

STATEMENT OF RESEARCH INTERESTS

MATTHEW L. SMITH

1. Introduction. My research has been in the area of Diophantine equations. The fundamental problems in this area concern the conditions under which a given equation or system of equations has an integral solution. Perhaps the most celebrated problem in the field of Diophantine equations is Fermat's Last Theorem, now a theorem of Wiles [26], which states that for any integer $n \geq 3$, the equation $x^n + y^n = z^n$ has no solutions for which $x, y, z \in \mathbb{Z} \setminus \{0\}$. Another well-known problem in this area is Waring's problem, which asks, given a natural number $k \geq 2$, what is the smallest number s such that the equation

$$x_1^k + \dots + x_s^k = N$$

has solutions $(x_1, \dots, x_s) \in \mathbb{N}^s$ for all sufficiently large N ?

There are further examples of Diophantine problems in which the logical approach to determine the circumstances under which a given system of one or more equations possesses a solution in a given set of numbers \mathcal{A} is to determine instead the conditions under which the system does *not* possess a solution. Given a set $\mathcal{A} \subset \mathbb{N}$, we define $\mathcal{A}_N = \mathcal{A} \cap [1, N]$ and define the *density* δ_N of \mathcal{A}_N to be

$$\delta_N = |\mathcal{A}_N|/N.$$

We further define the *upper density* of \mathcal{A} by $\limsup_{N \rightarrow \infty} \delta_N$. If the set \mathcal{A} contains no solutions to a given system of one or more equations, one may ask how the density behaves as N increases, and in particular whether or not the upper density is zero. If, for example, the upper density of a solution-free set may be shown to be zero, then the contrapositive of this remark states that a set with positive upper density must necessarily contain a solution to the given equation or system of equations. Such problems have recently risen in prominence as a result of the celebrated work of Gowers [10, 11] and Green and Tao [14] on long arithmetic progressions.

My research has been primarily concerned with problems of the latter type in the field of additive combinatorics, which deals with counting additive structures in sets. Given a system of equations, I have been concerned not only with showing that a solution-free set of natural numbers necessarily has zero upper density, but also with determining an upper bound for the rate of decay of the density δ_N of such a set. In particular, my research has been concerned with the first instance of a system of equations which is not purely linear and the analysis of which is dominated by the nonlinear components. The methods I use are primarily harmonic analytic, and combine the classical Hardy-Littlewood circle method as used in the work of Roth [17] on three-term progressions with the methods developed by Gowers [11] in his proof of Szemerédi's theorem on sets with long arithmetic progressions.

2. Motivation. Much of the initial interest in solution-free sets associated to systems of one or more equations grew out of a 1936 conjecture by Erdős and Turán [7] that a set \mathcal{A} of natural numbers with positive upper density necessarily contains arbitrarily long arithmetic progressions, or sequences of the form $\{a, a + d, a + 2d, \dots\}$, where $d \in \mathbb{N} \setminus \{0\}$. That is to say, \mathcal{A} necessarily contains solutions to the system of $k - 2$ linear equations

$$x_i - 2x_{i+1} + x_{i+2} = 0 \quad (1 \leq i \leq k - 2)$$

for any $k \geq 3$. The case $k = 3$ was settled in 1953 by Roth [17], who proved using a version of the Hardy-Littlewood circle method that if \mathcal{A} contains no three-term progressions, then the density δ_N decays at least as rapidly as $(\log \log N)^{-1}$. Szemerédi [21, 22] showed using combinatorial and analytic methods that a set containing no progressions of length k for any $k \geq 3$ must have zero upper density, but was unable to establish a reasonable upper bound for the decay of δ_N .

It was not until recently that Gowers [10, 11], using analytic methods, established that the density of a set containing no progressions of length k decays at least as rapidly as $(\log \log N)^{-c}$, where c is a positive constant dependent only on k . Central to Gowers' proof is the fact that a set which is uniform under the standard definition may contain more than the expected number of arithmetic progressions. The example of

$$\{s \in \mathbb{Z}/N\mathbb{Z} : |s^2| \leq N/10\},$$

a uniform set which contains more than the expected number of four-term progressions, is given in [11] to illustrate the need for a new definition of pseudorandomness. This new definition is based on polynomial uniformity of degree d . Gowers showed that if a set \mathcal{A}_N is not \mathfrak{a} -uniform of degree d with the uniformity parameter \mathfrak{a} appropriately bounded above in terms of the density δ_N , then \mathcal{A}_N is concentrated in some long arithmetic progression \mathcal{P} , and one may instead search for k -term arithmetic progressions in \mathcal{P} .

One should not, however, assume that if a set with no solutions to a given system has zero upper density, then the converse is also true. For example, in 2006, Green and Tao [14] showed that the primes, despite having zero upper density in \mathbb{N} , nevertheless contain arbitrarily long arithmetic progressions.

The systems for which the properties of upper densities of solution-free sets have been studied have consisted almost exclusively of linear equations. Even those which are not purely linear, such as the equation

$$x_1 - x_2 = y^k \quad (k \geq 2)$$

first studied by Furstenberg [8] in the case $k = 2$ and Sárközy [18] in the general case, are dominated by the linear components. Though I have studied some examples of such systems, as I shall discuss later, the main result of my dissertation research addresses for the first time the behaviour of the upper density of solution-free sets associated to a system which is not purely linear and which is dominated by its nonlinear components.

3. Current research. As part of my research as a graduate student and as a postdoctoral associate, I have been concerned with the system

$$(1) \quad \lambda_1 x_1^j + \dots + \lambda_s x_s^j = 0 \quad (1 \leq j \leq k),$$

where $k \geq 2$ is fixed, s is sufficiently large in terms of k , and the coefficients λ_i are nonzero integers satisfying

$$(2) \quad \lambda_1 + \dots + \lambda_s = 0.$$

The condition (2) on the λ_i guarantees that the system (1) is invariant under translation and dilation of the vector (x_1, \dots, x_s) . Moreover, one may easily verify that it is possible to generate solutions to the system (1) by simply setting $x_1 = \dots = x_s$. We refer to such solutions as trivial solutions.

In my dissertation, I established the following upper bound for the decay of δ_N in a set \mathcal{A} which furnishes no non-trivial solutions to the system (1).

Theorem 1. *Suppose that $s \geq s_1$, where s_1 is a constant depending on k , and that \mathcal{A}^s contains no non-trivial solutions (x_1, \dots, x_s) to the system (1). Then for N sufficiently large in terms of the λ_i , there exists a constant $c > 0$ dependent only on k such that $\delta_N \ll (\log \log N)^{-c}$.*

Here \ll and \gg denote the familiar Vinogradov notation. In the case $k = 2$, discussed in [19], I show that when $s \geq 9$, one may take a value for c of 8×10^{-7} . The general case is discussed in [20].

In my proofs of Theorem 1, I combine the Hardy-Littlewood circle method as used by Roth in [17] with the methods developed by Gowers in [11]. I begin by assuming that \mathcal{A}_N is \mathfrak{a} -uniform of degree k and approximating the number of solutions to the system (1) in \mathcal{A}_N^s by the number of solutions I in $[1, N]^s$ multiplied by δ_N^s . Using methods from [11] and [17], I obtain an upper bound for the error $|\delta_N^s \cdot I - \mathcal{N}|$ involved in this approximation. I then apply the classical Hardy-Littlewood method to obtain a lower bound for the number of solutions in $[1, N]^s$. Central to this stage of the proof is the fact that the system (1) is invariant under translation and dilation of the vector (x_1, \dots, x_s) .

The final stage of the proof of Theorem 1 in the uniform case involves assuming that the uniformity parameter \mathfrak{a} is sufficiently small in terms of δ_N as to be dominated by the other terms in the bound on the approximation error $|\delta_N^s \cdot I - \mathcal{N}|$. Under this assumption, I show that for N sufficiently large in terms of the λ_i , the number of solutions \mathcal{N} satisfies the lower bound

$$(3) \quad \mathcal{N} \gg \delta_N^s N^{s-k(k+1)/2},$$

a bound consistent with the expected product of local densities. In the non-uniform case, I remove the assumption that the uniformity parameter \mathfrak{a} is small in terms of δ_N and apply the concentration argument developed in [11] to show that the set must be dense in some long arithmetic progression \mathcal{P} . As this process results in a non-negligible increase in δ_N , one cannot use this method to obtain an analogue of (3) in the non-uniform case.

4. Programme for future research. The field of additive combinatorics has seen rapid development in the wake of Gowers' work on Szemerédi's theorem and the proof of the existence of long arithmetic progressions in the primes by Green and Tao. As such, there are many different possible paths which I plan to explore in future research, paths involving both extensions of my current research and also more ambitious projects in and beyond number theory.

Quantitative refinements. The lower bounds on $s_0(k)$ in the current proof of Theorem 1 are a consequence of the Vinogradov mean value theorem (see, for example, Chapter 5 in [25]), which is used in the applications of the Hardy-Littlewood circle method to obtain a lower bound for the number of solutions to the system (1) in $[1, N]^s$. I anticipate that, by using newer methods when applying the circle method, it should be possible to improve the lower bounds on $s_0(k)$ for small values of k .

For example, in the case $k = 2$, at one stage in the minor arc analysis one must obtain an upper bound for the integral

$$(4) \quad \int_0^1 \int_0^1 \left| \sum_{1 \leq x \leq N} e(\alpha_2 x^2 + \alpha_1 x) \right|^t d\alpha_2 d\alpha_1,$$

where $e(\theta) = \exp(2\pi i\theta)$. The method of Vinogradov may be used to show that the integral (4) is of size $\ll N^{t-3+\epsilon}$ for any $\epsilon > 0$, provided that $t \geq 6$. This leads to a value of $s_0(2) = 9$. However, if we instead estimate the integral (4) using the restriction theory developed by Bourgain in [4] and applied to the problem of finding arithmetic progressions in the primes by Green and Tao in [13], it should be possible to obtain an upper bound for the integral (4) of size $\ll N^{t-3}$ for $t > 6$. This would lead to an improved lower bound of $s_0(2) = 6 + \delta$ for any $\delta > 0$, the theoretical best possible number of variables for which the circle method is effective when applied to the system (1).

Additionally, the upper bound on the decay of δ_N for a set containing no three-term progressions has been improved considerably since Roth's original proof. In particular, Heath-Brown [15] and Szemerédi [23] proved that $\delta_N \ll (\log N)^{-c}$ for an absolute, positive constant c , and Bourgain [5]

has refined this further to $\delta_N \ll (\log \log N / \log N)^{1/2}$. Where the minor arc estimate in the original proof uses Parseval's inequality, the method of Heath-Brown and Szemerédi uses instead a form of the large sieve inequality for a set of “well-spaced” points to produce a sharper minor arc estimate. It should be possible to adapt the approach used in their proofs to the application of the circle method in the proof of Theorem 1, which should lead to a similar refinement for the stated bound.

Additive combinatorics over different fields. The theory of the Hardy-Littlewood circle method in fields other than \mathbb{R} is well established, and analogues of numerous examples of Waring's problem over algebraic number fields and function fields have been studied in recent years. However, while many results from additive combinatorics have been studied over finite fields, such as the work of Green [12] on Szemerédi's theorem in \mathbb{F}_p^n , the possibility that they possess analogues over infinite fields other than \mathbb{R} remains largely unexplored. It would be interesting to see if, for example, the upper density of a subset of the ring of integers \mathcal{O}_K of a number field K containing no solutions to a system of the form in (1) with coefficients in \mathcal{O}_K is zero, and if one may obtain an explicit upper bound for the decay of the density as a function of the norm $|\cdot|$ of elements of \mathcal{O}_K . The proof of such a result would involve applying the Hardy-Littlewood method over the given number field K rather than over \mathbb{Q} as in the proof of Theorem 1. The error estimate would require a well-defined analogue of \mathfrak{a} -uniformity of degree d for sets

$$\mathcal{A}_N \subset \{x \in \mathcal{O}_K : |x| \leq N\},$$

where $|\cdot|$ is the norm over \mathbb{Q} .

Additive equations over difference sets. Sárközy proved in [18] that if a set $\mathcal{A} \subset \mathbb{N}$ has positive upper density, then for any integer $k \geq 2$, the difference set $\mathcal{A} - \mathcal{A}$ necessarily contains k^{th} powers of integers. Bergelson and Leibman [2] generalised this result using ergodic theory to show that, given any polynomials $P_1(x), \dots, P_k(x)$ with integer coefficients and zero constant term, then a set \mathcal{A} of positive upper density necessarily furnishes infinitely many solutions to the system

$$x_0 - x_j = P_j(m) \quad (1 \leq j \leq k),$$

where $x_0, \dots, x_k \in \mathcal{A}$ and $m \in \mathbb{Z}$. More recently, Tao and Ziegler [24] have proven an analogous result for subsets of the primes of positive upper density.

By adapting the methods developed by Balog *et al.* [1] in their refinement of Sárközy's work on powers in difference sets, I anticipate that it should be possible to prove the following two-dimensional version of Sárközy's result. If \mathcal{A} has positive upper density, then the system

$$b_{i1}\alpha_1 + b_{i2}\alpha_2 = m^{k_i} \quad (i = 1, 2),$$

where the $b_{ij} \in \mathbb{Z}$ are fixed and $k_1, k_2 \geq 2$ are fixed integers, necessarily has solutions with $\alpha_1, \alpha_2 \in \mathcal{A} - \mathcal{A}$ and $m \in \mathbb{Z}$. This proof would involve showing that the Fourier sum

$$\sum_{a_1, a_2 \in \mathcal{A}_{N/2}} e(((b_{11}a_1 + b_{12}a_2)t + (b_{21}a_1 + b_{22}a_2)s) / BN),$$

where $B = \max_{i=1,2} b_{i1} + b_{i2}$, takes on large values for many pairs (t, s) . It would be interesting to see if similar multi-dimensional analogues of the results in [2] and [24] also exist.

Ergodic proofs. There may be proofs of the results I have obtained using methods other than the techniques from harmonic analysis I have used in my proofs. For example, the result of Szemerédi that a set containing no arithmetic progressions of a given length k has zero upper density was proven using ergodic theory by Furstenberg in [8], while Kra [16] has given an ergodic theoretic perspective on the result of Green and Tao on long arithmetic progressions in the primes.

It seems probable that Theorem 1 may be proven using ergodic methods, which may in turn lead to further results in both number theory and ergodic theory. Indeed, an ergodic result of Furstenberg and Katznelson [9] may be used to prove that a set furnishing no non-trivial solutions

to the system (1) has zero upper density, though this does not lead to an explicit upper bound for δ_N .

Forms in many variables. The methods of Gowers hold much promise for establishing bounds for the upper density of solution-free sets for other systems of equations. For example, let $\mathbf{x} = (x_1, \dots, x_s)$ and $f_i(\mathbf{x})$, $1 \leq i \leq r$ be a system of homogeneous forms of degree d with rational coefficients. Davenport [6] showed that a single cubic form (the case $r = 1$ and $d = 3$) in $s \geq 32$ variables necessarily possesses a non-trivial solution in \mathbb{Z}^s . Birch [3] generalised this result to show that if the system defined by $f_i(\mathbf{x}) = 0$, $1 \leq i \leq r$ has non-singular real and p -adic solutions for all rational primes p , then the system necessarily possesses a non-trivial solution in \mathbb{Z}^s provided that the number of variables s satisfies

$$s - \dim[V^*] > r(r+1)(d-1)2^{d-1},$$

where V^* is the singular locus of the system defined by the $f_i(\mathbf{x})$.

Davenport and Birch's proofs use modified versions of the Hardy-Littlewood circle method. It seems probable that the techniques of Gowers may be combined with this modified circle method to establish a bound for the upper density of a solution-free subset of \mathbb{Z} for a system of homogenous forms satisfying the hypotheses of Birch's result.

Other Diophantine problems. The Hardy-Littlewood circle method, which I employ in my proof of Theorem 1, has many applications beyond additive combinatorics. For example, the method may be used to count rational points of bounded height on algebraic varieties, or one may use it to count points in lattices, and there are further applications in coding theory and cryptography. Moreover, the Hardy-Littlewood method provides many possible topics for undergraduate research, such as estimates of the $L^p[0, 1]$ norms of certain exponential sums. The questions in these and other applications of the circle method provide ample possibilities for future research.

REFERENCES

- [1] A. Balog, J. Pelikán, J. Pintz, and E. Szemerédi, Difference sets without k^{th} powers, *Acta Math. Hungar.* (2), **65** (1994), 165-187.
- [2] V. Bergelson and A. Leibman, Polynomial extensions of van der Waerden's and Szemerédi's theorems, *J. Amer. Math. Soc.*, **9** (1996), No.3, 725-753.
- [3] B. J. Birch, Forms in many variables, *Proc. Roy. Soc. Ser. A*, **265** (1962), 245-263.
- [4] J. Bourgain, On $\Lambda(p)$ -subsets of squares, *Israel J. Math.*, **67** (1989), 291-311.
- [5] J. Bourgain, On triples in arithmetic progression, *Geom. Funct. Anal.*, **9** (1999), 968-984.
- [6] H. Davenport, Cubic forms in 32 variables, *Phil. Trans. A*, **251** (1959), 193-232.
- [7] P. Erdős and P. Turán, On some sequences of integers, *J. London Math. Soc.*, **11** (1936), 261-264.
- [8] H. Furstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, *J. Analyse Math.* **31** (1977), 204-256.
- [9] H. Furstenberg and Y. Katznelson, An ergodic Szemerédi theorem for commuting transformations, *J. Analyse Math.* **34** (1978), 275-291.
- [10] W. T. Gowers, A new proof of Szemerédi's Theorem for arithmetic progressions of length four, *Geom. Funct. Anal.*, **8** (1998), 529-551.
- [11] W. T. Gowers, A new proof of Szemerédi's Theorem, *Geom. Funct. Anal.*, **11** (2001), 465-588.
- [12] B. J. Green, Finite fields models in additive combinatorics, in *Surveys in combinatorics 2005*, London Math. Soc. Lecture Note Ser., 327, Cambridge University Press, 2005, 1-27.
- [13] B. J. Green and T. C. Tao, Restriction theory of the Selberg sieve, with applications, *J. Théor. Nombres Bordeaux*, **18** (2006), 137-172.
- [14] B. J. Green and T. C. Tao, The primes contain arbitrarily long arithmetic progressions, to appear in *Ann. Math.* (2007).
- [15] D. R. Heath-Brown, Integer sets containing no arithmetic progressions, *J. London Math. Soc.* (2), **35** (1987), 385-394.
- [16] B. Kra, The Green-Tao theorem on arithmetic progressions in the primes: an ergodic point of view, *Bull. Amer. Math. Soc. (N.S.)*, **43** (2006), No.1, 3-23.

- [17] K. F. Roth, On certain sets of integers, *J. London Math. Soc.*, **28** (1953), 104-109.
- [18] A. Sárközy, On difference sets of sequences of integers III, *Acta Math. Acad. Sci. Hungar.* **31** (1978), No.3-4, 355-386.
- [19] M. L. Smith, On solution-free sets for simultaneous quadratic and linear equations, *J. London Math. Soc.* **79** (2009), 273-293.
- [20] M. L. Smith, On solution-free sets for simultaneous additive equations, preprint.
- [21] E. Szemerédi, On sets of integers containing no four elements in arithmetic progression, *Acta Math. Acad. Sci. Hungar.* **20** (1969), 89-104.
- [22] E. Szemerédi, On sets of integers containing no k elements in arithmetic progression, *Acta Arith.* **27** (1975), 299-345.
- [23] E. Szemerédi, Integer sets containing no arithmetic progressions, *Acta Math. Hungar.* **56** (1990), 155-158.
- [24] T. C. Tao and T. Ziegler, The primes contain arbitrarily long polynomial progressions, preprint.
- [25] R. C. Vaughan, *The Hardy-Littlewood method*, 2nd ed., Cambridge Tracts in Math. vol.125, Cambridge University Press, 1997.
- [26] A. J. Wiles, Modular elliptic curves and Fermat's last theorem, *Ann. of Math. (2)* **141** (1995), No.3, 443-551.