SOME PROBLEMS IN COMBINATORIAL NUMBER THEORY

[1] Mei-Chu Chang
Department of Mathematics
University of California
Riverside, CA 92521
mcc@math.ucr.edu

— *dedicated to Mel Nathanson on his 60th birthday*

Abstract

We state and discuss various problems in the general area of arithmetic combinatorics and recent developments related to the 'sum-product phenomenon' in the ring of integers, the real and complex numbers and finite fields. In particular, we discuss applications and connections to the theory of exponential sums, Burgess' estimate, the subspace theorem and to Szemeredi-Trotter type results.

In recent years, some of the developments in combinatorial number theory relate to a line of research initiated by Erdős and Szemerédi in their seminal paper [ES] on the size of sum and product sets of sets of integers. The motivation for this renewed interest has several distinct sources. One of them belongs to harmonic analysis and the so-called Kakeya problem. Another relates to the search for deterministic forms of randomness in computer science. It is certainly not our purpose here to review how the Erdős-Szemerédi sum-product problems are relevant to these other fields and what progress they have generated. Our presentation will be mostly limited to a discussion of what is new for the questions raised in [ES] (obviously a lot remains to be solved). We will also explain the mathematical connections and tools that have been brought into play. They will clearly demonstrate that the questions raised in [ES] have significant connections to several major themes in number theory, such as character sums and the subspace theorem.

Let A and B be subsets of a ring. The *sum set* and the *product set* of A and B are

$$A+B=\{a+b\,:\,a\in A,\ \text{and } b\in B\}$$

and

$$AB=\{ab\,:\,a\in A,\ \text{and } b\in B\},$$

respectively. The study of sum-product sets aims in particular to estimate the sizes of $A+A$ and AA asymptotically when N is very large. The following conjecture [ES] is well-known.

[1] partially supported by NSA grant No. MDA 904-03-1-0045.
2000 Mathematics Subject Classification. 05A99.

1. Erdős-Szemerédi Conjecture. *For all $\varepsilon > 0$,*

$$\min_{\substack{A\subset\mathbb{Z},\\ |A|=N}} |A+A| + |AA| > c_\varepsilon N^{2-\varepsilon}.$$

What they proved is that for any $A \subset \mathbb{R}$, $|A+A| + |AA| > c|A|^{1+\delta}$, where $\delta > 0$ is an absolute constant. Results with explicit δ are: $\delta = \frac{1}{31}$ by Nathanson [N], $\delta = \frac{1}{6}$ by Chen [Ch], $\delta = \frac{1}{15}$ by Ford [F], $\delta = \frac{1}{4}$ by Elekes [E], and $\delta = \frac{3}{11+\varepsilon}$ by Solymosi [S].

At this point it is fair to say that there is a consensus that the Erdős-Szemerédi conjecture, if true, is deep and its solution will likely be a substantial achievement. But it is unclear what kind of mathematics will be involved. The limited progress in research has brought a wide area of techniques into play (graph theory, incidence geometry, harmonic analysis, algebraic number theory etc.). One modest and immediate goal would be to

2. *Improve Solymosi's bound.*

Solymosi made the following more general conjecture.

3. Solymosi Conjecture. $\displaystyle\min_{\substack{A,B,C\subset\mathbb{Z}\\ |A|=|B|=|C|=N}} |A+B| + |AC| > N^{2-\varepsilon}$, *for all $\varepsilon > 0$.*

The next problem is a special case of Problem 3.

4. *Is it true that if $|AA| < |A|^{1+\varepsilon}$, then $|A+A| > |A|^{2-\delta}$, where $\delta = \delta(\varepsilon) \to 0$ as $\varepsilon \to 0$? or more generally that $|A+B| > |A|^{2-\delta}$, for all B with $|A| = |B|$?*

The answer is yes [C2], if $A \subset \mathbb{Z}$. If $A \subset \mathbb{R}$, then the more general conclusion is true [C4] under the stronger assumption that $|AA| < K|A|$ for some (possibly large) constant K. The proof of the last statement uses some consequences of the subspace theorem of Evertse, Schlickewei and Schmidt [ESS].

Switching addition and multiplication, one may ask the following counterpart of Problem 4.

5. *Is it true that if $|A+A| < |A|^{1+\varepsilon}$, then $|AA| > |A|^{2-\delta}$, where $\delta = \delta(\varepsilon) \to 0$ as $\varepsilon \to 0$? or more generally that $|AB| > |A|^{2-\delta}$, for all B with $|A| = |B|$?*

Again, the answer is yes for $A \subset \mathbb{R}$, if $|AA| < K|A|$ for some (possibly large) constant K. (See [C3], [C4].)

Let $hA = A + \cdots + A$, and $A^h = A \cdots A$ be the h-fold sumset and h-fold product set of A. Erdős and Szemerédi made the following more general conjecture.

6. Erdős-Szemerédi Conjecture'. *Let h be fixed. Then*

$$\min_{\substack{A\subset\mathbb{Z},\\ |A|=N}} |hA| + |A^h| > c_\varepsilon N^{h-\varepsilon},$$

for all $\varepsilon > 0$.

Regarding Problem 6, Elekes, Nathanson and Ruzsa [ENR] have the following result.

Theorem (Elekes-Nathanson-Ruzsa).

$$|hA|\,|A^h| > cN^{3-2^{1-h}}.$$

They and Konyagin made the following conjecture.

7. Elekes-Nathanson-Ruzsa-Konyagin Conjecture. *For any $b \in \mathbb{N}$, there exists $h = h(b)$ such that for any A with $|A| > 1$, $|hA|\,|A^h| > c|A|^b$.*

In [BC1], Bourgain and Chang proved the conjecture is true for $A \subset \mathbb{Z}$ with $h \sim c^{b^4}$. The method relies on the prime factorization for integers, which suggests that it cannot be generalized to the real numbers.

Let $\mathcal{G} \subset A \times A$ be a graph. Then the sum and product along the graph $\mathcal{G}$ are the sets

$$A \overset{\mathcal{G}}{+} A = \{a + a' : (a, a') \in \mathcal{G}\}$$

and

$$A \overset{\mathcal{G}}{\times} A = \{aa' : (a, a') \in \mathcal{G}\},$$

respectively. Erdős and Szemerédi asked the following question.

8. Erdős-Szemerédi's Question. *Assuming $|\mathcal{G}| > |A|^{1+\varepsilon}$, is it true that $|A \overset{\mathcal{G}}{+} A| + |A \overset{\mathcal{G}}{\times} A| > |\mathcal{G}|^{1-\varepsilon'}$, for all $\varepsilon' > 0$?*

Assume that $|\mathcal{G}| > \delta|A|^2$. It is known that if $A \subset \mathbb{Z}$ and $|A \overset{\mathcal{G}}{+} A| < c|A|$, then $|A \overset{\mathcal{G}}{\times} A| > |\mathcal{G}|^{1-\varepsilon'}$. Also, if $A \subset \mathbb{R}$ and $|A \overset{\mathcal{G}}{\times} A| < c|A|$, then $|A \overset{\mathcal{G}}{+} A| > C(\delta, c)|\mathcal{G}|$. (See [C3] and [C2].)

The restrictive assumption $|\mathcal{G}| > \delta|A|^2$ allows one to apply the following theorem [LR].

Laczkovich-Ruzsa Theorem. *Assume that $\mathcal{G} \subset A \times A$ satisfies $|\mathcal{G}| > \delta|A|^2$ and that $|A \overset{\mathcal{G}}{+} A| < C\,|A|$. Then there is a subset $A' \subset A$ such that*

$$|A' + A'| < K|A|$$

and

$$|(A' \times A') \cap \mathcal{G}| > \frac{1}{K}|A|^2,$$

where $K = K(\delta, C)$.

Return to Question 8. In the original question one assumes $A \subset \mathbb{Z}$ and $|\mathcal{G}| = \delta|A|$. However, for subsets $A \subset \mathbb{R}$ this assumption on $|\mathcal{G}|$ does not give the right conclusion. For example, if one takes

$$A = \{\sqrt{i} \pm \sqrt{j} : 1 \le i, j \le k, \text{ and } i, j \text{ are square free }\}$$

and $\mathcal{G} = \{(\sqrt{i} + \sqrt{j}, \sqrt{i} - \sqrt{j}) : i, j \text{ as in } A\}$, then $|A| \sim |\mathcal{G}| \sim k^2$ and $|A \overset{\mathcal{G}}{+} A| \sim |A \overset{\mathcal{G}}{\times} A| \sim k$. Hence the assumption that $|\mathcal{G}| > |A|^{1+\varepsilon}$ is essential for $A \subset \mathbb{R}$.

Using the Szemerédi-Trotter Theorem as in Elekes' proof, one may show

$$|A \overset{\mathcal{G}}{+} A| + |A \overset{\mathcal{G}}{\times} A| > \frac{|\mathcal{G}|^{\frac{3}{2}}}{|A|}.$$

The Balog-Szemerédi Theorem is a powerful tool for studying sum-product problems. Below is the Gowers' version [G] (or see [C7] for an explicit expression of δ).

Balog-Szemerédi-Gowers Theorem. *Let $A \subset \mathbb{Z}$ with $|A| = N$. If for some constant $\alpha > 0$,*

$$|\{(a_1, a_2, a_3, a_4) \,:\, a_i \in A,\ a_1 - a_2 + a_3 - a_4 = 0\}| > \alpha N^3,$$

then there is a subset $A' \subset A$ satisfying

$$|A'| > \alpha^{1+\varepsilon} N$$

and

$$|A' - A'| < \delta N,$$

where $\delta = 2^{21}(\alpha \log \frac{1}{\alpha})^5$.

9. *It would be interesting to know the optimal dependence of δ on α.*

The following problem posed by Ruzsa turns out to be related to Problem 8. (See [C7].)

10. Ruzsa's Distance Conjecture. *Let $k_1, \cdots, k_N \in \mathbb{Z}$. Define $D = \{k_i^2 + k_j^2 : 1 \leq i, j \leq N\}$. Then $|D| > N^{2-\varepsilon}$.*

In [C7] we used the result of Bombieri-Granville-Pintz [BGP] and Freiman's Theorem to show that $|D| > N(\log N)^{\frac{1}{12}-\varepsilon}$.

We think an affirmative answer to the next question would have applications to coding theory.

11. Noncommutative Setting. *Let $A \subset Sym(d, \mathbb{Z}_p)$ be a subset of symmetric matrices over $\mathbb{Z}_p$ with $p >> 0$ and $|A| >> 0$. Then is $|A + A| + |AA| > |A|^{1+\varepsilon}$ for some $\varepsilon > 0$?*

This is true for $A \subset \mathrm{Sym}(d, \mathbb{R})$.

As for sum-product in $\mathbb{F}_p$ for p prime, there is the following theorem. It was first proved by Bourgain, Katz and Tao [BKT] for $p^\delta < |A| < p^{1-\delta}$, then improved by Bourgain, Glibichuk and Konyagin [BGK].

Theorem BKT-BGK. *Let p be a prime. Given $\delta > 0$, there is $\varepsilon = \varepsilon(\delta) > 0$ such that if $A \subset \mathbb{F}_p$ and*

$$1 < |A| < p^{1-\delta},$$

then

$$|2A| + |A^2| > c|A|^{1+\epsilon},$$

where $c = c(\delta)$.

12. *Find good explicit bounds on ε.*

See the recent work of Hart-Iosevich-Solymosi [HIS], M. Garaev [Ga], and Katz-Shen [KS] on this question.

Warning. *For k fixed, there are infinitely many p such that there exists $A \subset \mathbb{F}_p$ with $|A| \sim p^{\frac{1}{2}}$, $|A^k| < p^{\frac{3}{4}}$, and $|kA| < kp^{\frac{3}{4}}$.*

Proof.
Claim. There are infinitely many primes p such that $\mathbb{F}_p^*$ has a multiplicative subgroup of order $\sim p^{\frac{3}{4}}$.
Proof. We will show that there are infinitely many primes p such that $p-1$ has a divisor $q \sim p^{\frac{1}{4}}$.

Let $\Lambda(n)$ be the Von Mangoldt function defined on $\mathbb{Z}^+$,

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^\alpha \text{ for some } \alpha > 0 \\ 0 & \text{otherwise,} \end{cases}$$

and let

$$\psi(x; q, 1) = \sum_{\substack{n \le x \\ n \equiv 1 \bmod q}} \Lambda(n). \tag{12.1}$$

By Bombieri-Vinogradov Theorem (see Theorem 17.1 in [IK]), we have

$$\sum_{1<q<2x^{\frac{1}{4}}} \left|\psi(x;q,1) - \frac{x}{\phi(q)}\right| < \frac{x}{\log x},$$

where ϕ is the Euler-phi function.

Applying the Bombieri-Vinogradov Theorem again with $\frac{x}{2}$ in combination with the inequality above, (taking possibly fewer summands) we have

$$\sum_{\frac{1}{2}x^{\frac{1}{4}}<q<x^{\frac{1}{4}}} \left|\psi(x;q,1) - \psi\big(\frac{x}{2};q,1\big) - \frac{x}{2\phi(q)}\right| < \frac{2x}{\log x}.$$

It follows that for some $\frac{1}{2}x^{\frac{1}{4}} < q < x^{\frac{1}{4}}$,

$$\left|\psi(x;q,1) - \psi\big(\frac{x}{2};q,1\big) - \frac{x}{2\phi(q)}\right| < \frac{4x^{\frac{3}{4}}}{\log x}. \tag{12.2}$$

Next, from (12.1), it is clear that

$$\psi(x;q,1) = \sum_{\substack{p \le x \\ p\equiv 1 \bmod q}} \log p \quad + O(\sqrt{x}).$$

Hence, (12.2) implies

$$\left| \sum_{\substack{\frac{x}{2} < p \le x \\ p \equiv 1 \bmod q}} \log p - \frac{x}{2\phi(q)}\right| < \frac{4x^{\frac{3}{4}}}{\log x} \quad + O(\sqrt{x}).$$

Therefore,

$$\sum_{\substack{\frac{x}{2}<p\leq x\\ p\equiv 1 \bmod q}} \log p > \frac{x}{2\phi(q)} - \frac{5x^{\frac{3}{4}}}{\log x} > \frac{x}{3\phi(q)} > 0,$$

since $\phi(q) < q \sim x^{\frac{1}{4}}$.

This shows that for any x large, there is at least one prime $\frac{1}{2}x < p \leq x$, and an integer $\frac{1}{2}x^{\frac{1}{4}} < q < x^{\frac{1}{4}}$, such that $p \equiv 1 \bmod q$. □

Take $H < \mathbb{F}_p^*$ with $|H| \sim p^{\frac{3}{4}}$, a subgroup provided by the claim. For $\delta > 0$, define $\rho(\delta)$ such that

$$p^{\rho(\delta)} = \max_{x\in\mathbb{F}_p} |H \cap \{x, x+1, \cdots, x+[p^\delta]\}|.$$

Clearly, $\rho(\delta) \leq \delta$.

It is easy to see that $\rho(\delta) \geq \delta - \frac{1}{4}$. Indeed,

$$\begin{aligned} p^{\rho(\delta)} &\geq \frac{1}{p}\sum_{x\in\mathbb{F}_p}\sum_{j=1}^{[p^\delta]} \chi_H(x+j)\\ &= \frac{|H|p^\delta}{p}\\ &= p^{\delta-\frac{1}{4}}. \end{aligned}$$

In particular, $\rho(\frac{1}{4}) \leq \frac{1}{4}$ and $\rho(1) \geq \frac{3}{4}$. Moreover, from the definition of ρ, we have

$$\rho(\delta) < \rho(\delta + \log_p 2) \leq \rho(\delta) + \log_p 2.$$

Hence, there exists δ_0 such that $\rho(\delta_0) \sim \frac{1}{2}$.

Let $x \in \mathbb{F}_p$ such that $|H \cap \{x, x+1, \cdots, x+[p^{\delta_0}]\}| = p^{\rho(\delta_0)}$. Take $A = H \cap \{x, x+1, \cdots, x+[p^{\delta_0}]\}$. Then $A^k \subset H$ and $|kA| \leq k[p^{\delta_0}] \sim kp^{\frac{3}{4}}$. □

Remark. M. Garaev pointed out that for our purpose we could take $H = \{g^n : n < p^{3/4}\}$ in the warning, where g is a primitive root modulo p, instead of taking a subgroup H. So there is no need of the Bombieri-Vinogradov Theorem. This works for any p and we have $|A| \sim p^{1/2}$ and $|kA| + |A^k| < kp^{3/4}$.

The above example shows that the analogues of the generalized Erdős-Szemerédi Conjecture and Elekes-Nathanson-Ruzsa-Konyagin Conjecture for $\mathbb{F}_p$ are false. Below is a modification of Problem 7.

13. *Let $A \subset \mathbb{F}_p$ with $|A| = p^a$. Then for all $b < \frac{1}{2}$, is it true that $|kA| + |A^k| > p^b$ for $k > k(a,b)$?*

In [BGK] the following result was proved.

Given $\varepsilon > 0$, for all $H < \mathbb{F}_p^*$ with $|H| > p^\varepsilon$, there exists $\delta = \delta(\varepsilon)$ such that

$$\max_{(a,p)=1} \left| \sum_{x\in H} e_p(ax) \right| \leq c|H|^{1-\delta}.$$

14. *Make δ more explicit.*

For $\varepsilon > \frac{1}{4}$, explicit bounds are due to Heath-Brown and Konyagin.

Let χ be a nontrivial multiplicative character (mod p) and let $I \subset \{0, 1, \cdots, p-1\}$ be an arbitrary interval with $|I| > p^{\frac{1}{4}+\varepsilon}$. We have the

Burgess' estimate. $|\sum\limits_{x\in I} \chi(x)| < c|I|^{1-\delta}$, *with* $\delta = \delta(\varepsilon)$.

15. *Prove the estimate above under the weaker assumption $|I| > p^{\varepsilon}$, for some $\varepsilon > 0$.*

An affirmative answer to problem 15 would imply the next problem, which is obviously of relevance to the well-known problem on the least quadratic non-residue.

16. *Let $I \subset \{0, 1, \cdots, p-1\}$ be an interval with $|I| > p^{\varepsilon}$. Prove that $I^k = I \cdots I = \mathbb{F}_p^*$, for some k.*

Note that the case for $|I| > p^{\frac{1}{4}+\varepsilon}$ follows from Burgess' estimate, and we thus obtain $k < k(\varepsilon)$.

17. *Given $\varepsilon > 0$, for any $H < \mathbb{F}_p^*$ with $|H| > p^{\varepsilon}$, we have $kH = \mathbb{F}_p$ for some $k < k(\varepsilon)$. Find a good dependence of k on ε.*

Bourgain, Glibichuk and Konyagin [BGK] proved that $\log k(\varepsilon) < c(\frac{1}{\varepsilon})^c$.

18. *We say a subgroup $H < \mathbb{F}_p^*$ is uniformly distributed in $\mathbb{F}_p$ if*

$$\max_{(a,p)=1} |\sum_{x\in H} e_p(ax)| < o(|H|).$$

How large must $|H|$ be?

Again, in [BGK] Bourgain, Glibichuk and Konyagin proved that

$$\log|H| > \frac{\log p}{(\log\log p)^{\rho}},$$

where ρ is an explicit constant. Perhaps $\log|H| \gg \log\log p$ is sufficient.

19. *Find an analogue of the Szemerédi-Trotter Theorem for $\mathbb{F}_p$ or $\mathbb{C}$.*

In [BKT], a Szemerédi-Trotter type Theorem for point-line incidence in $\mathbb{F}_p^n$ is proven. This result is significantly weaker (although nontrivial) than the corresponding result in the Euclidean plane. The argument in [BKT] is by contradiction, it leads to violation of the sum-product theorem in $\mathbb{F}_p$ and is therefore non-explicit.

An extension of the Szemerédi-Trotter Theorem to "pseudo-line systems" has been obtained in the Euclidean plane. (See for instance, work by J. Pach and M. Sharir [PS].)

20. *Is there a variant of the Szemerédi-Trotter Theorem when straight lines are replaced by families of algebraic curves over $\mathbb{F}_p$?*

Progress on these questions would have significant applications to the theory of exponential sums. For some recent work related to Question 20, see [HIS].

Acknowledgement. The author would like to thank M. Garaev for helpful comments.

References

[BC1]. J. Bourgain, M-C. Chang, *On the size of k-fold Sum and Product Sets of Integers*, JAMS.

[BC2]. ———, *Exponential sum estimates over subgroups and almost subgroups of $\mathbb{Z}_q^*$, where q is composite with few prime factors*, GAFA.

[BGK]. J. Bourgain, A. Glibichuk, S. Konyagin, *Estimate for the number of sums and products and for exponential sums in fields of prime order*, submitted to J. London MS.

[BKT]. J. Bourgain, N. Katz, T. Tao, *A sum-product estimate in finite fields and their applications*, GAFA 14 (2004), n1, 27–57.

[C1]. M-C. Chang, *A Polynomial bound in Freiman's theorem*, Duke Math. J. Vol. 113, No. 3, (2002), 399-419.

[C2]. ———, *Erdös-Szemeredi problem on sum set and product set*, Annals of Math. 157 (2003), 939-957.

[C3]. ———, *Factorization in generalized arithmetic progressions and applications to the Erdös-Szemerédi sum-product problems*, GAFA Vol. 113, (2002), 399-419.

[C4]. ———, *Sum and product of different sets*, preprint 2004.

[C5]. ———, *Additive and multiplicative structure in matrix spaces*, preprint 2004.

[C6]. ———, *A sum-product estimate in algebraic division algebras*, Israel JM.

[C7]. ———, *On problems of Erdös and Rudin*, J. of Functional Analysis, 207 (2004), 444-460.

[Ch]. Y.G. Chen, *Private communication.*

[E]. G. Elekes, *On the number of sums and products*, Acta Arithm., 81, 4 (1997), 365-367.

[ENR]. G. Elekes, M. Nathanson, I. Ruzsa, *Convexity and sumsets*, J. Number Theory, (to appear).

[ER]. G. Elekes, J. Ruzsa, *Few sums, many products*, Studia Sci. Math. Hungar. 40 (2003).

[ES]. P. Erdös, E. Szemerédi, *On sums and products of integers*, In P. Erdös, L. Alpàr, G. Haláz (editors), Stud. Pure Math., to the memory of P. Turán, p. 215–218.

[ESS]. J.-H. Evertse, H. Schlickewei, W. Schmidt, *Linear equations in variables which lie in a multiplicative group*, Annals Math 155, (2002), 807-836.

[Ga]. M. Garaev, *An explicit sum-product estimate in $\mathbb{F}_p$*, (preprint).

[G]. T. Gowers, *A new proof of Szemeredi's theorem for arithmetic progressi ons of length four*, GAFFA 8 (1998), no3, 529-551.

[HIS]. D. Hart, A. Iosevich, J. Solymosi, *Sum product estimates in finite fields via Kloosterman sums*, IMRN (to appear).

[IK]. H. Iwaniec, E. Kowalski, *Analytic number theory*, AMS Colloquium Publications, Vol 53 (2004).

[KS]. N. Katz, C.-Y. Shen, *A slight improvement to Garaev's sum product estimate (preprint)*.■

[LR]. Laczkovich, I. Ruzsa, *The number of homothetic subsets,*, in 'The mathemattics of P. Erdős, II. (R.L. Grham, J. Nesetril, eds.), Springer, Algorithms Combin. 14 (1997), 294-302.

[N]. M. Nathanson, *Additive Number Theory*, Springer (1996).

[PS]. J. Pach, M. Sharir, *On the number of incidences between points and curves*, Combinatorics, Probability and Computing 7 (1998), 121–127.

[R]. W. Rudin, *Trigonometric series with gaps*, J. Math. Mech. 9 (1960), 203–227.

[ST]. E. Szemerédi, W. Trotter, *Extremal problems in Discrete Geometry*, Combinatorics, 3 (3-4), 387–392 (1983).