V}$.

Mathematics Subject Classification 2000 Primary 03H05, 03H15, 11B05, 11P70; Secondary 11U10, 11B13, 28D05, 28E05

Keywords: upper Banach density, upper asymptotic density, inverse problem, Freiman's conjecture, nonstandard analysis

1 Duality between null ideal and meager ideal

Given an ordered measure space Ω such as the Lebesgue measure space on the real line with the natural order, one can discuss the relationship between measurable sets and open sets[1]. The null ideal on Ω is the collection of null sets, i.e. the sets with measure zero, and the meager ideal is the collection of all meager sets[2]. The sets in an ideal are often considered to be small. The duality between null ideal and meager ideal means that there exists a meager set with positive measure, i.e. the smallness in terms of null ideal is incomparable with the smallness in terms of meager ideal. However, it is usually true that if the space also has an additive structure, then the sum of two set with positive measure may not be meager. See Corollary 1.2 for example. What can we say about a Loeb space?

Let H be a hyperfinite integer and let $[0, H]$ be an interval of integers. The term $[a, b]$ in this article always means the interval of integers between a and b including a and b if they are also integers. On $[0, H]$ one can construct a Loeb measure generated by the normalized counting measure. By a Loeb space we always mean the hyperfinite set $[0, H]$ with the Loeb measure generated by the normalized counting measure[3]. On $[0, H]$ there is also a natural order and an additive structure. However, the order topology on $[0, H]$ is discrete, therefore uninteresting. In [19] a U-topology is introduced for each cut $U \subseteq [0, H]$ that gives meaningful analogy of the order topology on the real line. An infinite initial segment U of non-negative integers is called a cut if it is closed under addition. For example $\mathbb{N}$, the set of all standard non-negative integers, is the smallest cut. We often write $x > U$ for a positive integer x and a cut U if $x \notin U$. Let $U \subseteq [0, H]$ be a cut. A set $A \subseteq [0, H]$ is called U-open if for every $x \in A$, there exists a positive integer $y > U$ such that $[x - y, x + y] \cap [0, H] \subseteq A$. A U-topology is the collection of all U-open sets and a U-meager set is a meager set in terms of U-topology. In [19] it was proven that for any cut $U \subseteq [0, H]$ there is always a U-meager set of Loeb measure one in $[0, H]$. The question 9.13 in [19] asked whether the sum (modulo $H + 1$) of two sets in $[0, H]$ with positive Loeb measure can

[1] An order on a space can generate a topology called order topology on the space so that a set is open if it is the union of open intervals.

[2] A set A in a topological space is nowhere dense if every non-empty open set O contains another non-empty open set R disjoint from A. A meager set is the union of finitely many or countably many nowhere dense sets. A meager set is also called a set of the first category.

[3] The Loeb space here is often called the hyperfinite uniform Loeb space.

be U-meager for some cut $U \subseteq [0, H]$. For two sets A and B and a binary operation $\circ$ between A and B, we write $A \circ B$ for the set $\{a \circ b : a \in A \text{ and } b \in B\}$. For a number k, we write kA for the set $\{ka : a \in A\}$. In [11] we prove the following theorem.

Theorem 1.1 *Let H be a hyperfinite integer and $U \subseteq [0, H]$ be any cut. If $A, B \subseteq [0, H]$ are two internal sets with positive Loeb measure, then $A \oplus_H B$ is not U-nowhere dense, where $\oplus_H$ is the usual addition modulo $H + 1$.*

Note that Theorem 1.1 yields a negative answer to [19, Problem 9.13]. Also note that the theorem is still true if $\oplus_H$ is replaced by the usual addition $+$ and the sumset $A + B$ is considered to live in $[0, 2H]$. Theorem 1.1 has several corollaries in the standard world. If one lets U be the cut $\bigcap_{n \in \mathbb{N}} [0, \frac{H}{n}]$, then Theorem 1.1 implies the following well known non-trivial fact.

Corollary 1.2 *If A and B are two sets of reals with positive Lebesgue measure, then $A + B$ must contain a non-empty open interval of reals.*

Corollary 1.2 was credited to Steinhaus in [21].

Let $A \subseteq \mathbb{N}$ be infinite. The upper Banach density $BD(A)$ of A is defined by

$$BD(A) = \lim_{k \to \infty} \sup_{n \in \mathbb{N}} \frac{|A \cap [n, n+k]|}{k+1}.$$

A set $C \subseteq \mathbb{N}$ is called piecewise syndetic if there is a positive integer k such that $C + [0, k]$ contains arbitrarily long sequence of consecutive numbers. The definition of upper Banach density, syndeticity, and piecewise syndeticity can be found in [1, 7]. If one let $U = \mathbb{N}$, then Theorem 1.1 implies the following result.

Corollary 1.3 *Let $A, B \subseteq \mathbb{N}$. If $BD(A) > 0$ and $BD(B) > 0$, then $A+B$ is piecewise syndetic.*

Corollary 1.3 was pointed out to me by Steve Leth. By choosing other cuts U one can have more corollaries. These corollaries also have their own corollaries. The reader can find more of them in [11, 18].

Corollary 1.2 and Corollary 1.3 can also be proven using standard methods [13]. However, Theorem 1.1 does not have a standard version. The generality of Theorem 1.1 shows the advantages of the nonstandard methods. Theorem 1.1 reveals a universal phenomenon, which says that if two sets are large in terms of “measure”, then $A + B$ must not be small in terms of “order-topology”.

2 Buy-one-get-one-free scheme

Excited by the results such as Corollary 1.3, I was eager to let people know what I had obtained. After a talk I gave at a meeting in 1997, I was informed by a member in the audience that Corollary 1.3 had probably already been proven in [1] or in [7]. This made me rush to the library to check out the book and the paper; I was anxious to see whether my efforts were a waste of time. Fortunately, they weren't; In fact, Corollary 1.3 complemented a theorem in [7] which says that if a set $A \subseteq \mathbb{N}$ has positive upper Banach density, then $A - A$ is syndetic. From [1, 7] I also learned of terms such as *upper Banach density, syndeticity, piecewise syndeticity*, etc. the first time.

One thing which caught my eye when I read [1, 7] was the use of Birkhoff Ergodic Theorem. It is natural for a nonstandard analyst to think what one can achieve if Birkhoff Ergodic Theorem is applied to some problems in a Loeb measure space setting. With that in mind, I derived Theorem 2.1 and Theorem 2.2 as lemmas in [12].

Given a set $A \subseteq \mathbb{N}$, the lower asymptotic density $\underline{d}(A)$ of A is defined by

$$\underline{d}(A) = \liminf_{n\to\infty} \frac{|A \cap [1,n]|}{n},$$

where $|X|$ means the cardinality of X when X is finite. Later, we will use $|X|$ representing the internal cardinality of X when X is a hyperfinite set. For a set A and a number x we often write $A \pm x$ for $A \pm \{x\}$ and write $x \pm A$ for $\{x\} \pm A$.

Theorem 2.1 *Suppose $A \subseteq \mathbb{N}$ with $BD(A) = \alpha$. Then there is an interval of hyperfinite length $[H, K]$ such that for almost all $x \in [H, K]$ in terms of the Loeb measure on $[H, K]$, we have $\underline{d}((^*A - x) \cap \mathbb{N}) = \alpha$. On the other hand, if $A \subseteq \mathbb{N}$ and there is a positive integer x such that $\underline{d}((^*A - x) \cap \mathbb{N}) \geqslant \alpha$, then $BD(A) \geqslant \alpha$.*

Given a set $A \subseteq \mathbb{N}$, the Shnirel'man density $\sigma(A)$ of A is defined by

$$\sigma(A) = \inf_{n\geqslant 1} \frac{|A \cap [1,n]|}{n}.$$

Theorem 2.2 *Suppose $A \subseteq \mathbb{N}$ with $BD(A) = \alpha$. Then there is a positive integer x such that $\sigma((^*A - x) \cap \mathbb{N}) = \alpha$.*

Theorem 2.1 is [12, Lemma 2] and Theorem 2.2 is the combination of [12, Lemma 3, Lemma 4, and Lemma 5]. It is often the case that a result involving Shnirel'man density is obtained first. Then people explore possible generalizations to some results involving lower asymptotic density. The behaviors of these two densities are quite similar. From Theorem 2.1 and Theorem 2.2 we can see that the behavior of upper Banach density is also similar to the behavior of lower asymptotic density or Shnirel'man density. We can now claim that there is a theorem involving upper Banach density parallel to each existing theorem involving lower asymptotic density or Shnirel'man density. This is the scheme that can be called *buy-one-get-one-free* because we can get a parallel theorem involving upper Banach density for free as soon as a theorem involving lower asymptotic density or Shnirel'man density is obtained. I would now like to briefly describe how this works.

Given a set $A \subseteq \mathbb{N}$ with $BD(A) = \alpha$, there is a positive integer x (may be nonstandard) such that $\underline{d}((^*A - x) \cap \mathbb{N}) = \alpha$ (or $\sigma((^*A - x) \cap \mathbb{N}) = \alpha$). This means that in $x + \mathbb{N}$, a copy of $\mathbb{N}$ above x, the set $^*A$ has lower asymptotic density (or Shnirel'man density) α. Now apply the existing theorem involving $\underline{d}$ (or σ) to the set $^*A \cap (x + \mathbb{N})$ to obtain a result about $^*A$. Finally, pushing down the result to the standard world, one can obtain a parallel theorem involving upper Banach density. Corollary 2.3 and Corollary 2.4 below are the results obtained using this scheme.

The first one is a corollary parallel to Mann's Theorem. Mann's Theorem says that if two sets $A, B \subseteq \mathbb{N}$ both contain 0, then $\sigma(A + B) \geqslant \min\{\sigma(A) + \sigma(B), 1\}$. Mann's Theorem is an important theorem; In [20] it is referred as one of three pearls in number theory. We can now easily prove the following corollary of Theorem 2.2.

Corollary 2.3 $BD(A + B + \{0, 1\}) \geqslant \min\{BD(A) + BD(B), 1\}$ *for all* $A, B \subseteq \mathbb{N}$.

The addition of $\{0, 1\}$, which substitutes the condition $0 \in A \cap B$ in Mann's Theorem, is necessary because without it, Corollary 2.3 is no longer true. For example if A and B both are the set of all even numbers, then $BD(A) = BD(B) = BD(A+B) = \frac{1}{2}$.

The second corollary is parallel to Plünnecke's Theorem [24, p.225] which says that if $B \subseteq \mathbb{N}$ is a basis of order h, then $\sigma(A + B) \geqslant \sigma(A)^{1-\frac{1}{h}}$ for every set $A \subseteq \mathbb{N}$. A set $B \subseteq \mathbb{N}$ is called a basis of order h if every non-negative integer n is the sum of at most h non-negative integers (repetition is allowed) from B. In the upper Banach density setting we can define piecewise basis of order h. A set $B \subseteq \mathbb{N}$ is called a piecewise basis of order h if there is a sequence of intervals $[a_k, b_k] \subseteq \mathbb{N}$ with $\lim_{k\to\infty}(b_k - a_k) = \infty$ such that every integer $n \in [0, b_k - a_k]$ is the sum of at most h integers (repetition allowed) from $(B - a_k) \cap [0, b_k - a_k]$. If B is a basis of order h, then it must also be a piecewise basis of order h because one can take $b_k = k$ and $a_k = 0$.

Corollary 2.4 *If B is a piecewise basis of order h, then $BD(A + B) \geqslant BD(A)^{1-\frac{1}{h}}$ for every set $A \subseteq \mathbb{N}$.*

Corollary 2.3 and Corollary 2.4 can also be proven using the standard methods [13] from Ergodic Theory. For more results similar to Corollary 2.3 and Corollary 2.4, see [12].

3 From Kneser to Banach

In the last section we didn't use the full power of Theorem 2.1. Suppose $A \subseteq \mathbb{N}$ has upper Banach density α. We use only one x such that $\underline{d}((^*A - x) \cap \mathbb{N}) = \alpha$ while there are almost all x in an interval $[H, K]$ of hyperfinite length such that $\underline{d}((^*A - x) \cap \mathbb{N}) = \alpha$. We can take this advantage and prove a theorem involving upper Banach density parallel to Kneser's Theorem [9]. For two sets $A, B \subseteq \mathbb{N}$ we write $A \sim B$ if $(A \smallsetminus B) \cup (B \smallsetminus A)$ is a finite set. Kneser's Theorem[4] says that for all $A, B \subseteq \mathbb{N}$, if $\underline{d}(A + B) < \underline{d}(A) + \underline{d}(B)$, then there is a positive integer g and a set $G \subseteq [0, g - 1]$ such that $A + B \subseteq G + g\mathbb{N}$, $A + B \sim G + g\mathbb{N}$, and $\underline{d}(A + B) = \frac{|G|}{g} \geqslant \underline{d}(A) + \underline{d}(B) - \frac{1}{g}$.

Kneser's Theorem was motivated by Mann's Theorem. Can Mann's Theorem be true if one replaces Shnirel'man density with lower asymptotic density? There are obvious counterexamples. Let d, k, and k' be positive integers. Suppose $G = \{0, d, 2d, \ldots, (k - 1)d\}$ and $G' = \{0, d, 2d, \ldots, (k' - 1)d\}$. Suppose also that $g > (k+k'-2)d$, $A = G+g\mathbb{N}$, and $B = G'+g\mathbb{N}$. Then $\underline{d}(A+B) = \frac{k+k'-1}{g} = \underline{d}(A)+\underline{d}(B)-\frac{1}{g}$. Roughly speaking, Kneser's Theorem says that the only kind of counterexamples which make the inequality false in Mann's Theorem with σ replaced by $\underline{d}$ are similar to the one just described.

[4]The version of Kneser's Theorem in [9] is about the addition of multiple sets. We stated the version here for the addition of two sets just for simplicity.

In [13] a parallel theorem [13, Theorem 3.8] was obtained. Let $A, B \subseteq \mathbb{N}$ with $BD(A) = \alpha$ and $BD(B) = \beta$. Then there are intervals $[a_n, b_n]$ and $[c_n, d_n]$ such that $\lim_{n\to\infty}(b_n - a_n) = \infty$, $\lim_{n\to\infty}(d_n - c_n) = \infty$, $\lim_{n\to\infty} \frac{|A\cap[a_n,b_n]|}{b_n-a_n+1} = \alpha$, and $\lim_{n\to\infty} \frac{|B\cap[c_n,d_n]|}{d_n-c_n+1} = \beta$. We hope to characterize the structure of $A + B$ inside the intervals $[a_n+c_n, b_n+d_n]$. However, [13, Theorem 3.8] when restricted to the addition of two sets, only characterized the structure of $A+B$ on a very small part of $\mathbb{N}$. The reason for this is because we used only one x and one y with $\underline{d}((^*A - x) \cap \mathbb{N}) = \alpha$ and $\underline{d}((^*B - y) \cap \mathbb{N}) = \beta$ in the proof.

During the summer of 2003 my undergraduate research partner Prerna Bihani and I conducted an undergraduate research project funded by the College of Charleston to work on theorems parallel to Kneser's Theorem. The work done during the summer and the following year produced the paper [2], which contains Theorem 3.1. To avoid some technical difficulties we considered only the sum of two copies of the same set in [2].

Theorem 3.1 *Let A be a set of non-negative integers such that $BD(A) = \alpha$ and $BD(A + A) < 2\alpha$. Let $\{[a_n, b_n] : n \in \mathbb{N}\}$ be a sequence of intervals such that $\lim_{n\to\infty}(b_n - a_n) = \infty$ and $\lim_{n\to\infty} \frac{|A\cap[a_n,b_n]|}{b_n-a_n+1} = \alpha$. Then there are $g \in \mathbb{N}$, $G \subseteq [0, g-1]$, and $[c_n, d_n] \subseteq [a_n, b_n]$ for each $n \in \mathbb{N}$ such that*

(1). $\lim_{n\to\infty} \frac{d_n-c_n}{b_n-a_n} = 1$,

(2). $A + A \subseteq G + g\mathbb{N}$,

(3). $(A + A) \cap [2c_n, 2d_n] = (G + g\mathbb{N}) \cap [2c_n, 2d_n]$ *for all* $n \in \mathbb{N}$,

(4). $BD(A + A) = \frac{|G|}{g} \geqslant 2\alpha - \frac{1}{g}$.

Note that (1) above shows that the structure of $A+A$ is characterized on a large portion of $[2a_n, 2b_n]$. Note also that we cannot replace $[2c_n, 2d_n]$ with $[2a_n, 2b_n]$ in (3) because all conditions for A still hold if we delete any elements from $A \cap ([a_n, b_n] \setminus [c_n, d_n])$.

The proof of Theorem 3.1 can be described in several steps. Given a hyperfinite integer N, we know that for almost all $x, y \in [a_N, b_N]$ we have $\underline{d}((^*A - x) \cap \mathbb{N}) = \underline{d}((^*A - y) \cap \mathbb{N}) = \alpha$. We can also assume that $\underline{d}((x - {^*A}) \cap \mathbb{N}) = \underline{d}((y - {^*A}) \cap \mathbb{N}) = \alpha$. Step one: characterize the structure of $^*(A+A)$ in $x+y+\mathbb{Z}$ using Kneser's Theorem, where $\mathbb{Z}$ is the set of all standard integers. Step two: show that the structures of $^*(A+A)\cap(x+y+\mathbb{Z})$ for almost all $x,y\in[a_N,b_N]$ are consistent with one another so that these structures can be combined into one structure. Hence we can characterize the structure of $^*(A+A)$ in $[2c_N,2d_N]$, where $\frac{d_N-c_N}{b_N-a_N}\approx 1$. Step three: prove that for different hyperfinite integers N and N', the structure of $^*(A+A)$ in $[2a_N,2b_N]$ and the structure of $^*(A+A)$ in $[2a_{N'},2b_{N'}]$ are consistent so that these structures of $^*(A+A)$ in $[2a_N,2b_N]$ for all hyperfinite integers N can be combined into one structure of $^*(A+A)$ in $\bigcup\{[a_N,b_N]: N \text{ is hyperfinite}\}$. Step five: pushing down the structure of $^*(A+A)$ to the standard world results Theorem 3.1.

The methods developed in [13] do not seem to be enough for proving Theorem 3.1. So it is interesting to see whether one can produce a reasonably nice and short standard proof of the theorem.

In [3] the structure of A was characterized when $\underline{d}(A)$ is very small and $\underline{d}(A+A)\leqslant c\underline{d}(A)$ for some constant $c\geqslant 2$. It is also interesting to see how one can characterize the structure of $A+A$ when $BD(A+A)\leqslant cBD(A)$ for some constant $c\geqslant 2$.

4 Inverse problem for upper asymptotic density

In January of 2000, I was invited to give a talk at the DIMACS workshop "Unusual Applications of Number Theory". One of the workshop organizers was Melvyn Nathanson to whom I am grateful for being the first number theorist to express an interest in my research on number theory not to mention his continued encouragement. During the workshop I had a chance to meet another number theorist G. A. Freiman who is well-known for his work on inverse problems in additive and combinatorial number theory. He gave me a preprint of his list of open problems [5]. This list and the book [24] have since gotten me interested in the inverse problems.

Inverse problems study the properties of A when $A+A$ satisfies certain conditions. Freiman discovered a phenomenon that if $A+A$ is small, then A must have some arithmetic structure. In fact Kneser's Theorem and Theorem 3.1 can be viewed as two examples of the phenomenon. One can characterize the arithmetic structure of A from the structure of $A+A$ in Theorem 3.1 and characterize the structure of A and the structure of B from the structure of $A+B$ in Kneser's Theorem (see [2] for details). In this section we characterize the structure of A when the upper asymptotic density of $A+A$ is small. Given $A\subseteq\mathbb{N}$, the upper asymptotic density $\bar{d}(A)$ of A is defined by

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

Without loss of generality we always assume $0 \in A$ in this section. We can also assume that $\gcd(A) = 1$ because if $\gcd(A) = d > 1$, then we can recover the structure of A from the structure of A', where $A' = \{a/d : a \in A\}$. When $0 \in A$ and $\gcd(A) = 1$, one can easily prove, using Freiman's result (1) at the beginning of the next section, that $\bar{d}(A+A) \geqslant \frac{3}{2}\bar{d}(A)$ if $\bar{d}(A) \leqslant \frac{1}{2}$ and $\bar{d}(A+A) \geqslant \frac{1+\bar{d}(A)}{2}$ if $\bar{d}(A) \geqslant \frac{1}{2}$. The following two examples show that the lower bounds above are optimal.

Example 4.1 *For every real number* $0 \leqslant \alpha \leqslant 1$*, let*

$$A = \{0\} \cup \bigcup_{n=1}^{\infty} [\lceil (1-\alpha)2^{2^n} \rceil, 2^{2^n}].$$

Then $\bar{d}(A) = \alpha$*,* $\bar{d}(A+A) = \frac{1+\alpha}{2}$ *if* $\alpha \geqslant \frac{1}{2}$*, and* $\bar{d}(A+A) = \frac{3}{2}\alpha$ *if* $\alpha \leqslant \frac{1}{2}$*.*

Example 4.2 *Let* $k, m \in \mathbb{N}$ *be such that* $k \geqslant 4$ *and* $0, m, 2m$ *are pairwise distinct modulo* k*. Let* $A = k\mathbb{N} \cup (m + k\mathbb{N})$*. Then* $\bar{d}(A) = \frac{2}{k} = \alpha \leqslant \frac{1}{2}$ *and* $\bar{d}(A+A) = \frac{3}{k} = \frac{3}{2}\alpha$*. It is easy to choose* k, m *such that* $\gcd(A) = 1$*.*

So we can say that $\bar{d}(A+A)$ is small when $\bar{d}(A+A) = \min\{\frac{3}{2}\bar{d}(A), \frac{1+\bar{d}(A)}{2}\}$ and we need to characterize the structure of A when $\bar{d}(A+A)$ is small.

Clearly the characterization of the structure of A should cover the cases in both Example 4.1 and Example 4.2. We hope to show that A must have the structure described in one of the examples above when $\bar{d}(A+A)$ is small. However, some variations of the examples are unavoidable. If the set A is replaced by $A' \subseteq A$ with $\bar{d}(A') = \alpha$ in both Example 4.1 and Example 4.2, then $\bar{d}(A' + A')$ is also small. Furthermore, if $A = A' \cup A''$ where

$$A' = \{0\} \cup \bigcup_{n=1}^{\infty} [\lceil (1-\alpha)2^{2^{2n}} \rceil, 2^{2^{2n}}]$$

and A'' is an arbitrary subset of

$$\bigcup_{n=1}^{\infty} [\lceil (1-\alpha)2^{2^{2n+1}} \rceil, 2^{2^{2n+1}}],$$

then again $\bar{d}(A+A)$ is small. This example shows that we can only hope to characterize the structure of A along the increasing sequence h_n such that $\lim_{n\to\infty} \frac{|A\cap[1,h_n]|}{h_n} = \alpha$.

It was a long journey for me to arrive at the most recent result in [16] due to the technical difficulties of the proof. First the structure of A was characterized in [14] when $\bar{d}(A + A + \{0, 1\})$ is small. Later the structure of A was characterized in [15] when $\bar{d}(A + A)$ is small and A contains two consecutive numbers. Finally in [16] the following theorem was proven.

Theorem 4.3 *Let* $\bar{d}(A) = \alpha > 0$.

Part I: Assume $\alpha > \frac{1}{2}$. *Then* $\bar{d}(A + A) = \frac{1+\alpha}{2}$ *implies that for every increasing sequence* $\{h_n : n \in \mathbb{N}\}$ *with* $\lim_{n\to\infty} \frac{|A\cap[0,h_n]|}{h_n+1} = \alpha$, *we have*

$$\lim_{n\to\infty} \frac{|(A + A) \cap [0, h_n]|}{h_n + 1} = \alpha.$$

Part II: Assume $\alpha < \frac{1}{2}$ *and* $\gcd(A) = 1$. *Then* $\bar{d}(A + A) = \frac{3}{2}\alpha$ *implies that either (a) there exist* $k > 4$ *and* $c \in [1, k - 1]$ *such that* $\alpha = \frac{2}{k}$ *and* $A \subseteq k\mathbb{N} \cup (c + k\mathbb{N})$ *or (b) for every increasing sequence* $\{h_n : n \in \mathbb{N}\}$ *with* $\lim_{n\to\infty} \frac{|A\cap[0,h_n]|}{h_n+1} = \alpha$, *there exist two sequences* $0 \leqslant c_n \leqslant b_n \leqslant h_n$ *such that*

$$\lim_{n\to\infty} \frac{|A \cap [b_n, h_n]|}{h_n - b_n + 1} = 1,$$

$$\lim_{n\to\infty} \frac{c_n}{h_n} = 0,$$

and $[c_n + 1, b_n - 1] \cap A = \emptyset$ *for every* $n \in \mathbb{N}$.

Part III: Assume $\alpha = \frac{1}{2}$ *and* $\gcd(A) = 1$. *Then* $\bar{d}(A + A) = \frac{3}{2}\alpha$ *implies that either (a) there exists* $c \in \{1, 3\}$ *such that* $A \subseteq 4\mathbb{N} \cup (c + 4\mathbb{N})$ *or (b) for every increasing sequence* $\{h_n : n \in \mathbb{N}\}$ *with* $\lim_{n\to\infty} \frac{|A\cap[0,h_n]|}{h_n+1} = \alpha$, *we have*

$$\lim_{n\to\infty} \frac{|(A + A) \cap [0, h_n]|}{h_n + 1} = \alpha.$$

I would like to make some remarks here on Theorem 4.3. First, the proof of Part I is easy; The most difficult part is Part II. Second, Part I and (b) of Part III cannot be improved so that set A has the structure similar to the structure described in (b) of Part II. For example, if one lets

$$A = \{0\} \cup \bigcup_{n=1}^{\infty} \left([3 \cdot 2^{2^n-3}, 4 \cdot 2^{2^n-3}] \cup [5 \cdot 2^{2^n-3}, 2^{2^n}] \right),$$

then $\bar{d}(A) = \frac{1}{2}$ and $\bar{d}(A+A) = \frac{1+\bar{d}(A)}{2}$. Clearly A does not have the structure described in (b) of Part II.

The main ingredient of the proof of Theorem 4.3 is the following lemma in nonstandard analysis. For an internal set $A \subseteq [0, H]$ and a cut $U \subseteq [0, H]$ we define the lower U-density $\underline{d}_U(A)$ by

$$\underline{d}_U(A) = \sup \left\{ \inf \left\{ st \left(\frac{|A \cap [0, n]|}{n+1} \right) : n \in U \setminus [0, m] \right\} : m \in U \right\},$$

where st means the standard part map. Note that if $U = \mathbb{N}$ and $A \subseteq \mathbb{N}$, then $\underline{d}(A) = \underline{d}_U({}^*A)$. A set $I = \{a, a+d, a+2d, \ldots\}$ is called an arithmetic progression with difference d. An arithmetic progression can be finite (hyperfinite) or infinite. If an arithmetic progression is finite (hyperfinite), then its cardinality (internal cardinality) is its length. A set $I \cup J$ is called a bi-arithmetic progression if both I and J are arithmetic progressions with the same difference d and $I + I$, $I + J$, and $J + J$ are pairwise disjoint. A finite (hyperfinite) bi-arithmetic progression $I \cup J$ has its length $|I| + |J|$. Let U be a cut. A bi-arithmetic progression $B \subseteq U$ is called U-unbounded if both I and J are upper unbounded in U.

Lemma 4.4 *Let H be hyperfinite and $U = \bigcap_{n \in \mathbb{N}} [0, \frac{H}{n}]$. Suppose $A \subseteq [0, H]$ be such that $0 < \underline{d}_U(A) = \alpha < \frac{2}{3}$. If $A \cap U$ is neither a subset of an arithmetic progression of difference greater than 1 nor a subset of a U-unbounded bi-arithmetic progression, then there is a standard positive real number $\gamma > 0$ such that for every $N > U$, there is a $K \in A$, $U < K < N$, such that*

$$\frac{|(A + A) \cap [0, 2K]|}{2K + 1} \geqslant 3 \frac{|A \cap [0, K]|}{2K + 1} + \gamma.$$

Lemma 4.4 is motivated by Kneser's Theorem. It basically says that either $A + A$ is large in an interval $[0, 2K]$ with $K > \frac{H}{n}$ for some standard n or A has desired arithmetic structure in an interval $[0, K]$ with $K > \frac{H}{n}$ for some standard n. The proof uses the fact that U is an additive semi-group. This can be done only in a nonstandard setting. It is interesting to see whether this lemma can be replaced by a standard argument with a reasonable length.

Recently G. Bordes [4] generalized Part II of Theorem 4.3 for sets A with small upper asymptotic density. He characterized the structure of A when $\bar{d}(A) \leqslant \alpha_0$ for some small positive number α_0 and $\bar{d}(A+A) < \frac{5}{3}\bar{d}(A)$. It is interesting to see whether one can replace α_0 by a relatively large value, say $\frac{2}{5}$, in Bordes' Theorem.

5 Freiman's $3k-3+b$ conjecture

After Theorem 4.3 was proven, I realized that the same methods used there could also be used to advance the existing results towards the solution of Freiman's $3k-3+b$ conjecture [5]. This is important because the conjecture is about the inverse problem for the addition of *finite sets*. Let A be a finite set of integers with cardinality $k>0$. It is easy to see that $|A+A| \geqslant 2k-1$. On the other hand, if $|A+A|=2k-1$, then A must be an arithmetic progression. In the early 1960s, Freiman obtained the following generalizations [6].

(1) Let $A \subseteq \mathbb{N}$. Suppose $k=|A|$, $0=\min A$, and $n=\max A$. Suppose also $\gcd(A)=1$. Then $|A+A| \geqslant 3k-3$ if $n \geqslant 2k-3$ and $|A+A| \geqslant k+n$ if $n \leqslant 2k-3$.

(2) If $k>3$ and $|2A|=2k-1+b<3k-3$, then A is a subset of an arithmetic progression of length at most $k+b$

(3) If $k>6$ and $|2A|=3k-3$, then either A is a subset of an arithmetic progression of length at most $2k-1$ or A is a bi-arithmetic progression.

In [6] a result was also mentioned without proofs for characterizing the structure of A when $k>10$ and $|A+A|=3k-2$. In [10] an interesting generalization of (3) above was obtained by Hamidoune and Plagne, where the condition $|2A|=3k-3$ is replaced by $|A+tA|=3k-3$ for every integer t. However, no further progress of this kind had been made for a larger value of $|A+A|$ before my recent work. In fact, Freiman made the following conjecture in [5] five years ago.

Conjecture 5.1 *There exists a natural number K such that for any finite set of integers A with $|A|=k>K$ and $|A+A|=3k-3+b<\frac{10}{3}k-5$ for some $b \geqslant 0$, A is either a subset of an arithmetic progression of length at most $2k-1+2b$ or a subset of a bi-arithmetic progression of length at most $k+b$.*

Note that the conclusion of Conjecture 5.1 could be false if one allows $|A+A|=\frac{10}{3}k-5$. Simply let A be the union of three intervals $[0,a-1]$, $[b,b+a-1]$, and $[2b,2b+a-1]$, where $k=3a$ and b is a sufficiently large integer. Clearly $|A+A|=\frac{10}{3}k-5$. Since b can be as large as we want, we can choose a b so that set A is neither a subset of an arithmetic progression of a restricted length nor a subset of a bi-arithmetic progression of a restricted length.

Using nonstandard methods such as Lemma 4.4, I was able to prove the following theorem in [17].

Theorem 5.2 *Suppose $f : \mathbb{N} \mapsto \mathbb{N}$ is a function with $\lim_{n\to\infty} \frac{f(n)}{n} = 0$. There exists a natural number K such that for any finite set of integers A with $|A| = k$, if $k > K$ and $|A + A| = 3k - 3 + b$ for some $0 \leqslant b \leqslant f(k)$, then A is either a subset of an arithmetic progression of length at most $2k - 1 + 2b$ or a subset of a bi-arithmetic progression of length at most $k + b$.*

Theorem 5.2 gives a new result even for $f(x) \equiv 2$. However, we still have a long way before solving Conjecture 5.1. It is already interesting to see whether we can obtain the same result with $f(x) = \alpha x$ for some positive real number α.

The ideas for proving Theorem 5.2 are similar to the proof of Theorem 4.3, but much more technical. Suppose Theorem 5.2 is not true; Then one can find a sequence of counterexamples A_n such that $|A_n| \to \infty$. Given a hyperfinite integer N, let $A = A_N$. Without loss of generality, we can assume that $0 = \min A$, $H = \max A$, $\gcd(A) = 1$, and $\alpha \approx \frac{|A|}{H+1} \gg 0$. Note that $\frac{|A+A|-3|A|+3}{H} \approx 0$. Hence $|A + A|$ is almost the same as $3|A| - 3$ from the nonstandard point of view. Using the case-by-case argument, we can show that if $\frac{|A|}{H+1} \ll \frac{1}{2}$, then A is a subset of a bi-arithmetic progression. If $\frac{|A|}{H+1} \approx \frac{1}{2}$ and $b = |A + A| - 3|A| + 3$, then we can show that $H + 1 \leqslant 2|A| - 1 + 2b$ when A is not a subset of a bi-arithmetic progression. The proof for the case $\frac{|A|}{H+1} \approx \frac{1}{2}$ is much harder than the proof for the case $\frac{|A|}{H+1} \ll \frac{1}{2}$ although the former depends on the latter. In both cases Lemma 4.4 was used to get structural information of A on an interval with length longer than $\frac{H}{n}$ for some standard positive integer n.

There are some similarities between our methods and analytic methods. In order to detect some structural properties of $A \subseteq [0, n]$, one may need to show that either A is uniformly distributed on $[0, n]$ or A has a greater density on a well formed subset of $[0, n]$. The analytic methods usually look for a large Fourier coefficient $\bar{A}(r)$ (cf. [8, Corollary 2.5]) or a large exponential sum $\sum_{i=0}^{n-1} A(i)e^{\frac{2\pi i}{n}}$ (cf.[24, Theorem 2.9]) to detect the greater density on a well formed subset of $[0, n]$ when n is a prime number. When n is not a prime number then one needs to replace it with a prime number $p > n$ and consider A in $[0, p]$ instead. This replacement may not work well for Conjecture 5.1 as the structure of A needs to be very precise. In our methods we look for the greater density of $A \subseteq [0, H]$ on an interval $[0, K]$ for some $K > U$ by checking the value of $\underline{d}_U(A)$, where $U = \bigcap_{n\in\mathbb{N}} [0, \frac{H}{n}]$. If $\underline{d}_U(A) \geqslant \frac{2}{3}$, then the density of A on $[0, K]$ for some $K > U$ is significantly greater than $|A|/H$, which will lead to a contradiction that $|A+A|$ is almost the same as $3k-3$. If $\underline{d}_U(A)=0$, then the density of A on $[K,H]$ is significantly greater than $|A|/H$, which will again lead to a contradiction. Otherwise either $|(A+A)\cap[0,2K]|$ is large, which is impossible by the fact that $\frac{|A+A|}{2H+1}\lessapprox\frac{3}{2}\alpha$, or A has very nice structural properties on $[0,K]$ following Lemma 4.4, which will force A to have the structure we hope for.

References

[1] *V. Bergelson*, Ergodic Ramsey theory–an update, Ergodic theory of $\mathbb{Z}^d$ actions (Warwick, 1993-1994) 1-61, London Mathematical Society Lecture Note Ser. 228, Cambridge University Press, Cambridge 1996.

[2] *P. Bihani and R. Jin*, Kneser's Theorem for upper Banach density, submitted.

[3] *Y. Bilu*, Addition of sets of integers of positive density, the Journal of Number Theory, **64** (1997), No. 2, 233—275.

[4] *G. Bordes*, Sum-sets of small upper density, Acta Arithmetica, to appear.

[5] *G. A. Freiman*, Structure theory of set addition. II. Results and problems. Paul Erdös and his mathematics, I (Budapest, 1999), 243–260, Bolyai Soc. Math. Stud., 11, János Bolyai Math. Soc., Budapest, 2002.

[6] *G. A. Freiman*, **Foundations of a structural theory of set addition**. Translated from the Russian. Translations of Mathematical Monographs, Vol 37. American Mathematical Society, Providence, R. I., 1973

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

[8] *T. Gowers*, A new proof of Szemerédi's theorem, Geometric and Functional Analysis, 11 (2001), 465-588.

[9] *H. Halberstam and K. F. Roth*, **Sequences**, Oxford University Press, 1966.

[10] *Y. O. Hamidoune and A. Plagne*, A generalization of Freiman's $3k-3$ theorem. Acta Arith. 103 (2002), no. 2, 147–156.

[11] *R. Jin*, Sumset phenomenon, The Proceedings of American Mathematical Society, Vol. 130, No.3 (2002), pp.855—861.

[12] *R. Jin*, Nonstandard methods for upper Banach density problems, The Journal of Number Theory, 91 (2001), 20—38.

[13] *R. Jin*, Standardizing nonstandard methods for upper Banach density problems, in the DIMACS series **Unusual Applications of Number Theory**, edited by M. Nathanson, Vol. 64 (2004)

[14] *R. Jin*, Inverse problem for upper asymptotic density, The Transactions of American Mathematical Society, 355 (2003), No. 1, pp. 57—78.

[15] *R. Jin*, Inverse problem for upper asymptotic density II, to appear.

[16] *R. Jin*, Solution to the inverse problem for upper asymptotic density, to appear.

[17] *R. Jin*, Freiman's $3k-3+b$ conjecture and nonstandard methods, preprint.

[18] *R. Jin and H. J. Keisler*, Abelian group with layered tiles and the sumset phenomenon, The Transactions of American Mathematical Society, Vol. 355 (2003), No. 1, pp. 79 —97.

[19] *H. J. Keisler and S. Leth*, Meager sets on the hyperfinite time line, The Journal of Symbolic Logic, Vol. 56 (1991), 71—102.

[20] *A. I. Khinchin*, Three pearls of number theory, Translated from the 2d (1948) rev. Russian ed. by F. Bagemihl, H. Komm, and W. Seidel, Rochester, N.Y., Graylock Press, 1952.

[21] *S. C. Leth*, Applications of nonstandard models and Lebesgue measure to sequences of natural numbers. The Transactions of American Mathematical Society, 307 (1988), no. 2, 457–468.

[22] *S. C. Leth*, Sequences in countable nonstandard models of the natural numbers, Studia Logica 47 (1988), no. 3, 243–263.

[23] *S. C. Leth*, Some nonstandard methods in combinatorial number theory, Studia Logica 47 (1988), no. 3, 265–278.

[24] *M. B. Nathanson*, Additive Number Theory–Inverse Problems and the Geometry of Sumsets, Springer, 1996.