Generalizing the Hardy-Littlewood Method for Primes

Ben Green (reporting on joint work with Terry Tao)

Clay Institute/University of Cambridge

August 20, 2006

Using these slides

These are slides from a talk I will give at the ICM in Madrid in August 2006. They contain a few deliberate inaccuracies which I will draw attention to verbally in the talk. For the benefit of anyone wishing to use these slides in the future, an accompanying document is available on my webpage which draws attention to these points.

The Hardy-Littlewood Method, I

Hardy and Littlewood developed their method in the early part of the 20th century, initially to deal with Waring's problem on representing

$$N = x_1^k + \cdots + x_s^k.$$

They also showed that

$$N = p_1 + p_2 + p_3$$

for odd N, though an unproved assumption in the direction of GRH was required.

The Hardy-Littlewood Method, II

Vinogradov (1937) removed any dependence on GRH and simplified the proof. Van der Corput and Chowla used the same method to show that

$$p_1 + p_3 = 2p_2$$

has infinitely many nontrivial solutions (that is, the primes contain infinitely many 3-term arithmetic progressions).

If one had to summarise the Hardy-Littlewood method in a short sentence, one would say that it is "a method of harmonic analysis". In this context Fourier transforms are often referred to as *exponential sums*. Much of this talk will be about "going beyond harmonic analysis".

The Hardy-Littlewood Method – Asymptotics

The method often gives, when it applies, an asymptotic for the number of solutions. In the case of the primes, such asymptotics are most conveniently stated using the von Mangoldt function:

$$\Lambda(n) := \begin{cases} \log p & \text{if } n = p^k,\ p \text{ prime;} \\ 0 & \text{otherwise.} \end{cases}$$

Example (3-term APs: Chowla/Van der Corput)

$$\sum_{n_1,n_2 \leqslant N} \Lambda(n_1)\Lambda(n_1+n_2)\Lambda(n_1+2n_2) = \mathfrak{S}N^2 + o(N^2)$$

as $N \to \infty$, where $\mathfrak{S} = 2\prod_{p \geqslant 3}\left(1 - \frac{1}{(p-1)^2}\right) \approx 1.32032.$

The Hardy-Littlewood Method – Asymptotics

It is not hard to recover unweighted counts from estimates involving Λ by using the prime number theorem.

Example

3-term progressions, unweighted version

$$\#\{n_1, n_2 \leqslant N : n_1, n_1 + n_2, n_1 + 2n_2 \text{ are all prime}\}$$
$$= \mathfrak{S}\frac{N^2}{(\log N)^3} + o\left(\frac{N^2}{(\log N)^3}\right)$$

as $N \to \infty$, where $\mathfrak{S} \approx 1.32032$ is the same constant as before.

Systems of linear forms, I

How can we generalize the result of Chowla and Van der Corput, which gave an asymptotic for

$$\sum_{n_1,n_2 \leqslant N} \Lambda(n_1)\Lambda(n_1+n_2)\Lambda(n_1+2n_2)?$$

Let us be ambitious, and ask for an asymptotic in which (n_1, n_1+n_2, n_1+2n_2) is replaced by Ψ, a general t-tuple of linear forms.

Asymptotic for linear forms in primes?

Is it true that

$$\sum_{\vec{n} \in K} \Lambda(\psi_1(\vec{n}))\Lambda(\psi_2(\vec{n}))\ldots\Lambda(\psi_t(\vec{n})) = \mathfrak{S}N^d + o_\Psi(N^d)$$

as $N \to \infty$, for some easily describable $\mathfrak{S}$?

Systems of linear forms, II

On the previous slide, the notation is as follows:

- $K \subseteq [-N, N]^d$ is a convex body;
- Each ψ_i is an affine linear form with integer coefficients, i.e. a map from $\mathbb{Z}^d$ to $\mathbb{Z}$ having the form

$$\psi_i(\vec{n}) = L_{i1} n_1 + \cdots + L_{id} n_d + l_i.$$

We use the letter Ψ to denote a system $(\psi_i)_{i=1}^t$ of linear forms like this.

In fact there is a conjecture, basically due to Dickson, which predicts that there *is* such an asymptotic and gives a formula for $\mathfrak{S}$.

Dickson's conjecture, I

Suppose that $\Psi = (\psi_i)_{i=1}^t$ is a system of linear forms, and let $K \subseteq [-N, N]^d$ be a convex body. For any integer q, we define the local von Mangoldt function $\Lambda_{\mathbb{Z}/q\mathbb{Z}} : \mathbb{Z}/q\mathbb{Z} \to \mathbb{R}$ by

$$\Lambda_{\mathbb{Z}/q\mathbb{Z}}(n) := \begin{cases} q/\phi(q) & \text{if } n \in (\mathbb{Z}/q\mathbb{Z})^* \\ 0 & \text{otherwise} \end{cases}$$

Definition (Local factors)

Let $q \geqslant 1$ be an integer. Define the local factor β_q by

$$\beta_q := \mathbb{E}_{\vec{n} \in (\mathbb{Z}/q\mathbb{Z})^d} \Lambda_{\mathbb{Z}/q\mathbb{Z}}(\psi_1(\vec{n})) \dots \Lambda_{\mathbb{Z}/q\mathbb{Z}}(\psi_t(\vec{n})).$$

Define also

$$\beta_\infty := \mathrm{vol}_d(K \cap \Psi^{-1}((\mathbb{R}^+)^d)).$$

Dickson's conjecture, II

As on the last slide, let $\Psi = (\psi_i)_{i=1}^d$ be a system of linear forms, and let $K \subseteq [-N, N]^d$ be a convex body.

Conjecture (Dickson's conjecture)

We have

$$\sum_{\vec{n} \in K} \Lambda(\psi_1(\vec{n}))\Lambda(\psi_2(\vec{n})) \dots \Lambda(\psi_t(\vec{n})) = \mathfrak{S} N^d + o_\Psi(N^d)$$

as $N \to \infty$, *where* $\mathfrak{S} = \beta_\infty \prod_p \beta_p$.

A refinement is possible allowing for the possibility that the constant terms of the ψ_i grow with N (cf. Vinogradov's 3-primes theorem).

Dickson's conjecture, III: examples

Example (Progressions of length 3)

Take $d = 2$, $\Psi = (n_1, n_1 + n_2, n_1 + 2n_2)$ and $K = [0, N]^2$.

$$\mathfrak{S} = 2 \prod_{p \geqslant 3} \left(1 - \frac{1}{(p-1)^2}\right) \approx 1.32032.$$

Complexity $= 1$

Example (Progressions of length 4)

Take $d = 2$, $\Psi = (n_1, n_1 + n_2, n_1 + 2n_2, n_1 + 3n_2)$ and $K = [0, N]^2$.

$$\mathfrak{S} = \frac{9}{2} \prod_{p \geqslant 5} \left(1 - \frac{3p-1}{(p-1)^3}\right) \approx 2.85825.$$

Complexity $= 2$

Example (Twin primes)

Take $d = 1$, $\Psi = (n_1, n_1 + 2)$ and $K = [0, N]$.

$$\mathfrak{S} = 2 \prod_{p \geqslant 3} \left(1 - \frac{1}{(p-1)^2}\right) \approx 1.32032.$$

Complexity $= \infty$

Complexity of a linear system

Definition (Complexity)

Let $\Psi = (\psi_1, \dots, \psi_t)$ be a system of affine-linear forms. If $1 \leqslant i \leqslant t$ and $s \geqslant 0$, we say that Ψ has *i-complexity at most s* if one can cover the $t-1$ forms $\{\psi_j : j \in [t]\backslash\{i\}\}$ by $s+1$ classes, such that ψ_i does not lie in the affine-linear span of any of these classes. The *complexity* of the Ψ is defined to be the least s for which the system has i-complexity at most s for all $1 \leqslant i \leqslant t$, or ∞ if no such s exists.

Example (Progressions of length 4)

Take $\Psi = (n_1, n_1 + n_2, n_1 + 2n_2, n_1 + 3n_2)$. $\psi_1 = n_1$ does not lie in the affine-linear span of any individual form ψ_2, ψ_3, ψ_4, but it does lie in the span of any *two* of these forms.

Our results and goals

We hope to be able to prove Dickson's conjecture for systems of complexity $s < \infty$. A system of linear forms only has infinite complexity if some two of the forms are affine multiples of one another (e.g. $n_1, n_1 + 2$).

Theorem (G.–Tao 2006)

Let $s \geqslant 1$ be an integer. Assume two conjectures, the Gowers Inverse conjecture GI(s) *and the Möbius Nilsequences conjecture* MN(s). *Then Dickson's conjecture holds for all linear systems of complexity s.*

Theorem (H.-L. (1920s) + Vinogradov (1937) + ε)

The conjectures GI(1) *and* MN(1) *are true.*

Theorem (G.–Tao 2006)

The conjectures GI(2) *and* MN(2) *are true.*

Structure and randomness

Fix a system $\Psi = (\psi_i)_{i=1}^t$ of linear forms with complexity s and a convex body $K \subseteq [-N, N]^d$. If $f_1, \dots, f_t : \{1, \dots, N\} \to \mathbb{R}$ are functions, define

$$T(f_1, \dots, f_t) := \sum_{\vec{n} \in K \cap \mathbb{Z}^d} f_1(\psi_1(\vec{n})) \dots f_t(\psi_t(\vec{n})).$$

We are interested in $T(\Lambda, \dots, \Lambda)$. The key idea is to decompose

$$\Lambda = \Lambda^\sharp + \Lambda^\flat = \text{structured} + \text{pseudorandom}$$

Then we may expand $T(\Lambda, \dots, \Lambda)$ as a sum of 2^t terms. The term $T(\Lambda^\sharp, \dots, \Lambda^\sharp)$ gives the main term $\mathfrak{S}N^d$ in Dickson's conjecture. The other $2^t - 1$ terms, each of which involves at least one $\Lambda^\flat$, are "error" terms and we show they are $o(N^d)$.

A decomposition

The actual decomposition

$$\Lambda = \Lambda^\sharp + \Lambda^\flat$$

is a fairly standard one in analytic number theory.

Recall that

$$\Lambda(n) = -\sum_{d|n} \mu(d) \log d,$$

where μ is the Möbius function

$$\mu(n) = \begin{cases} (-1)^k & \text{if } n = p_1 \dots p_k \text{ is squarefree} \\ 0 & \text{otherwise.} \end{cases}$$

We "morally" define

$$\Lambda^\sharp := -\sum_{d|n, d<N^\theta} \mu(d) \log d, \qquad \Lambda^\flat := -\sum_{d|n, d\geqslant N^\theta} \mu(d) \log d,$$

for some small $\theta = \theta(\Psi)$ (in real life there is some smoothing).

Gowers norms, I

What do we mean by structured and pseudorandom?

Let $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{R}$ be a function. Define

$$\|f\|_{U^k} := \left(\mathbb{E}_{n \in \mathbb{Z}/N\mathbb{Z}, \vec{h} \in (\mathbb{Z}/N\mathbb{Z})^k} \prod_{\vec{\omega} \in \{0,1\}^k} f(n + \vec{\omega} \cdot \vec{h})\right)^{1/2^k}.$$

In fact an extension to $\mathbb{C}$-valued functions is possible by insertion of appropriate bars. For example

$$\|f\|_{U^2} := \left(\mathbb{E}_{n,h_1,h_2} f(n)\overline{f(n+h_1)f(n+h_2)}f(n+h_1+h_2)\right)^{1/4}.$$

This is a kind of sum of f over parallelograms.

Gowers norms, II

The U^3 norm expands explicitly as

$$\begin{aligned}\|f\|_{U^3} := \big(\mathbb{E}_{n,h_1,h_2,h_3} f(n)\overline{f(n+h_1)f(n+h_2)f(n+h_3)}\times \\ \times f(n+h_1+h_2)f(n+h_1+h_3)f(n+h_2+h_3)\times \\ \times \overline{f(n+h_1+h_2+h_3)}\big)^{1/8}.\end{aligned}$$

This is a kind of sum of f over 3-dimensional parallelepipeds.

Generalised von Neumann Theorems

Suppose we are thinking about a system $\Psi = (\psi_i)_{i=1}^t$ of complexity s. Recall the associated average

$$T(f_1, \dots, f_t) := \mathbb{E}_{\vec{n} \in (\mathbb{Z}/N\mathbb{Z})^d} f_1(\psi_1(\vec{n})) \dots f_t(\psi_t(\vec{n})).$$

The $(s+1)$st Gowers norm $\| \cdot \|_{U^{s+1}}$ controls such averages.

Theorem (Generalised von Neumann theorem)

For "reasonably general" functions $f_1, \dots, f_t : \mathbb{Z}/N\mathbb{Z} \to \mathbb{R}$ and for any i we have the estimate $|T(f_1, \dots, f_i, \dots, f_t)| \ll \|f_i\|_{U^{s+1}}$.

Functions bounded by 1 are "reasonably general". So (essentially) is Λ.

The Inverse Question for the Gowers Norms

Key point

If we are studying a system $\Psi = (\psi_i)_{i=1}^t$ of complexity s, then a function $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ should be thought of as *pseudorandom* if $\|f\|_{U^{s+1}}$ is small.

The Inverse Question for the Gowers Norms

Let $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ be a function with $\|f\|_\infty \leqslant 1$, and let $\delta > 0$. Suppose that

$$\|f\|_{U^{s+1}} \geqslant \delta.$$

What can we say about f?

In fact we need to ask the same question for "reasonable" functions f which are not bounded by 1 (such as $f = \Lambda^\flat$), but let us start with the bounded case.

Inverse theorem for U^2

Using Fourier analysis (or "exponential sums" in the usual parlance of the Hardy-Littlewood method) one may prove the following.

Theorem (Inverse theorem for U^2)

Suppose that $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ is a function with $\|f\|_\infty \leqslant 1$. Suppose that $\|f\|_{U^2} \geqslant \delta$. Then there is some $\theta \in \mathbb{R}/\mathbb{Z}$ such that

$$|\mathbb{E}_{n \leqslant N} f(n) e(\theta n)| \geqslant \delta^2.$$

Recall that $e(\alpha) := e^{2\pi i \alpha}$.

We won't give the (easy) proof, but the key fact is the existence of the *formula* $\|f\|_{U^2} = \|\widehat{f}\|_4$ involving the discrete Fourier transform on $\mathbb{Z}/N\mathbb{Z}$.

Inverse theorem for U^3, I

An inverse theorem for the U^3 norm cannot be so simple.

Example (Quadratic phases)

Write $f(n) := e(n^2\sqrt{2})$. Then $\|f\|_{U^3} = 1$. However, for every $\theta \in \mathbb{R}/\mathbb{Z}$,

$$|\mathbb{E}_{n \leqslant N} f(n) e(\theta n)| = o(1).$$

Indeed, writing $\phi(n) = n^2\sqrt{2}$,

$$\begin{aligned}\|f\|_{U^3}^8 &= \mathbb{E}_{n,h_1,h_2,h_3} e(\phi(n) - \phi(n+h_1) - \phi(n+h_2) - \phi(n+h_3) \\ &\qquad + \phi(n+h_1+h_2) + \phi(n+h_1+h_3) + \phi(n+h_2+h_3) \\ &\qquad - \phi(n+h_1+h_2+h_3)) \\ &= \mathbb{E}_{n,h_1,h_2,h_3} e(\phi'''(h_1,h_2,h_3)) = 1.\end{aligned}$$

Inverse theorem for U^3, II

In fact an inverse theorem for the U^3-norm must look rather exotic.

Example (Generalised quadratic phases)

Write $f(n) = e(\{n\sqrt{2}\}\{n\sqrt{3}\})$, for $n = 1, \dots, N$. Then $\|f\|_{U^3} \gg 1$, but f does not correlate with any genuinely linear or quadratic phase function.

Theorem (G.–Tao, 2005)

Suppose that $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ has $\|f\|_\infty \leqslant 1$, and that $\|f\|_{U^3} \geqslant \delta$. Then there exists a generalised quadratic polynomial

$$\phi(n) = \sum_{r,s \leqslant C(\delta)} \beta_{rs}\{\theta_r n\}\{\theta_s n\} + \sum_{r \leqslant C(\delta)} \gamma_r \{\theta_r n\},$$

where $\beta_{rs}, \gamma_r, \theta_r \in \mathbb{R}$, such that

$$|\mathbb{E}_{n \leqslant N} f(n) e(\phi(n))| \gg_\delta 1.$$

Inverse theorem for the U^3-norm, III

The last theorem wasn't very pretty. It turns out that generalised quadratics like $e(\{n\sqrt{2}\}\{n\sqrt{3}\})$ may be interpreted in a natural way in terms of 2-*step nilsequences*.

Definition (Nilsequences)

Let G be a connected, simply-connected Lie Group which is s-step nilpotent. Thus if we write $G_0 = G_1 = G$ and $G_{i+1} = [G, G_i]$ for $i \geqslant 1$ then we have $G_{s+1} = \{1\}$. Let $\Gamma \subseteq G$ be a discrete, cocompact subgroup. The quotient G/Γ is called an s-step nilmanifold. The group G acts on G/Γ by left multiplication. For any $x \in G/\Gamma$ and $g \in G$ we may consider the orbit $(g^n \cdot x)_{n\in\mathbb{N}}$ of x under multiplication by g. If $F : G/\Gamma \to \mathbb{C}$ is a bounded, Lipschitz function then we call the sequence $(F(g^n \cdot x))_{n\in\mathbb{N}}$ an s-step nilsequence.

Inverse theorem for the U^3-norm, IV

As we remarked, generalised quadratics such as $e(\{n\sqrt{2}\}\{n\sqrt{3}\})$ can be interpreted in terms of 2-step nilsequences, in particular 2-step nilsequences arising from the Heisenberg group $G = \begin{pmatrix} 1 & \mathbb{R} & \mathbb{R} \\ 0 & 1 & \mathbb{R} \\ 0 & 0 & 1 \end{pmatrix}$.

In this way it is possible to reformulate our inverse theorem for the U^3-norm in terms of nilmanifolds. We omit the details.

Theorem (The GI(2) conjecture, G.–Tao, 2005)

Let $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ be a function with $\|f\|_\infty \leqslant 1$. Suppose that $\|f\|_{U^3} \geqslant \delta$. Then there is a 2-step nilsequence $(F(g^n \cdot x))_{n\in\mathbb{N}}$ on some 2-step nilmanifold G/Γ such that

$$|\mathbb{E}_{n\in\mathbb{Z}/N\mathbb{Z}} f(n)F(g^n \cdot x)| \gg 1.$$

Everything in sight – the dimension of G/Γ, the Lipschitz constant of F, and the implied constant in the $\gg$ notation – depends only on δ.

The Gowers Inverse Conjectures GI(s)

Given the last slide, it is not hard to guess the formulation.

Conjecture (Gowers Inverse conjecture GI(s))

Let $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ be a function with $\|f\|_\infty \leqslant 1$. Suppose that $\|f\|_{U^{s+1}} \geqslant \delta$. Then there is an s-step nilsequence $(F(g^n \cdot x))_{n\in\mathbb{N}}$ on some s-step nilmanifold G/Γ such that

$$|\mathbb{E}_{n\in\mathbb{Z}/N\mathbb{Z}} f(n)F(g^n \cdot x)| \gg 1.$$

Everything in sight – the dimension of G/Γ, the Lipschitz constant of F, and the implied constant in the $\gg$ notation – depends only on δ.

Bootstrapping the GI(s) conjecture

The GI(s) conjecture does not say anything useful about the function $f = \Lambda$, which is not bounded.

Theorem (Bootstrapped GI(s), G.–Tao 2006)

The Gowers Inverse conjecture GI(s) *implies a stronger version of itself, in which the function* $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{R}$ *need not be bounded by* 1. *Instead, it need only be bounded by a "pseudorandom measure".*

We will not define the term "pseudorandom measure" here. The existence of pseudorandom measures bounding Λ, or rather a somewhat modified version of Λ, was one of the key ingredients in our paper *The primes contain arbitrarily long arithmetic progressions*.

Gowers norms – summary

Summary

To study linear systems of complexity s, for example in the primes, the right "harmonics" to use are the s-step nilsequences. These may be brought into play via the Gowers norms $\|\cdot\|_{U^{s+1}}$.

Recall.....

We talked about a decomposition

$$\begin{aligned}\Lambda &= \Lambda^\sharp + \Lambda^\flat \\ &= \text{structured} + \text{pseudorandom}\end{aligned}$$

where we "morally" defined

$$\Lambda^\sharp := -\sum_{d|n, d<N^\theta} \mu(d)\log d, \qquad \Lambda^\flat := -\sum_{d|n, d\geqslant N^\theta} \mu(d)\log d,$$

for some small $\theta = \theta(\Psi)$.

We now know, in terms of the Gowers norms, what an appropriate notion of "pseudorandom" is. Our task, then, is to establish that

$$\|\Lambda^\flat\|_{U^{s+1}} \text{ is small.}$$

Möbius and Nilsequences, I

To show that

$$\|\Lambda^\flat\|_{U^{s+1}} \text{ is small,}$$

it suffices (assuming GI(s), and hence bootstrapped GI(s)) to show that

$$|\mathbb{E}_{n\leqslant N}\Lambda^\flat(n)F(g^n\cdot x)| \text{ is small}$$

for every fixed s-step nilsequence $(F(g^n\cdot x))_{n\in\mathbb{N}}$.

Substituting in the definition of $\Lambda^\flat$ and rearranging, one may reduce this to showing that

$$|\mathbb{E}_{n\leqslant N}\mu(n)F(g^n\cdot x)| \text{ is really rather small}$$

for every fixed s-step nilsequence $(F(g^n\cdot x))_{n\in\mathbb{N}}$.

Möbius and Nilsequences, II

Conjecture (Möbius-Nilsequences conjecture, MN(s))

Fix an s-step nilmanifold G/Γ and a bounded Lipschitz function $F : G/\Gamma \to \mathbb{C}$. Then we have the estimate

$$|\mathbb{E}_{n \leqslant N} \mu(n) F(g^n \cdot x)| \ll_A \log^{-A} N$$

as $N \to \infty$, for any $A > 0$.

This accords well with the well-known "Möbius randomness heuristic":

Möbius randomness heuristic

Let $F : \mathbb{N} \to \mathbb{R}$ be any bounded "low complexity" function. Then we expect

$$|\mathbb{E}_{n \leqslant N} \mu(n) F(n)| \text{ to be small.}$$

Future plans

- Prove the GI(s) and MN(s) conjectures.
  This is work in progress. We are close to the "finite field model" version of GI(s), certainly for $s = 3$. The techniques we used for MN(2) ought to extend to MN(3), MN(4), ...
- Quantitative issues; error terms.
  Relevant to this would be good bounds in Freiman's theorem, in particular the "Polynomial Freiman-Ruzsa conjecture".
- A more conceptual way of discovering nilsequences?
  No serious ideas in this direction at present.
- Look at non-linear equations in the primes.
  $p_1 p_2 - p_3 p_4 = 2$ would be extremely interesting!