SOME OPEN PROBLEMS ON MULTIPLE ERGODIC AVERAGES

NIKOS FRANTZIKINAKIS

1. Problems related to sequences arising from smooth functions

In this section we give a list of problems related to the study of multiple ergodic averages involving iterates given by sequences arising from smooth functions, and related applications to multiple recurrence and number theory.

We are going to restrict ourselves, almost entirely, to a class of non-oscillatory functions that is rich enough to contain several interesting examples. Its formal definition is the following: Let B be the collection of equivalence classes of real valued functions defined on some half-line (c, ∞), where we identify two functions if they agree eventually. A *Hardy field* is a subfield of the ring $(B, +, \cdot)$ that is closed under differentiation. With $\mathcal{H}$ we denote the *union of all Hardy fields.* It is easy to check that if a function belongs in $\mathcal{H}$, then it is eventually monotonic and the same holds for its derivatives, so if $a \in \mathcal{H}$, all limits $\lim_{t\to\infty} a^{(k)}(t)$ exist (they may be infinite). We call a *Hardy sequence* any sequence of the form $([a(n)])$ where $a \in \mathcal{H}$.

An explicit example of a Hardy field to keep in mind is the set $\mathcal{LE}$ that consists of all *logarithmico-exponential functions* (introduced by Hardy in [13]), meaning all functions defined on some half-line (c, ∞) using a finite combination of the symbols $+, -, \times, :, \log, \exp$, operating on the real variable t and on real constants. For example, all rational functions and the functions $t^{\sqrt{2}}$, $t \log t$, $t^{\sqrt{\log\log t}}/\log(t^2 + 1)$ belong in $\mathcal{LE}$. Let us stress though that the set $\mathcal{H}$ is much more extensive than the set $\mathcal{LE}$; it contains all antiderivatives of elements of $\mathcal{LE}$, the Riemann zeta function ζ, the Euler Gamma function Γ, etc.

The main advantage we get by working with elements of $\mathcal{H}$ is that it is possible to relate their growth rates with the growth rates of their derivatives.[1] As a consequence, a single growth condition encodes a lot of useful information and this enables us to give more transparent and appetizing statements.

Background material on Hardy fields can be found in [6, 7, 8, 13, 14, 17].

1.1. **Powers of a single transformation.**

1.1.1. *Hardy sequences of polynomial growth.* To avoid repetition we remark that in this subsection we always work with a family $\mathcal{F} := \{a_1(t), \dots, a_\ell(t)\}$ of functions of *polynomial growth* (meaning $a_i(t)/t^k \to 0$ for some $k \in \mathbb{N}$) that belong to the same Hardy field. With $\text{span}^*(\mathcal{F})$ we denote the set of all *non-trivial* linear combinations of elements of $\mathcal{F}$.

Date: March 2011.

[1]If $a \in \mathcal{H}$ and $b \in \mathcal{LE}$, then there exists a Hardy field that contains both a and b. As a consequence, the limit $\lim_{t\to\infty} a'(t)/b'(t)$ exists (it may be infinite), and so assuming that $a(t), b(t) \to \infty$, we get (using L'Hospital's rule) that the quotients $a(t)/b(t)$ and $a'(t)/b'(t)$ have the same limit as $t \to \infty$. We deduce, for instance, that if $a \in \mathcal{H}$ satisfies $a(t)/t^2 \to \infty$, then $a'(t)/t \to \infty$ and $a''(t) \to \infty$.

We first state two problems from [10] related to the mean convergence of multiple ergodic averages involving iterates given by Hardy sequences. The following result was proved in [10] (the case $\ell = 1$ was first handled in [9]):

Theorem. *Let $a \in \mathcal{H}$ have polynomial growth. Then the sequence $([a(n)])$ is good for multiple convergence of powers if and only if one of the following conditions is satisfied:*

- $|a(t) - cp(t)|/\log t \to \infty$ *for every* $c \in \mathbb{R}$ *and* $p \in \mathbb{Z}[t]$; *or*
- $a(t) - cp(t) \to d$ *for some* $c, d \in \mathbb{R}$; *or*
- $|a(t) - t/m| \ll \log t$ *for some* $m \in \mathbb{Z}$.

For instance, the sequences (n^2), $([n^{3/2}])$, $([n \log n])$, $([n^2 + (\log n)^2])$, $([n^2 + n\sqrt{2} + \log\log n])$ are all good for multiple convergence of powers, but the sequences $([n^2 + \log n])$, $([n^2\sqrt{2} + \log\log n])$ are not good for 1-convergence. Unlike the case of polynomial sequences, if $a \in \mathcal{H}$ satisfies $a(t)/t^{k-1} \to \infty, a(t)/t^k \to 0$ for some $k \in \mathbb{N}$, then the sequence $([a(n)])$ takes odd (respectively even) values in arbitrarily large intervals. As a consequence, when T is the rotation by $1/2$ on the circle and $f := \mathbf{1}_{[0,1/2]}$, the $L^2(\mu)$-limit $\lim_{N\to\infty} \frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} T^{[a(n)]}f$ does not exist for some appropriately chosen Følner sequence $(\Phi_N)_{N\in\mathbb{N}}$ of subsets of $\mathbb{N}$.

The next problem seeks to give similar necessary and sufficient conditions for ℓ-convergence of arbitrary collections of sequences arising from functions of polynomial growth that belong to the same Hardy field. We remind the reader that in such circumstances one is seeking to prove mean convergence for averages of the form

$$\frac{1}{N}\sum_{n=1}^{N} T^{[a_1(n)]}f_1 \cdot \ldots \cdot T^{[a_\ell(n)]}f_\ell. \tag{1}$$

Problem 1. *Let $\mathcal{F}$ be as above. Show that the collection of sequences $\{([a_1(n)]), \ldots, ([a_\ell(n)])\}$ is good for ℓ-convergence of a single transformation if and only if every function $a \in$ span$^*(\mathcal{F})$ satisfies one of the following conditions:*

- $|a(t) - cp(t)|/\log t \to \infty$ *for every* $c \in \mathbb{R}$ *and* $p \in \mathbb{Z}[t]$; *or*
- $a(t) - cp(t) \to d$ *for some* $c, d \in \mathbb{R}$; *or*
- $|a(t) - t/m| \ll \log t$ *for some* $m \in \mathbb{Z}$.

Convergence was proved in [10] under much more restrictive conditions than those advertised here. The collection of sequences $\{([n \log n]), ([n^2 \log n]), \ldots, ([n^\ell \log n])\}$ is an explicit example that is expected to be good for ℓ-convergence of a single transformation but this is not known yet (not even for all weak mixing systems, or all nilsystems).

When the multiple ergodic averages of a collection of Hardy sequences of polynomial growth converge in the mean one would like to have an explicit formula for their limit. In general, such a limit formula can be extremely complicated but when the sequences are in "general position" the limit is expected to be very simple:

Problem 2. *Let $\mathcal{F}$ be as above and suppose that for every function $a \in$ span$^*(\mathcal{F})$ we have $|a(t) - cp(t)|/\log t \to \infty$ for every $c \in \mathbb{R}$ and $p \in \mathbb{Z}[t]$. Show that for every invertible measure preserving system $(X, \mathcal{B}, \mu, T)$ and functions $f_1, \ldots, f_\ell \in L^\infty(\mu)$ we have*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} T^{[a_1(n)]}f_1 \cdot \ldots \cdot T^{[a_\ell(n)]}f_\ell = \mathbb{E}(f_1|\mathcal{I}_T) \cdot \ldots \cdot \mathbb{E}(f_\ell|\mathcal{I}_T) \tag{2}$$

where the convergence takes place in $L^2(\mu)$.

The identity is known when $a_i(t) = t^{c_i}$, $i = 1, \ldots, \ell$, where $c_1, \ldots, c_\ell \in \mathbb{R} \setminus \mathbb{Z}$ are different and positive [10] (this was established first in [4] when $c_i \in (0,1)$, or when the system is weak mixing).[2] A particular collection of sequences for which the identity is not known is the one mentioned before: $\{([n \log n]), ([n^2 \log n]), \ldots, ([n^\ell \log n])\}$. If some function $a \in \text{span}(\mathcal{F})$ satisfies $|a(t) - cp(t)| \ll \log t$ for some $c \in \mathbb{R}$ and $p \in \mathbb{Z}[t]$ with $\deg(p) \geq 2$, then one easily sees that (2) fails for T given by an appropriate rotation on $\mathbb{T}^\ell$.

An intermediate step that would help solve the previous two problems is to find suitable characteristic factors for the relevant multiple ergodic averages. We state a problem from [10] of this sort that we find of independent interest:

Problem 3. *Let* $\mathcal{F}$ *be as above and suppose that* $a_i(t)/\log t \to \infty$ *and* $(a_i(t) - a_j(t))/\log t \to \infty$ *whenever* $i \neq j$. *Show that for every invertible measure preserving system* $(X, \mathcal{X}, \mu, T)$ *the factor* $\mathcal{Z}_T := \bigvee_{d \in \mathbb{N}} \mathcal{Z}_{d,T}$ *is characteristic for mean convergence of the averages* (1).

This is known when for some $\varepsilon > 0$ we have $a_i(t)/t^\varepsilon \to \infty$ and $(a_i(t) - a_j(t))/t^\varepsilon \to \infty$ whenever $i \neq j$ [10], and the methods of [10] (see the proof of Theorem 2.4 there) can be used to show that it also holds when $a_i(t) = ia(t)$ for $i = 1, \ldots, \ell$ and $a(t)/\log t \to \infty$. On the other hand, in the generality stated, the problem is open even for weak mixing systems:

Special Case of Problem 3. *Let* $\mathcal{F}$ *be as above and suppose that* $a_i(t)/\log t \to \infty$ *and* $(a_i(t) - a_j(t))/\log t \to \infty$ *whenever* $i \neq j$. *Show that for every weak mixing system the averages* (1) *converge in the mean to the product of the integrals of the individual functions.*

One can check that the stated assumptions are necessary.

Next we state some problems related to multiple recurrence. The following result was proved in [10] (see also [11] for a special case):

Theorem. *Let* $a \in \mathcal{H}$ *have polynomial growth and suppose that* $|a(t) - cp(t)| \to \infty$ *for every* $c \in \mathbb{R}$ *and* $p \in \mathbb{Z}[t]$. *Then the sequence* $([a(n)])$ *is good for multiple recurrence of powers.*

It follows that the sequences $([n^{\sqrt{2}}])$, $([n \log n])$, $([n^2 + (\log n)^2])$, $([n^2 + \log n])$, $([n^2\sqrt{2} + \log\log n])$ $([n^2 + n\sqrt{2}])$ are all good for multiple recurrence of powers. The previous result does not handle the case $a(t) = cp(t) + d + e(t)$ where p is an integer polynomial with zero constant term, $e \in \mathcal{H}$ is non-negative and converges to zero, and $d \in \mathbb{R}$. If $d = 0$, then one can show that the sequence $([a(n)])$ is good for multiple recurrence of powers. The case where $d \neq 0$ is trickier. For instance, the sequence $([\sqrt{5n+1}])$ is good for multiple recurrence of powers but the sequence $([\sqrt{5n+2}])$ is not good for 1-recurrence for the rotation on the circle by $1/\sqrt{5}$.

Next we state a problem from [10] that seeks to give necessary conditions for ℓ-recurrence of arbitrary collections of sequences arising from functions of polynomial growth that belong to the same Hardy field. We remind the reader that in such circumstances one is seeking to prove that whenever $\mu(A) > 0$ we have $\mu(A \cap T^{-[a_1(n)]}A \cap \cdots \cap T^{-[a_\ell(n)]}A) > 0$ for some $n \in \mathbb{N}$.

Problem 4. *Let* $\mathcal{F}$ *be as above and suppose that for every function* $a \in span^*(\mathcal{F})$ *we have* $|a(t) - cp(t)| \to \infty$ *for every* $c \in \mathbb{R}$ *and* $p \in \mathbb{Z}[t]$. *Show that the collection of sequences* $\{([a_1(n)]), \ldots, ([a_\ell(n)])\}$ *is good for* ℓ*-recurrence of a single transformation.*

[2] More generally, the identity was shown in [10] when the functions $a_1, \ldots, a_\ell$ and their pairwise differences belong to the set $\mathcal{LE} \cap \{a\colon a(t)/t^{k+\varepsilon} \to \infty, a(t)/t^{k+1} \to 0, \text{ for some } k \geq 0 \text{ and } \varepsilon > 0\}$.

As stated above, this is known for $\ell = 1$ [11].

Lastly, we mention an interesting multiple recurrence problem involving a collection of sequences not covered by the previous problem. Using consecutive values of positive fractional powers is expected to produce good multiple return times, a result that fails for integral powers:

Problem 5. *Let c be a positive real number that is not an integer. Show that for every $\ell \in \mathbb{N}$, the collection of sequences $\{([n^c]), ([(n+1)^c]), \ldots, ([(n+\ell)^c])\}$ is good for $(\ell+1)$-recurrence of a single transformation.*

If the integer ℓ is greater than $[c]$, then some non-trivial linear combination of the functions $t^c, (t+1)^c, \ldots, (t+\ell)^c$ converges to 0; so in this particular instance the assumptions of Problem 4 are not be satisfied. The conclusion fails trivially when $c = 1$, to see that it fails for $c = 2, 3, \ldots$ it suffices to consider appropriate rotations on the circle.

1.1.2. *Hardy sequences of super-polynomial growth.* Despite the fact that multiple recurrence and convergence properties of Hardy sequences of polynomial growth are relatively well understood, when it comes down to sequences that grow faster than polynomials, even the most basic problems are open.

Problem 6. *Find an example of a function $a \in \mathcal{H}$ that grows faster than polynomials, meaning $a(t)/t^k \to \infty$ for every $k \in \mathbb{N}$, such that the sequence $[a(n)]$ is good for multiple recurrence and convergence of powers.*

The sequences $([n^{(\log n)^a}])$, $([e^{n^b}])$, where $a > 0$ and $b \in (0,1)$, seem to be natural candidates; unfortunately they are extremely hard to work with. Even when $\ell = 1$ the relevant exponential sum estimates needed to prove convergence appear to be out of reach in most cases; for the first sequence such estimates are available only when $a \in (0, 1/2)$ [15], and no estimates are available for the second sequence. On the other hand, a slower growing sequence, like the sequence $([n^{\log\log n}])$ may be easier to handle. But even for this sequence, 2-recurrence and 2-convergence is not known for all weak mixing systems or all nilsystems.

1.1.3. *Hardy sequences evaluated at the primes.* With p_n we denote the n-th prime.

Problem 7. *Let $c \in \mathbb{R} \setminus \mathbb{Z}$ be a positive real number. Show that the sequence $([p_n^c])$ is good for multiple recurrence and convergence of powers.*

Proving multiple recurrence is trivial when $c < 1$ since in this case the sequence $([p_n^c])$ misses at most finitely many positive integer values. It is known that if $c \in \mathbb{R} \setminus \mathbb{Z}$ is positive, then the sequence of fractional parts $(\{p_n^c\})$ is equidistributed in the unit interval (see [18] or [22] for $c < 1$ and [16] for $c > 1$). Probably the techniques used to prove these equidistribution results suffice to prove 1-recurrence and 1-convergence (it suffices to show that the sequence $(\{p_n^c \alpha\})$ is equidistributed in the unit interval for every non-zero $\alpha \in \mathbb{R}$), but the problem is open for ℓ-recurrence and ℓ-convergence when $\ell \geq 2$.

1.1.4. *Oscillatory sequences.* All the previous problems deal with sequences that do not oscillate. Multiple recurrence and convergence properties of oscillatory sequences are not well studied and analyzing some simple looking sequences leads to interesting problems:

Problem 8. *Show that the sequence $([n \sin n])$ is good for multiple convergence of powers.*

Quite likely one can say more; the averages $\frac{1}{N}\sum_{n=1}^{N} T^{a(n)}f_1 \cdot T^{2a(n)}f_2 \cdot \ldots \cdot T^{\ell a(n)}f_\ell$ have the same limiting value when $a(n) = [n \sin n]$ and $a(n) = n$. This is known for $\ell = 1$, it follows from equidistribution results in [1] (see also related results in [2, 3]). As far as I know the problem has not been studied when $\ell \geq 2$, even for particular classes of measure preserving systems, like nilsystems or weakly mixing systems.

1.2. **Commuting transformations.** As mentioned before if a Hardy sequence has polynomial growth and stays away from constant multiples of integer polynomials, then it is going to be good for multiple recurrence of powers. The next problem seeks to extend this to the case of commuting transformations. We remind the reader that in such circumstances one is seeking to prove that whenever $\mu(A) > 0$ we have $\mu(A \cap T_1^{-[a(n)]}A \cap \cdots \cap T_\ell^{-[a(n)]}A) > 0$ for some $n \in \mathbb{N}$.

Problem 9. *Let $a \in \mathcal{H}$ have polynomial growth and suppose that $|a(t) - cp(t)| \to \infty$ for every $c \in \mathbb{R}$ and $p \in \mathbb{Z}[t]$. Then the sequence $([a(n)])$ is good for multiple recurrence of commuting transformations.*

The proof that any such $([a(n)])$ is good for multiple recurrence of powers relies crucially on the precise algebraic structure of suitable characteristic factors for the corresponding multiple ergodic averages; an advantage that is lost when one works with commuting transformations.

Problem 10. *Let $(X, \mathcal{X}, \mu)$ be a probability space, $T_1, \ldots, T_\ell \colon X \to X$ be commuting invertible measure preserving transformations, and $f_1, \ldots, f_\ell \in L^\infty(\mu)$. Show that for every positive real number c the following limit exists in $L^2(\mu)$*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} T_1^{[n^c]}f_1 \cdot \ldots \cdot T_\ell^{[n^c]}f_\ell, \tag{3}$$

and if c is not an integer, then it is equal to $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} T_1^n f_1 \cdot \ldots \cdot T_\ell^n f_\ell$.

For $c = 1$ the existence of the limit (3) is known [20]. The case $0 < c < 1$ can be easily reduced to the case $c = 1$. So the interesting case is when $c > 1$ in which case the problem is open even when $\ell = 2$ and all transformations are assumed to be weak mixing.

Problem 11. *Let $(X, \mathcal{X}, \mu)$ be a probability space, $T_1, \ldots, T_\ell \colon X \to X$ be commuting invertible measure preserving transformations, and $f_1, \ldots, f_\ell \in L^\infty(\mu)$. Let $c_1, \ldots, c_\ell \in \mathbb{R} \setminus \mathbb{Z}$ be positive and distinct. Show that*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} T_1^{[n^{c_1}]}f_1 \cdot \ldots \cdot T_\ell^{[n^{c_\ell}]}f_\ell = \mathbb{E}(f_1|\mathcal{I}_{T_1}) \cdot \ldots \cdot \mathbb{E}(f_\ell|\mathcal{I}_{T_\ell}),$$

where the convergence takes place in $L^2(\mu)$.

The identity is known when all the transformations are equal [10] and is also known when all the exponents c_i are smaller than 1 [10] (in which case the assumption that the transformations commute is not needed). The interesting case is when $\ell \geq 2$ and all the exponents are greater than 1. Easy examples show that the limit formula fails if one of the powers is an integer different than 1.

1.3. **Configurations in the primes.** As we mentioned in the introduction, the theorem of Szemerédi on arithmetic progressions [19], and its polynomial extension [5], have been instrumental in proving that the primes contain arbitrarily long arithmetic progressions [12] and polynomial progressions [21]. It is then natural to expect that the various available Hardy field extensions of the theorem of Szemerédi [10, 11] can be used to prove that the primes contain the corresponding Hardy field patterns. For instance:

Problem 12. *Let $\ell \in \mathbb{N}$ and $c, c_1, \ldots, c_\ell$ be positive real numbers. Show that the prime numbers contain patterns of the form*

$$\{m, m+[n^c], m+2[n^c], \ldots, m+\ell[n^c]\} \quad \text{and} \quad \{m, m+[n^{c_1}], \ldots, m+[n^{c_\ell}]\}.$$

When all exponents are rational the existence of such patterns follows immediately from [21].

References

[1] D. Berend, G. Kolesnik. Distribution modulo 1 of some oscillating sequences. *Israel J. Math.* **71** (1990), no. 2, 161-179.

[2] D. Berend, M. Boshernitzan, G. Kolesnik. Distribution modulo 1 of some oscillating sequences II. *Israel J. Math.* **92** (1995), no. 1-3, 125-147.

[3] D. Berend, M. Boshernitzan, G. Kolesnik. Distribution modulo 1 of some oscillating sequences III. *Acta Math. Hungar.* **95** (2002), no. 1-2, 1-20.

[4] V. Bergelson, I. Håland-Knutson. Weak mixing implies mixing of higher orders along tempered functions. *Ergodic Theory Dynam. Systems* **29** (2009), no. 5, 1375–1416

[5] V. Bergelson, A. Leibman. Polynomial extensions of van der Waerden's and Szemerédi's theorems. *J. Amer. Math. Soc.* **9** (1996), 725–753.

[6] M. Boshernitzan. New "Orders of Infinity". *J. Analyse Math.* **41** (1982), 130–167.

[7] M. Boshernitzan. Homogeneously distributed sequences and Poincaré sequences of integers of sublacunary growth. *Monatsh. Math.* **96** (1983), no. 3, 173–181.

[8] M. Boshernitzan. Uniform distribution and Hardy fields. *J. Analyse Math.* **62** (1994), 225–240.

[9] M. Boshernitzan, G. Kolesnik, A. Quas, M. Wierdl. Ergodic averaging sequences. *J. Analyse Math.* **95** (2005), 63–103.

[10] N. Frantzikinakis. Multiple recurrence and convergence for Hardy sequences of polynomial growth. *J. Analyse Math.* **112** (2010), 79–135.

[11] N. Frantzikinakis, M. Wierdl. A Hardy field extension of Szemerédi's theorem. *Adv. Math.* **222**, (2009), 1–43.

[12] B. Green, T. Tao. The primes contain arbitrarily long arithmetic progressions. *Annals Math.* **167** (2008), 481–547.

[13] G. Hardy. Properties of logarithmico-exponential functions. *Proc. London Math. Soc.* (2) **10** (1912), 54–90.

[14] G. Hardy. Orders of Infinity. *Cambridge Tracts in Math. and Math. Phys.*, 12 (2nd edition), Cambridge, (1924).

[15] A. Karatsuba. Estimates for trigonometric sums by Vinogradov's method, and some applications. *Trudy Mat. Inst. Steklov*, **112**, (1971), 241–255; English Transl., *Proc. Steklov Inst. Math.* **112** (1973), 251–265.

[16] D. Leitmann. On the uniform distribution of some sequences. *J. London Math. Soc.* (2) **14** (1976), no. 3, 430-432.

[17] M. Rosenlicht. Hardy fields. *J. Math. Anal. Appl.* **91** (1983), no. 2, 297–311.

[18] I. Stux. On the uniform distribution of prime powers. *Comm. Pure Appl. Math.* **27** (1974), 729-740.

[19] E. Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith.* **27** (1975), 299–345.

[20] T. Tao. Norm convergence of multiple ergodic averages for commuting transformations. *Ergodic Theory Dynam. Systems* **28** (2008), no. 2, 657–688.

[21] T. Tao, T. Ziegler. The primes contain arbitrarily long polynomial progressions. *Acta Math.* **201** (2008), 213–305.

[22] D. Wolke. Zur Gleichverteilung einiger Zahlenfolgen. *Math. Z.* **142** (1975), 181-184.

(Nikos Frantzikinakis) University of Crete, Department of mathematics, Knossos Avenue, Heraklion 71409, Greece
E-mail address: frantzikinakis@gmail.com