PRIMES IN INTERVALS OF BOUNDED LENGTH

ANDREW GRANVILLE

ABSTRACT. In April 2013, Yitang Zhang proved the existence of a finite bound B such that there are infinitely many pairs of distinct primes which differ by no more than B. This is a massive breakthrough, makes the twin prime conjecture look highly plausible (which can be re-interpreted as the conjecture that one can take $B = 2$) and his work helps us to better understand other delicate questions about prime numbers that had previously seemed intractable. The original purpose of this talk was to discuss Zhang's extraordinary work, putting it in its context in analytic number theory, and to sketch a proof of his theorem.

Zhang had even proved the result with $B = 70\,000\,000$. Moreover, a co-operative team, *polymath8*, collaborating only on-line, had been able to lower the value of B to 4680. Not only had they been more careful in several difficult arguments in Zhang's original paper, they had also developed Zhang's techniques to be both more powerful and to allow a much simpler proof. Indeed the proof of Zhang's Theorem, that will be given in the write-up of this talk, is based on these developments.

In November, inspired by Zhang's extraordinary breakthrough, James Maynard dramatically slashed this bound to 600, by a substantially easier method. Both Maynard, and Terry Tao who had independently developed the same idea, were able to extend their proofs to show that for any given integer $m \geqslant 1$ there exists a bound B_m such that there are infinitely many intervals of length B_m containing at least m distinct primes. We will also prove this much stronger result herein, even showing that one can take $B_m = e^{8m+5}$.

If Zhang's method is combined with the Maynard-Tao set up then it appears that the bound can be further reduced to 576. If all of these techniques could be pushed to their limit then we would obtain $B(= B_2)= 12$, so new ideas are still needed to have a feasible plan for proving the twin prime conjecture.

The article will be split into two parts. The first half, which appears here, we will introduce the work of Zhang, Polymath8, Maynard and Tao, and explain their arguments that allow them to prove their spectacular results. As we will discuss, Zhang's main novel contribution is an estimate for primes in relatively short arithmetic progressions. The second half of this article sketches a proof of this result; the Bulletin article will contain full details of this extraordinary work.

Part 1. Primes in short intervals

sec:intro

1. Introduction

1.1. **Intriguing questions about primes.**

Early on in our mathematical education we get used to the two basic rules of arithmetic, addition and multiplication. When we

1991 *Mathematics Subject Classification.* 11P32.

To Yitang Zhang, for showing that one can, no matter what.

define a prime number, simply in terms of the number's multiplicative properties, we discover a sequence of numbers, which is easily defined, yet difficult to gain a firm grasp of, perhaps since the primes are defined in terms of what they are not (i.e. that they *cannot* be factored into two smaller integers)).

When one writes down the sequence of prime numbers:

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, \ldots$$

one sees that they occur frequently, but it took a rather clever construction of the ancient Greeks to even establish that there really are infinitely many. Looking further at a list of primes, some patterns begin to emerge; for example, one sees that they often come in pairs:

$$3 \text{ and } 5;\ 5 \text{ and } 7;\ 11 \text{ and } 13;\ 17 \text{ and } 19;\ 29 \text{ and } 31;\ 41 \text{ and } 43;\ 59 \text{ and } 61, \ldots$$

One might guess that there are infinitely many such prime pairs. But this is an open, elusive question, the *twin prime conjecture*. Until recently there was little theoretical evidence for it. All that one could say was that there was an enormous amount of computational evidence that these pairs never quit; and that this conjecture (and various more refined versions) fit into an enormous network of *conjecture*, which build a beautiful elegant structure of all sorts of prime patterns. If the twin prime conjecture were false then the whole edifice would crumble.

The twin prime conjecture is certainly intriguing to both amateur and professional mathematicians alike, though one might argue that it is an artificial question, since it asks for a very delicate additive property of a sequence defined by its multiplicative properties. Indeed, number theorists had struggled, until very recently, to identify an approach to this question that seemed likely to make any significant headway. In this article we will discuss these latest shocking developments. In the first few sections we will take a leisurely stroll through the historical and mathematical background, so as to give the reader a sense of the great theorems that have been recently proved, from a perspective that will prepare the reader for the details of the proof.

1.2. **Other patterns.** Looking at the list of primes above we see other patterns that begin to emerge, for example, one can find four primes which have all the same digits, except the last one:

$$11, 13, 17 \text{ and } 19;\ \text{which is repeated with } 101, 103, 107 \text{ and } 109;$$

and one can find many more such examples – are there infinitely many? More simply how about prime pairs with difference 4:

$$3 \text{ and } 7;\ 7 \text{ and } 11;\ 13 \text{ and } 17;\ 19 \text{ and } 23;\ 37 \text{ and } 41;\ 43 \text{ and } 47;\ 67 \text{ and } 71, \ldots;$$

or difference 10:

$$3 \text{ and } 13;\ 7 \text{ and } 17;\ 13 \text{ and } 23;\ 19 \text{ and } 29;\ 31 \text{ and } 41;\ 37 \text{ and } 47;\ 43 \text{ and } 53, \ldots?$$

Are there infinitely many such pairs? Such questions were probably asked back to antiquity, but the first clear mention of twin primes in the literature appears in a paper of de Polignac from 1849. In his honour we now call any integer h, for which there are infinitely many prime pairs $p, p+h$, a *de Polignac number*.[1]

Then there are the *Sophie Germain pairs*, primes p and $q := 2p+1$, which prove useful in several simple algebraic constructions:[2]

$$2 \text{ and } 5;\ 3 \text{ and } 7;\ 5 \text{ and } 11;\ 11 \text{ and } 23;\ 23 \text{ and } 47;\ 29 \text{ and } 59;\ 41 \text{ and } 83; \ldots;$$

Now we have spotted all sorts of patterns, we need to ask ourselves whether there is a way of predicting which patterns can occur and which do not. Let's start by looking at the possible differences between primes: It is obvious that there are not infinitely many prime pairs of difference 1, because one of any two consecutive integers must be even, and hence can only be prime if it equals 2. Thus there is just the one pair, 2 and 3, of primes with difference 1. One can make a similar argument for prime pairs with odd difference. Hence if h is an integer for which there are infinitely many prime pairs of the form $p,\ q = p + h$ then h must be even. We have seen many examples, above, for each of $h = 2,\ h = 4$ and $h = 10$, and the reader can similarly construct lists of examples for $h = 6$ and for $h = 8$, and indeed for any other even h that takes her or his fancy. This leads us to bet on the *generalized twin prime conjecture*, which states that for any even integer $2k$ there are infinitely many prime pairs $p,\ q = p + 2k$.

What about prime triples? or quadruples? We saw two examples of prime quadruples of the form $10n+1,\ 10n+3,\ 10n+7,\ 10n+9$, and believe that there are infinitely many. What about other patterns? Evidently any pattern that includes an odd difference cannot succeed. Are there any other obstructions? The simplest pattern that avoids an odd difference is $n, n+2, n+4$. One finds the one example $3,\ 5,\ 7$ of such a prime triple, but no others. Further examination makes it clear why not: One of the three numbers is always divisible by 3. This is very similar to what happened with $n, n+1$; and one can verify that, similarly, one of $n, n+6, n+12, n+18, n+24$ is always divisible by 5. The general obstruction can be described as follows:

For a given set of distinct integers $a_1 < a_2 < \ldots < a_k$ we say that prime p is an *obstruction* if p divides at least one of $n + a_1, \ldots, n + a_k$, for every integer n. In other words, p divides

$$\mathcal{P}(n) = (n + a_1)(n + a_2) \ldots (n + a_k)$$

for every integer n; which can be classified by the condition that the set $a_1, a_2, \ldots, a_k \pmod{p}$ includes all of the residue classes mod p. If no prime is an obstruction then we say that $x + a_1, \ldots, x + a_k$ is an *admissible* set of forms.[3].

[1]Some authors make a slightly different definition: That p and $p + h$ should also be consecutive primes.

[2]The group of reduced residues mod q is a cyclic group of order $q-1 = 2p$, and therefore isomorphic to $C_2 \times C_p$ if $p > 2$. Hence the order of each element in the group is either 1 (that is, $1 \pmod{q}$), 2 (that is, $-1 \pmod{q}$), p (the squares mod q) or $2p = q - 1$. Hence $g \pmod{q}$ generates the group of reduced residues if and only if g is not a square mod q and $g \not\equiv -1 \pmod{q}$.

[3]Notice that $a_1, a_2, \ldots, a_k \pmod{p}$ can occupy no more than k residue classes mod p and so, if $p > k$ then p cannot be an obstruction. Hence, to check whether a given set A of k integers is admissible, one needs only find one residue class $b_p \pmod{p}$, for each prime $p \leqslant k$, which does not contain any element of A.

In 1904 Dickson made the optimistic conjecture that if there is no such "obvious" obstruction to a set of linear forms being infinitely often prime, then they are infinitely often simultaneously prime. That is:

Conjecture: *If $x+a_1,\dots,x+a_k$ is an admissible set of forms then there are infinitely many integers n such that $n+a_1,\dots,n+a_k$ are all prime numbers.*

In this case, we call $n+a_1,\dots,n+a_k$ a *k-tuple* of prime numbers.

To date, this has not been proven for any $k>1$ though, following Zhang's work, we begin to get close for $k=2$. Indeed, Zhang has proved a weak variant of this conjecture for $k=2$, as we shall see. Moreover Maynard [29], and Tao [39], have gone on to prove a weak variant *for any $k \geqslant 2$.*

The above conjecture can be extended, as is, to all sets of k linear forms with integer coefficients in one variable (for example, the triple $n, 2n+1,\ 3n-2$), so long as we extend the notion of admissibility to also exclude the possible obstruction that two of the linear forms have different signs for all but finitely many n, (since, for example, n and $2-n$, can never be simultaneously prime); some people call this the "obstruction at the 'prime', -1". We can also extend the conjecture to more than one variable (for example the set of forms $m, m+n, m+4n$):

The prime k-tuplets conjecture: *If a set of k linear forms in n variables is admissible then there are infinitely many sets of n integers such that when we substitute these integers into the forms we get a k-tuple of prime numbers.*

There has been substantial recent progress on this conjecture. The famous breakthrough was Green and Tao's theorem [19] for the k-tuple of linear forms in the two variables a and d:

$$a,\ a+d,\ a+2d,\dots,\ a+(k-1)d$$

(in other words, there are infinitely many k-term arithmetic progressions of primes.) Along with Ziegler, they went on to prove the prime k-tuplets conjecture for *any* admissible set of linear forms, provided no two satisfy a linear equation over the integers, [20]. What a remarkable theorem! Unfortunately these exceptions include many of the questions we are most interested in; for example, $p,\ q=p+2$ satisfy the linear equation $q-p=2$; and $p,\ q=2p+1$ satisfy the linear equation $q-2p=1$).

Finally, we also believe that the conjecture holds if we consider any admissible set of k irreducible polynomials with integer coefficients, with any number of variables. For example we believe that n^2+1 is infinitely often prime, and that there are infinitely many prime triples $m,\ n,\ m^2-2n^2$.

1.3. **The new results; primes in bounded intervals.** In this section we state Zhang's main theorem, as well as the improvement of Maynard and Tao, and discuss a few of the more beguiling consequences:

Zhang's main theorem: *There exists an integer k such that the following is true: If $x + a_1, \ldots, x + a_k$ is an admissible set of forms then there are infinitely many integers n such that* at least two of $n + a_1, \ldots, n + a_k$ *are prime numbers.*

Note that the result states that only two of the $n + a_i$ are prime, not all (as would be required in the prime k-tuplets conjecture). Zhang proved this result for a fairly large value of k, that is $k = 3500000$, which has been reduced to $k = 105$ by Maynard. Of course if one could take $k = 2$ then we would have the twin prime conjecture,[4] but the most optimistic plan at the moment, along the lines of Zhang's proof, would yield $k = 5$.

To deduce that there are bounded gaps between primes from Zhang's Theorem we need only show the existence of an admissible set with k elements. This is not difficult, simply by letting the a_i be the first k primes $> k$.[5] Hence we have proved:

Corollary 1.1 (Bounded gaps between primes)**.** *There exists a bound B such that there are infinitely many integers pairs of prime numbers $p < q < p + B$.*

Finding the smallest B for a given k is a challenging question. The prime number theorem together with our construction above suggests that $B \leqslant k(\log k + C)$ for some constant C, but it is interesting to get better bounds. For Maynard's $k = 105$, Engelsma showed that one can take $B = 600$,[6] and that this is best possible.

Our Corollary further implies

Corollary 1.2. *There is an integer $h, 0 < h \leqslant B$ such that there are infinitely many pairs of primes $p, p + h$.*

That is, some positive integer $\leqslant B$ is a de Polignac number. In fact one can go a little further using Zhang's main theorem, and deduce that if A is *any* admissible set of k integers then there is an integer $h \in (A - A)^+ := \{a - b : \ a > b \in A\}$ such that there are infinitely many pairs of primes $p, p + h$. One can find many beautiful consequences of this; for example, that a positive proportion of even integers are de Polignac numbers.

Next we state the Theorem of Maynard and of Tao:

The Maynard-Tao theorem: *For any given integer $m \geqslant 2$, there exists an integer k such that the following is true: If $x + a_1, \ldots, x + a_k$ is an admissible set of forms then*

---

[4]And the generalized twin prime conjecture, and that there are infinitely many Sophie Germain pairs, and ...

[5]This is admissible since none of the a_i is $0 \pmod p$ for any $p \leqslant k$, and the $p > k$ were handled in the previous footnote.

[6]Sutherland's website `http://math.mit.edu/~primegaps/` gives Engelsma's admissible 105-tuple: 0, 10, 12, 24, 28, 30, 34, 42, 48, 52, 54, 64, 70, 72, 78, 82, 90, 94, 100, 112, 114, 118, 120, 124, 132, 138, 148, 154, 168, 174, 178, 180, 184, 190, 192, 202, 204, 208, 220, 222, 232, 234, 250, 252, 258, 262, 264, 268, 280, 288, 294, 300, 310, 322, 324, 328, 330, 334, 342, 352, 358, 360, 364, 372, 378, 384, 390, 394, 400, 402, 408, 412, 418, 420, 430, 432, 442, 444, 450, 454, 462, 468, 472, 478, 484, 490, 492, 498, 504, 510, 528, 532, 534, 538, 544, 558, 562, 570, 574, 580, 582, 588, 594, 598, 600.

there are infinitely many integers n such that at least m of $n+a_1,\dots,n+a_k$ *are prime numbers.*

This includes and extends Zhang's Theorem (which is the case $k=2$). The proof even allows one make this explicit (we will obtain $k \leqslant e^{8m+4}$, and Maynard improves this to $k \leqslant cm^2e^{4m}$ for some constant $c>0$).

Corollary 1.3 (Bounded intervals with m primes)**.** *For any given integer $m \geqslant 2$, there exists a bound B_m such that there are infinitely many intervals $[x, x+B_m]$ (with $x \in \mathbb{Z}$) which contain m prime numbers.*

We will prove that one can take $B_m = e^{8m+5}$ (and Maynard improves this to $B_m = cm^3e^{4m}$ for some constant $c>0$).

A *Dickson k-tuple* is a set of integers $a_1 < \dots < a_k$ such that there are infinitely many integers for which $n+a_1, n+a_2, \dots, n+a_k$ are each prime

Corollary 1.4. *A positive proportion of m-tuples of integers are Dickson m-tuples.*

Proof. With the notation as in the Maynard-Tao theorem let $R = \prod_{p\leqslant k} p$, select x to be a large integer multiple of R and let $\mathcal{N} := \{n \leqslant x : \ (n,R)=1\}$ so that $|\mathcal{N}| = \frac{\phi(R)}{R}x$. Any subset of k elements of $\mathcal{N}$ is admissible, since it does not contain any integer $\equiv 0 \pmod p$ for each prime $p \leqslant k$. There are $\binom{|\mathcal{N}|}{k}$ such k-tuples. Each contains a Dickson m-tuple by the Maynard-Tao theorem.

Now suppose that are $T(x)$ Dickson m-tuples with $1 \leqslant a_1 < \dots < a_m \leqslant x$. Any such m-tuple is a subset of exactly $\binom{|\mathcal{N}|-m}{k-m}$ of the k-subsets of $\mathcal{N}$, and hence

$$T(x) \cdot \binom{|\mathcal{N}|-m}{k-m} \geqslant \binom{|\mathcal{N}|}{k},$$

and therefore $T(x) \geqslant (|\mathcal{N}|/k)^m = (\frac{\phi(R)}{R}/k)^m \cdot x^m$ as desired. □

This proof yields that, as a proportion of the m-tuples in $\mathcal{N}$,

$$T(x)/\binom{|\mathcal{N}|}{m} \geqslant 1/\binom{k}{m}.$$

The $m=2$ case implies that at least $\frac{1}{5460}$th of the even integers are de Polignac numbers.

Zhang's Theorem and the Maynard-Tao theorem each hold for any admissible k-tuple of linear forms (not just those of the form $x+a$). With this we can prove several other amusing consequences:

• The last Corollary holds if we insist that the primes in the Dickson k-tuples are consecutive primes.

• There are infinitely many m-tuples of consecutive primes such that each pair in the m-tuple differ from one another by just two digits when written in base 10.

• For any $m \geqslant 2$ and coprime integers a and q, there are infinitely many intervals $[x, x + qB_m]$ (with $x \in \mathbb{Z}$) which contain exactly m prime numbers, each $\equiv a \pmod q$.[7]

• Let $d_n = p_{n+1} - p_n$ where p_n is the nth smallest prime. Fix $m \geqslant 1$. There are infinitely many n for which $d_n < d_{n+1} < \ldots < d_{n+m}$. There are also infinitely many n for which $d_n > d_{n+1} > \ldots > d_{n+m}$. This was a favourite problem of Paul Erdős, though we do not see how to deduce such a result for other orderings of the d_n.

1.4. **Bounding the gaps between primes. A brief history.** The young Gauss, examining Chernac's table of primes up to one million, surmised that "the density of primes at around x is roughly $1/\log x$". This was subsequently verified, as a consequence of the *prime number theorem*. Therefore we are guaranteed that there are infinitely many pairs of primes $p < q$ with $q - p \leqslant (1 + \epsilon) \log p$ for any fixed $\epsilon > 0$, which is not quite as small a gap as we are hoping for! Nonetheless this raises the question: Fix $c > 0$. Can we even prove that

There are infinitely many pairs of primes $p < q$ with $q < p + c \log p$?

This follows for all $c > 1$ by the prime number theorem, but it is not easy to prove such a result for any particular value of $c \leqslant 1$. The first such results were proved conditionally assuming the Generalized Riemann Hypothesis. This is surprising since the Generalized Riemann Hypothesis was formulated to better understand the distribution of primes in arithmetic progressions, so why would it appear in an argument about short gaps between primes? It is far from obvious by the argument used, and yet this connection deepened and broadened as the literature developed. We will discuss primes in arithmetic progressions in detail in the next section.

The first unconditional (though inexplicit) such result, bounding gaps between primes, was proved by Erdős in 1940 using the small sieve. In 1966, Bombieri and Davenport [2] substituted the Bombieri-Vinogradov theorem for the Generalized Riemann Hypothesis in earlier, conditional arguments, to prove this unconditionally for any $c \geqslant \frac{1}{2}$. The Bombieri-Vinogradov Theorem is also a result about primes in arithmetic progressions (as we will discuss later). In 1988 Maier [28] observed that one can easily modify this to obtain any $c \geqslant \frac{1}{2} e^{-\gamma}$; and he further improved this, by combining the approaches of Erdős and of Bombieri and Davenport, to obtain some bound a little smaller than $\frac{1}{4}$, in a technical *tour-de-force*.

The first big breakthrough occurred in 2005 when Goldston, Pintz and Yildirim [15] were able to show that there are infinitely many pairs of primes $p < q$ with $q < p + c \log p$, for *any* given $c > 0$. Indeed they extended their methods to show that, for any $\epsilon > 0$, there are infinitely many pairs of primes $p < q$ for which

$$q - p < (\log p)^{1/2+\epsilon}.$$

It is their method which forms the basis of the discussion in this paper. Like Bombieri and Davenport, they showed that one can better understand small gaps between primes

[7]Thanks to Tristan Freiberg for pointing this out to me.

by obtaining strong estimates on primes in arithmetic progressions, as in the Bombieri-Vinogradov Theorem. Even more, if one assumes a strong, but widely believed, conjecture about the equi-distribution of primes in arithmetic progressions, which extends the Bombieri-Vinogradov Theorem, then one can show that there are infinitely many pairs of primes $p < q$ which differ by no more than 12 (that is, $p < q \leqslant p + 12$)! In fact one can take $k = 5$ in Zhang's theorem, and then apply the result to the admissible 5-tuple, $\{0,\ 2,\ 6,\ 8,\ 12\}$ What an extraordinary statement! We know that if $p < q \leqslant p + 12$ then $q - p = 2,\ 4,\ 6,\ 8,\ 10$ or 12, and so at least one of these difference occurs infinitely often. That is, there exists a positive, even integer $2k \leqslant 12$ such that there are infinitely pairs of primes $p,\ p + 2k$. It would be good to refine this further.

After Goldston, Pintz and Yildirim, most of the experts tried and failed to obtain enough of an improvement of the Bombieri-Vinogradov Theorem to deduce the existence of some finite bound B such that there are infinitely many pairs of primes that differ by no more than B. To improve the Bombieri-Vinogradov Theorem is no mean feat and people have longed discussed "barriers" to obtaining such improvements. In fact a technique had been developed by Fouvry [10], and by Bombieri, Friedlander and Iwaniec [3], but this was neither powerful enough nor general enough to work in this circumstance.

Enter Yitang Zhang, an unlikely figure to go so much further than the experts, and to find exactly the right improvement and refinement of the Bombieri-Vinogradov Theorem to establish the existence of the elusive bound B such that there are infinitely many pairs of primes that differ by no more than B. By all accounts, Zhang was a brilliant student in Beijing from 1978 to the mid-80s, finishing with a master's degree, and then working on the Jacobian conjecture for his Ph.D. at Purdue, graduating in 1992. He did not proceed to a job in academia, working in odd jobs, such as in a sandwich shop, at a motel and as a delivery worker. Finally in 1999 he got a job at the University of New Hampshire as a lecturer (though with the same teaching load as tenure-track faculty). From time-to-time a lecturer devotes their energy to working on proving great results, but few have done so with such aplomb as Zhang. Not only did he prove a great result, but he did so by improving *technically* on the experts, having important key ideas that they missed and developing a highly ingenious and elegant construction concerning exponential sums. Then, so as not to be rejected out of hand, he wrote his difficult paper up in such a clear manner that it could not be denied. Albert Einstein worked in a patent office, Yitang Zhang in a Subway sandwich shop; both found time, despite the unrelated calls on their time and energy, to think the deepest thoughts in science. Moreover Zhang's breakthrough came at the relatively advanced age of 50 (or more). Truly *extraordinary.*

After Zhang, a group of researchers decided to team up online to push the techniques, created by Zhang, to their limit. This was the eighth incarnation of the *polymath* project, which is an experiment to see whether this sort of collaboration can help research develop beyond the traditional boundaries set by our academic culture. The original bound of $70,000,000$ was quickly reduced, and seemingly every few weeks, different parts of Zhang's argument could be improved, so that the bound came down in to the thousands. Moreover the polymath8 researchers found variants on Zhang's argument about the distribution of primes in arithmetic progressions, that allow one to avoid some of the deeper ideas that Zhang used. These modifications enabled your author to give an accessible complete proof in this article.

After these clarifications of Zhang's work, two researchers asked themselves whether the original "set-up" of Goldston, Pintz and Yildirim could be modified to get better results. James Maynard obtained his Ph.D. this summer at Oxford, writing one of the finest theses in sieve theory of recent years. His thesis work equipped him perfectly to question whether the basic structure of the proof could be improved. Unbeknownst to Maynard, at much the same time (late October), one of the world's greatest living mathematicians, Terry Tao, asked himself the same question. Both found, to their surprise, that a relatively minor variant made an enormous difference, and that it was suddenly much easier to prove Zhang's Main Theorem and to go far beyond, because one can avoid having to prove any difficult new results about primes in arithmetic progressions. Moreover it is now not difficult to prove results about m primes in a bounded interval, rather than just two.

2. The distribution of primes, divisors and prime k-tuplets

2.1. **The prime number theorem.** As we mentioned in the previous section, Gauss observed, at the age of 16, that "the density of primes at around x is roughly $1/\log x$", which leads quite naturally to the conjecture that

$$\#\{\text{primes } p \leqslant x\} \approx \int_2^x \frac{dt}{\log t} \sim \frac{x}{\log x} \quad \text{as } x \to \infty.$$

(We use the symbol $A(x) \sim B(x)$ for two functions A and B of x, to mean that $A(x)/B(x) \to 1$ as $x \to \infty$.) This was proved in 1896, the *prime number theorem*, and the integral provides a considerably more precise approximation to the number of primes $\leqslant x$, than $x/\log x$. However, this integral is rather cumbersome to work with, and so it is natural to instead weight each prime with $\log p$; that is we work with

$$\Theta(x) := \sum_{\substack{p \text{ prime} \\ p \leqslant x}} \log p$$

and the prime number theorem is equivalent to

$$\Theta(x) \sim x \quad \text{as } x \to \infty. \tag{2.1}$$ pnt2

2.2. **The prime number theorem for arithmetic progressions, I.** Any prime divisor of (a, q) is an obstruction to the primality of values of the polynomial $qx + a$, and these are the only such obstructions. The prime k-tuplets conjecture therefore implies that if $(a, q) = 1$ then there are infinitely many primes of the form $qn + a$. This was first proved by Dirichlet in 1837. Once proved, one might ask for a more quantitative result. If we look at the primes in the arithmetic progressions (mod 10):

$$\begin{aligned}&11,\ 31,\ 41,\ 61,\ 71,\ 101\\&3,\ 13,\ 23,\ 43,\ 53,\ 73,\ 83,\ 103\\&7,\ 17,\ 37,\ 47,\ 67,\ 97,\ 107\\&19,\ 29,\ 59,\ 79,\ 89,\ 109\end{aligned}$$

then there seem to be roughly equal numbers in each, and this pattern persists as we look further out. Let $\phi(q)$ denote the number of $a \pmod q$ for which $(a,q)=1$, and so we expect that

$$\Theta(x;q,a) := \sum_{\substack{p \text{ prime} \\ p\leqslant x \\ p\equiv a \pmod q}} \log p \sim \frac{x}{\phi(q)} \quad \text{as } x\to\infty.$$

This is the *prime number theorem for arithmetic progressions* and was first proved by suitably modifying the proof of the prime number theorem.

The function $\phi(q)$ was studied by Euler, who showed that it is *multiplicative*, that is

$$\phi(q) = \prod_{p^e\|q} \phi(p^e)$$

(where $p^e\|q$ means that p^e is the highest power of prime p dividing q) and that $\phi(p^e) = p^e - p^{e-1}$ for all $e\geqslant 1$.

2.3. **The prime number theorem and the Möbius function.** Multiplicative functions lie at the heart of much of the theory of the distribution of prime numbers. One, in particular, the Möbius function, $\mu(n)$, plays a prominent role. It is defined as $\mu(p)=-1$ for every prime p, and $\mu(p^m)=0$ for every prime p and exponent $m\geqslant 2$; the value at any given integer n is then deduced from the values at the prime powers, by multiplicativity: If n is squarefree then $\mu(n)$ equals 1 or -1 depending on whether n has an even or odd number of prime factors, respectively. One might guess that there are roughly equal numbers of each, which one can phrase as the conjecture that

$$\frac{1}{x}\sum_{n\leqslant x}\mu(n) \to 0 \quad \text{as} \quad n\to\infty.$$

This is a little more difficult to prove than it looks; indeed it is also equivalent to (2.1). That equivalence is proved using the remarkable identity

$$\sum_{ab=n}\mu(a)\log b \ = \begin{cases}\log p & \text{if } n=p^m, \text{ where } p \text{ is prime}, m\geqslant 1;\\ 0 & \text{otherwise.}\end{cases} \tag{2.2}$$

For more on this connection see the forthcoming book [18].

2.4. **A quantitative prime k-tuplets conjecture.** We are going to develop a heuristic to guesstimate the number of pairs of twin primes $p, p+2$ up to x. We start with Gauss's statement that "the density of primes at around x is roughly $1/\log x$. Hence the probability that p is prime is $1/\log x$, and the probability that $p+2$ is prime is

$1/\log x$ so, assuming that these events are independent, the probability that p and $p+2$ are simultaneously prime is

$$\frac{1}{\log x} \cdot \frac{1}{\log x} = \frac{1}{(\log x)^2};$$

and so we might expect about $x/(\log x)^2$ pairs of twin primes $p, p+2 \leqslant x$. However there is a problem with this reasoning, since we are implicitly assuming that the events "p is prime for an arbitrary integer $p \leqslant x$", and "$p+2$ is prime for an arbitrary integer $p \leqslant x$", can be considered to be independent. This is obviously false since, for example, if p is even then $p+2$ must also be.[8] So, to correct for the non-independence modulo small primes q, we determine the ratio of the probability that both p and $p+2$ are not divisible by q, to the probabiliity that p and p' are not divisible by q.

Now the probability that q divides an arbitrary integer p is $1/q$; and hence the probability that p is not divisible by q is $1 - 1/q$. Therefore the probability that both of two independently chosen integers are not divisible by q, is $(1-1/q)^2$.

The probability that q does not divide either p or $p+2$, equals the probability that $p \not\equiv 0$ or $-2 \pmod q$. If $q > 2$ then p can be in any one of $q-2$ residue classes mod q, which occurs, for a randomly chosen $p \pmod q$, with probability $1 - 2/q$. If $q = 2$ then p can be in any just one residue class mod 2, which occurs with probability 1/2. Hence the "correction factor" for divisibility by 2 is

$$\frac{(1-\frac{1}{2})}{(1-\frac{1}{2})^2} = 2,$$

whereas the "correction factor" for divisibility by any prime $q > 2$ is

$$\frac{(1-\frac{2}{q})}{(1-\frac{1}{q})^2}.$$

Now divisibility by different small primes is independent, as we vary over values of n, by the Chinese Remainder Theorem, and so we might expect to multiply together all of these correction factors, corresponding to each "small" prime q. The question then becomes, what does "small" mean? In fact, it doesn't matter much because the product of the correction factors over larger primes is very close to 1, and hence we can simply extend the correction to be a product over all primes q. (More precisely, the infinite product over all q, converges.) Hence we define the *twin prime constant* to be

$$C := 2 \prod_{\substack{q \text{ prime} \\ q \geqslant 3}} \frac{(1-\frac{2}{q})}{(1-\frac{1}{q})^2} \approx 1.3203236316,$$

and we conjecture that the number of prime pairs $p, p+2 \leqslant x$ is

$$\sim C \frac{x}{(\log x)^2}.$$

[8] Also note that the same reasoning would tell us that there are $\sim x/(\log x)^2$ prime pairs $p, p+1 \leqslant x$.

Computational evidence suggests that this is a pretty good guess. The analogous argument implies the conjecture that the number of prime pairs $p, p+2k \leqslant x$ is

$$\sim C \prod_{\substack{p|k \\ p\geqslant 3}} \left(\frac{p-1}{p-2}\right) \frac{x}{(\log x)^2}.$$

This argument is easily modified to make an analogous prediction for any k-tuple: Given $a_1, \dots, a_k$, let $\Omega(p)$ be the set of distinct residues given by $a_1, \dots, a_k \pmod p$, and then let $\omega(p) = |\Omega(p)|$. None of the $n + a_i$ is divisible by p if and only if n is in any one of $p - \omega(p)$ residue classes mod p, and therefore the correction factor for prime p is

$$\frac{(1-\frac{\omega(p)}{p})}{(1-\frac{1}{p})^k}.$$

Hence we predict that the number of prime k-tuplets $n+a_1, \dots, n+a_k \leqslant x$ is,

$$\sim C(a)\frac{x}{(\log x)^k} \quad \text{where} \quad C(a) := \prod_p \frac{(1-\frac{\omega(p)}{p})}{(1-\frac{1}{p})^k}.$$

An analogous conjecture, via similar reasoning, can be made for the frequency of prime k-tuplets of polynomial values in several variables. What is remarkable is that computational evidence suggests that these conjectures do approach the truth, though this rests on the rather shaky theoretical framework given here. A more convincing theoretical framework based on the *circle method* (so rather more difficult) was given by Hardy and Littlewood [21], which we will discuss in the extended (Bulletin) article.

3. Uniformity in arithmetic progressions

3.1. **When primes are first equi-distributed in arithmetic progressions.** There is an important further issue when considering primes in arithmetic progressions: In many applications it is important to know when we are first guaranteed that the primes are more-or-less equi-distributed amongst the arithmetic progressions $a \pmod q$ with $(a, q) = 1$; that is

$$\Theta(x; q, a) \sim \frac{x}{\phi(q)} \text{ for all } (a, q) = 1. \tag{3.1}$$

To be clear, here we want this to hold when x is a function of q, as $q \to \infty$.

Extensive calculations give evidence that, for any $\epsilon > 0$, if q is sufficiently large and $x \geqslant q^{1+\epsilon}$ then the primes up to x are equi-distributed amongst the arithmetic progressions $a \pmod q$ with $(a, q) = 1$, that is (3.1) holds. This is not only unproved at the moment, also no one really has a plausible plan of how to show such a result. However the slightly weaker statement that (3.1) holds for any $x \geqslant q^{2+\epsilon}$, can be shown to be true, assuming the Generalized Riemann Hypothesis. This gives us a clear plan for proving such a result, but one which has seen little progress in the last century!

The best unconditional results known are much weaker than we have hoped for, equidistribution only being proved once $x \geqslant e^{q^\epsilon}$. This is the *Siegel-Walfisz Theorem*, and it can be stated in several (equivalent) ways with an error term: For any $B > 0$ we have

$$\Theta(x;q,a) = \frac{x}{\phi(q)} + O\left(\frac{x}{(\log x)^B}\right) \text{ for all } (a,q)=1. \tag{3.2}$$ SW1

Or: for any $A > 0$ there exists $B > 0$ such that if $q < (\log x)^A$ then

$$\Theta(x;q,a) = \frac{x}{\phi(q)}\left\{1 + O\left(\frac{1}{(\log x)^B}\right)\right\} \text{ for all } (a,q)=1. \tag{3.3}$$ SW2

That x needs to be so large compared to q limited the number of applications of this result.

The great breakthough of the second-half of the twentieth century came in appreciating that for many applications, it is not so important that we know that equidistribution holds for *every* a with $(a,q)=1$, and *every* q up to some Q, but rather that this holds for *most* such q (with $Q = x^{1/2-\epsilon}$). It takes some juggling of variables to state the Bombieri-Vinogradov Theorem: We are interested, for each modulus q, in the size of the largest error term

$$\max_{\substack{a \bmod q\\(a,q)=1}} \left|\Theta(x;q,a) - \frac{x}{\phi(q)}\right|,$$

or even

$$\max_{y\leqslant x} \max_{\substack{a \bmod q\\(a,q)=1}} \left|\Theta(y;q,a) - \frac{y}{\phi(q)}\right|.$$

The bounds $0 \leqslant \Theta(x;q,a) \ll \frac{x}{q}\log x$ are trivial, the upper bound obtained by bounding the possible contribution from each term of the arithmetic progression. (Throughout, the symbol "$\ll$", as in "$f(x) \ll g(x)$" means "there exists a constant $c > 0$ such that $f(x) \leqslant cg(x)$.") We would like to improve on the "trivial" upper bound, perhaps by a power of $\log x$, but we are unable to do so for all q. However, what we can prove is that *exceptional q are few and far between*, and the Bombieri-Vinogradov Theorem expresses this in a useful form. The first thing we do is add up the above quantities over all $q \leqslant Q < x$. The "trivial" upper bound is then

$$\ll \sum_{q\leqslant Q} \frac{x}{q}\log x \ll x(\log x)^2.$$

The Bombieri-Vinogradov states that we can beat this trivial bound by an arbitrary power of $\log x$, provided Q is a little smaller than $\sqrt{x}$:

The Bombieri-Vinogradov Theorem. *For any given $A > 0$ there exists a constant $B = B(A)$, such that*

$$\sum_{q\leqslant Q} \max_{\substack{a \bmod q\\(a,q)=1}} \left|\Theta(x;q,a) - \frac{x}{\phi(q)}\right| \ll_A \frac{x}{(\log x)^A}$$

where $Q = x^{1/2}/(\log x)^B$.

In fact one can take $B = 2A+5$; and one can also replace the summand here by the expression above with the maximum over y (though we will not need to do this here).

3.2. **Breaking the $x^{1/2}$-barrier.** It is believed that estimates like that in the Bombieri-Vinogradov Theorem hold with Q significantly larger than $\sqrt{x}$; indeed Elliott and Halberstam conjectured [8] that one can take $Q = x^c$ for any constant $c < 1$:

The Elliott-Halberstam conjecture *For any given $A > 0$ and η, $0 < \eta < \frac{1}{2}$, we have*

$$\sum_{q \leqslant Q} \max_{\substack{a \bmod q \\ (a,q)=1}} \left| \Theta(x; q, a) - \frac{x}{\phi(q)} \right| \ll \frac{x}{(\log x)^A}$$

where $Q = x^{1/2+\eta}$.

However, it was shown in [13] that one *cannot* go so far as to take $Q = x/(\log x)^B$.

This conjecture was the starting point for the work of Goldston, Pintz and Yıldırım [15], that was used by Zhang [43] (which we give in detail in the next section). It can be applied to obtain the following result, which we will prove.

Theorem 3.1 (Goldston-Pintz-Yıldırım)**.** [15] *Let $k \geqslant 2$, $l \geqslant 1$ be integers, and $0 < \eta < 1/2$, such that*

$$1 + 2\eta > \left(1 + \frac{1}{2l+1}\right)\left(1 + \frac{2l+1}{k}\right). \tag{3.4}$$

Assume that the Elliott-Halberstam conjecture holds with $Q = x^{1/2+\eta}$. If $x+a_1, \dots, x+a_k$ is an admissible set of forms then there are infinitely many integers n such that at least two of $n + a_1, \dots, n + a_k$ are prime numbers.

The conclusion here is exactly the statement of Zhang's main theorem.

If the Elliott-Halberstam conjecture conjecture holds for some $\eta > 0$ then select l to be an integer so large that $\left(1 + \frac{1}{2l+1}\right) < \sqrt{1+2\eta}$. Theorem 3.1 then implies Zhang's theorem for $k = (2l+1)^2$.

The Elliott-Halberstam conjecture seems to be too difficult to prove for now, but progress has been made when restricting to one particular residue class: Fix integer $a \neq 0$. We believe that for any fixed η, $0 < \eta < \frac{1}{2}$, one has

$$\sum_{\substack{q \leqslant Q \\ (q,a)=1}} \left| \Theta(x; q, a) - \frac{x}{\phi(q)} \right| \ll \frac{x}{(\log x)^A}$$

where $Q = x^{1/2+\eta}$, which follows from the Elliott-Halberstam conjecture (but is weaker).

The key to progress has been to notice that if one can "factor" the key terms here then the extra flexibility allows one to make headway. For example by factoring the modulus q as, say, dr where d and r are roughly some pre-specified sizes. The simplest class of integers q for which this can be done is the *y-smooth integers*, those integers whose prime factors are all $\leqslant y$. For example if we are given a y-smooth integer q and we want $q = dr$ with d not much smaller than D, then we select d to be the largest divisor of q

that is $\leqslant D$ and we see that $D/y < d \leqslant D$. This is precisely the class of moduli that Zhang considered.

The other "factorization" concerns the sum $\Theta(x; q, a)$. The terms of this sum can be written as a sum of products, as we saw in (2.2); in fact we will decompose this further, partitioning the values of a and b into different ranges. This will be discussed in full detail in the accompanying article.

Theorem 3.2 (Yitang Zhang's Theorem)**.** *There exist constants $\eta, \delta > 0$ such that for any given integer a, we have*

$$\sum_{\substack{q \leqslant Q \\ (q,a)=1 \\ q \text{ is } y-\text{smooth} \\ q \text{ squarefree}}} \left| \Theta(x; q, a) - \frac{x}{\phi(q)} \right| \ll_A \frac{x}{(\log x)^A} \tag{3.5}$$

where $Q = x^{1/2+\eta}$ and $y = x^{\delta}$.

Zhang [43] proved his Theorem for $\eta/2 = \delta = \frac{1}{1168}$, and his argument works provided $414\eta + 172\delta < 1$. We will prove this result, by a somewhat simpler proof, provided $162\eta + 90\delta < 1$, and the more sophisticated proof of [34] gives (3.5) provided $43\eta + 27\delta < 1$. We expect that this estimate holds for every $\eta \in [0, 1/2)$ and every $\delta \in (0, 1]$, but just proving it for any positive pair $\eta, \delta > 0$ is an extraordinary breakthrough that has an enormous effect on number theory, since it is such an applicable result (and technique). This is the technical result that truly lies at the heart of Zhang's result about bounded gaps between primes, and sketching a proof of this is the focus of the second half of the complete paper (we will give a brief sketch at the end of this article).

4. Goldston-Pintz-Yildirim's argument

We now give a version of the combinatorial argument of Goldston-Pintz-Yıldırım [15], which lies at the heart of the proof that there are bounded gaps between primes. (Henceforth we will call it "the GPY argument".)

4.1. The set up.

Let $\mathcal{H} = (a_1 < a_2 < \ldots < a_k)$ be an admissible k-tuple, and take $x > a_k$. Our goal is to select a weight for which weight$(n) \geqslant 0$ for all n, such that

$$\sum_{x < n \leqslant 2x} \text{weight}(n) \left(\sum_{i=1}^{k} \theta(n + a_i) - \log 3x \right) > 0, \tag{4.1}$$

where $\theta(m) = \log m$ if $m = p$ is prime, and $\theta(m) = 0$ otherwise. If we can do this then there must exist an integer n such that

$$\text{weight}(n) \left(\sum_{i=1}^{k} \theta(n + a_i) - \log 3x \right) > 0.$$

In that case weight$(n) \neq 0$ so that weight$(n) > 0$, and therefore

$$\sum_{i=1}^{k} \theta(n+a_i) > \log 3x.$$

However each $n + a_i \leqslant 2x + a_k < 2x + x$ and so each $\theta(n + a_i) < \log 3x$. This implies that at least two of the $\theta(n + a_i)$ are non-zero, that is, at least two of $n + a_1, \ldots, n + a_k$ are prime.

A simple idea, but the difficulty comes in selecting the function weight(n) with these properties in such a way that we can evaluate the sum. Moreover in gpy [15] they also require that weight(n) is sensitive to when each $n + a_i$ is "almost prime". All of these properties can be acquired by using a construction championed by Selberg. In order that weight$(n) \geqslant 0$ one can simply take it to be a square. Hence we select

$$\text{weight}(n) := \Bigg(\sum_{\substack{d|\mathcal{P}(n)\\ d\leqslant R}} \lambda(d)\Bigg)^2,$$

where the sum is over the positive integers d that divide $\mathcal{P}(n)$, and

$$\lambda(d) := \mu(d) G\left(\frac{\log d}{\log R}\right),$$

where $G(.)$ is a measurable, bounded function, supported only on $[0, 1]$.[9], and μ is the Möbius function. Therefore $\lambda(d)$ is supported only on squarefree, positive integers, that are $\leqslant R$.

We can select $G(t) = (1 - t)^m/m!$ to obtain the results of this section but it will pay, for our understanding of the Maynard-Tao construction, if we prove the GPY result for more general $G(.)$.

4.2. **Evaluating the sums over n.** Now, expanding the above sum gives

$$\sum_{\substack{d_1,d_2\leqslant R\\ D:=[d_1,d_2]}} \lambda(d_1)\lambda(d_2) \Bigg(\sum_{i=1}^{k} \sum_{\substack{x<n\leqslant 2x\\ D|\mathcal{P}(n)}} \theta(n+a_i) - \log 3x \sum_{\substack{x<n\leqslant 2x\\ D|\mathcal{P}(n)}} 1\Bigg). \tag{4.2}$$ gpy2

Let $\Omega(D)$ be the set of congruence classes $m \pmod D$ for which $D|P(m)$; and let $\Omega_i(D)$ be the set of congruence classes $m \in \Omega(D)$ with $(D, m + a_i) = 1$. Hence the parentheses in the above line equals

$$\sum_{i=1}^{k} \sum_{m\in\Omega_i(D)} \sum_{\substack{x<n\leqslant 2x\\ n\equiv m \pmod D}} \theta(n+a_i) - \log 3x \sum_{m\in\Omega(D)} \sum_{\substack{x<n\leqslant 2x\\ n\equiv m \pmod D}} 1. \tag{4.3}$$ gpy3

Our first goal is to evaluate the sums over n. The final sum is easy; there are $x/D+O(1)$ integers in a given arithmetic progression with difference D, in an interval of length x.

[9]By *supported only on* we mean "can be non-zero only on".

The error term here is much smaller than the main term, and is easily shown to be irrelevant to the subsequent calculations.

Counting the number of primes in a given arithmetic progression with difference D, in an interval of length x. is much more difficult. We expect that (3.1) holds, so that each

$$\Theta(2x; D, m + a_i) - \Theta(x; D, m + a_i) \sim \frac{x}{\phi(D)}.$$

Here the error terms are larger and more care is needed. They can be handled by standard techniques, provided that the error terms are smaller than the main terms by an arbitrarily large power of $\log x$, at least on average. This shows why the Bombieri-Vinogradov Theorem is so useful, since it implies the needed estimate provided $R < x^{1/4-o(1)}$ so that each $D < x^{1/2-o(1)}$. Going any further is difficult, so that the $\frac{1}{4}$ is an important barrier. Goldston, Pintz and Yıldırım showed that if one can go just beyond $\frac{1}{4}$ then one can prove that there are bounded gaps between primes, but there did not seem to be any techniques available to them to do so.

For the purposes of the next part of this discussion let us not worry about the range in which such an estimate holds, nor about the size of the accumulated error terms, but rather make the substitution and see where it leads. First, though, we need to better understand the sets $\Omega(D)$ and $\Omega_i(D)$. Since they may be constructed using the Chinese Remainder Theorem from the sets with D prime, therefore if $\omega(D) := |\Omega(D)|$ then $\omega(.)$ is a multiplicative function. Moreover each $|\Omega_i(p)| = \omega(p) - 1$, which we denote by $\omega^*(p)$, and each $|\Omega_i(D)| = \omega^*(D)$, extending ω^* to be a multiplicative function. Putting this altogether we obtain in (4.3) a main term of

$$k\omega^*(D)\frac{x}{\phi(D)} - (\log 3x)\omega(D)\frac{x}{D} = x\left(k\frac{\omega^*(D)}{\phi(D)} - (\log 3x)\frac{\omega(D)}{D}\right).$$

This is typically negative which explains why we cannot simply take our weights, $\lambda(d)$, to all be positive. Substituting this in to (4.2) we obtain, in total, the sums

$$x\left(k\sum_{\substack{d_1,d_2\leqslant R\\ D:=[d_1,d_2]}}\lambda(d_1)\lambda(d_2)\frac{\omega^*(D)}{\phi(D)} - (\log 3x)\sum_{\substack{d_1,d_2\leqslant R\\ D:=[d_1,d_2]}}\lambda(d_1)\lambda(d_2)\frac{\omega(D)}{D}\right). \tag{4.4}$$

The two sums over d_1 and d_2 in (4.4) are not easy to evaluate: The use of the Möbius function leads to many terms being positive, and many negative, so that there is, in fact, a lot of cancelation. There are two techniques in analytic number theory that allow one to get accurate estimates for such sums, when there is a lot of cancelation, one more analytic ([15]), the other more combinatorial ([38], [16]). We will discuss them both, but only fully develop the latter.

4.3. **Evaluating the sums using Perron's formula.** Perron's formula allows one to study inequalities using complex analysis:

$$\frac{1}{2i\pi}\int_{\mathrm{Re}(s)=2}\frac{y^s}{s}\,ds=\begin{cases}1 & \text{if } y>1;\\ 1/2 & \text{if } y=1;\\ 0 & \text{if } 0<y<1.\end{cases}$$

(Here the subscript "$\mathrm{Re}(s)=2$" means that we integrate along the line $s:\ \mathrm{Re}(s)=2$; that is $s=2+it$, with $-\infty<t<\infty$.) So to determine whether $d<R$ we simply compute this integral with $y=R/d$. (The special case, $d=R$, has a negligible effect on our sums, and can be avoided by selecting $R\notin\mathbb{Z}$). Hence the second sum in (4.4) equals

$$\sum_{\substack{d_1,d_2\geqslant 1\\ D:=[d_1,d_2]}}\lambda(d_1)\lambda(d_2)\frac{\omega(D)}{D}\cdot\frac{1}{2i\pi}\int_{\mathrm{Re}(s_1)=2}\frac{(R/d_1)^{s_1}}{s_1}\,ds_1\cdot\frac{1}{2i\pi}\int_{\mathrm{Re}(s_2)=2}\frac{(R/d_2)^{s_2}}{s_2}\,ds_2.$$

Re-organizing this we obtain

$$\frac{1}{(2i\pi)^2}\int_{\substack{\mathrm{Re}(s_1)=2\\ \mathrm{Re}(s_2)=2}}\left(\sum_{\substack{d_1,d_2\geqslant 1\\ D:=[d_1,d_2]}}\frac{\lambda(d_1)\lambda(d_2)}{d_1^{s_1}d_2^{s_2}}\frac{\omega(D)}{D}\right)R^{s_1+s_2}\frac{ds_2}{s_2}\cdot\frac{ds_1}{s_1}\tag{4.5}$$

We will compute the sum in the middle in the special case that $\lambda(d)=\mu(d)$, the more general case following from a variant of this argument. Hence we have

$$\sum_{d_1,d_2\geqslant 1}\frac{\mu(d_1)\mu(d_2)}{d_1^{s_1}d_2^{s_2}}\frac{\omega([d_1,d_2])}{[d_1,d_2]}.\tag{4.6}$$

The summand is a multiplicative function, which means that we can evaluate it prime-by-prime. For any given prime p, the summand is 0 if p^2 divides d_1 or d_2 (since then $\mu(d_1)=0$ or $\mu(d_2)=0$). Therefore we have only four cases to consider: $p\nmid d_1,d_2$; $p|d_1,p\nmid d_2$; $p\nmid d_1,p|d_2$; $p|d_1,p|d_2$, so the pth factor is

$$1-\frac{1}{p^{s_1}}\cdot\frac{\omega(p)}{p}-\frac{1}{p^{s_2}}\cdot\frac{\omega(p)}{p}+\frac{1}{p^{s_1+s_2}}\cdot\frac{\omega(p)}{p}.$$

We have seen that $\omega(p)=k$ for all sufficiently large p so, in that case, the above becomes

$$1-\frac{k}{p^{1+s_1}}-\frac{k}{p^{1+s_2}}+\frac{k}{p^{1+s_1+s_2}}.$$

In the analytic approach, we compare the integrand to a (carefully selected) power of the Riemann-zeta function, $\zeta(s)$. The pth factor of $\zeta(s)$ is $\left(1-\frac{1}{p^s}\right)^{-1}$ so, as a first approximation, the last line is roughly

$$\left(1-\frac{1}{p^{1+s_1+s_2}}\right)^{-k}\left(1-\frac{1}{p^{1+s_1}}\right)^{k}\left(1-\frac{1}{p^{1+s_2}}\right)^{k}.$$

Substituting this back into (4.5) we obtain

$$\frac{1}{(2i\pi)^2}\int\int_{\substack{\mathrm{Re}(s_1)=2\\ \mathrm{Re}(s_2)=2}}\frac{\zeta(1+s_1+s_2)^k}{\zeta(1+s_1)^k\zeta(1+s_2)^k}G(s_1,s_2)\ \ R^{s_1+s_2}\frac{ds_2}{s_2}\cdot\frac{ds_1}{s_1}.$$

where

$$G(s_1,s_2) := \prod_{p \text{ prime}} \left(1-\frac{1}{p^{1+s_1+s_2}}\right)^k \left(1-\frac{1}{p^{1+s_1}}\right)^{-k} \left(1-\frac{1}{p^{1+s_2}}\right)^{-k} \left(1-\frac{\omega(p)}{p^{1+s_1}}-\frac{\omega(p)}{p^{1+s_2}}+\frac{\omega(p)}{p^{1+s_1+s_2}}\right).$$

The idea is to move both contours in the integral slightly to the left of $\mathrm{Re}(s_1) = \mathrm{Re}(s_2) = 0$, and show that the main contribution comes, via Cauchy's Theorem, from the pole at $s_1 = s_2 = 0$. This can be achieved using our understanding of the Riemann-zeta function, and noting that

$$G(0,0) := \prod_{p \text{ prime}} \left(1-\frac{\omega(p)}{p}\right)\left(1-\frac{1}{p}\right)^{-k} = C(a) \neq 0.$$

Remarkably when one does the analogous calculation with the first sum in (4.4), one takes $k-1$ in place of k, and then

$$G^*(0,0) := \prod_{p \text{ prime}} \left(1-\frac{\omega^*(p)}{p-1}\right)\left(1-\frac{1}{p}\right)^{-(k-1)} = C(a),$$

also. Since it is so unlikely that these two quite different products give the same constant by co-incidence, one can feel sure that the method is correct!

This was the technique used in [15] and, although the outline of the method is quite compelling, the details of the contour shifting can be complicated.

4.4. **Evaluating the sums using Selberg's combinatorial approach, I.** As discussed, the difficulty in evaluating the sums in (4.4) is that there are many positive terms and many negative terms. In developing his upper bound sieve method, Selberg encountered a similar problem and dealt with it in a surprising way, using combinatorial identities to remove this issue. The method rests on a *reciprocity law*: Suppose that $L(d)$ and $Y(r)$ are sequences of numbers, supported only on the squarefree integers. If

$$Y(r) := \mu(r) \sideset{}{'}\sum_{m:\ r|m} L(m) \text{ for all } r \geqslant 1,$$

then

$$L(d) = \mu(d) \sideset{}{'}\sum_{n:\ d|n} Y(n) \text{ for all } d \geqslant 1$$

Here, and henceforth, $\sum'$ denotes the restriction to squarefree integers that are $\leqslant R$. Let ϕ_ω be the multiplicative function (defined here, only on squarefree integers) for which $\phi_\omega(p) = p - \omega(p)$. We apply the above reciprocity law with

$$L(d) := \frac{\lambda(d)\omega(d)}{d} \quad \text{and} \quad Y(r) := \frac{y(r)\omega(r)}{\phi_\omega(r)}.$$

Now since $d_1d_2 = D(d_1,d_2)$ we have

$$\lambda(d_1)\lambda(d_2)\frac{\omega(D)}{D} = L(d_1)L(d_2)\ \frac{(d_2,d_2)}{\omega((d_2,d_2))}$$

and therefore

$$S_1 := \sideset{}{'}\sum_{\substack{d_1,d_2\\ D:=[d_1,d_2]}} \lambda(d_1)\lambda(d_2)\frac{\omega(D)}{D} = \sum_{r,s} Y(r)Y(s) \sideset{}{'}\sum_{\substack{d_1,d_2\\ d_1|r,\ d_2|s}} \mu(d_1)\mu(d_2)\frac{(d_1,d_2)}{\omega((d_1,d_2))}.$$

The summand (of the inner sum) is multiplicative and so we can work out its value, prime-by-prime. We see that if $p|r$ but $p \nmid s$ (or vice-versa) then the sum is $1-1=0$. Hence if the sum is non-zero then $r=s$ (as r and s are both squarefree). In that case, if $p|r$ then the sum is $1-1-1+p/\omega(p) = \phi_\omega(p)/\omega(p)$. Hence the sum becomes

$$S_1 = \sum_r Y(r)^2 \frac{\phi_\omega(r)}{\omega(r)} = \sum_r \frac{y(r)^2\omega(r)}{\phi_\omega(r)}. \tag{4.7}$$

squaresum

We will select

$$y(r) := F\left(\frac{\log r}{\log R}\right)$$

when r is squarefree, where $F(t)$ is measurable and supported only on $[0,1]$; and $y(r)=0$ otherwise. Hence we now have a sum with all positive terms so we do not have to fret about complicated cancelations.

4.5. **Sums of multiplicative functions.** An important theme in analytic number theory is to understand the behaviour of sums of multiplicative functions, some being easier than others. Multiplicative functions f for which the $f(p)$ are fixed, or almost fixed, were the first class of non-trivial sums to be determined. Indeed from the Selberg-Delange theorem,[10] one can deduce that

$$\sum_{n\leqslant x} \frac{g(n)}{n} \sim \kappa(g)\cdot \frac{(\log x)^k}{k!}, \tag{4.8}$$

SD

where

$$\kappa(g) := \prod_{p \text{ prime}} \left(1+\frac{g(p)}{p}+\frac{g(p^2)}{p^2}+\dots\right)\left(1-\frac{1}{p}\right)^k$$

when $g(p)$ is typically "sufficiently close" to some given positive integer k that the Euler product converges. Moreover, by partial summation, one deduces that

$$\sum_{n\leqslant x} \frac{g(n)}{n} F\left(\frac{\log n}{\log x}\right) \sim \kappa(g)(\log x)^k \cdot \int_0^1 F(t)\frac{t^{k-1}}{(k-1)!}dt. \tag{4.9}$$

SD+

We apply this in the sum above, noting that here $\kappa(g) = 1/C(a)$, to obtain

$$C(a)S_1 = C(a)\ \sum_r \frac{\omega(r)}{\phi_\omega(r)} F\left(\frac{\log r}{\log R}\right)^2 \sim (\log R)^k \cdot \int_0^1 F(t)^2 \frac{t^{k-1}}{(k-1)!}dt.$$

A similar calculation reveals that

$$C(a)\lambda(d) \sim \mu(d)\cdot(1-v_d)^k \int_{v_d}^1 F(t)\frac{t^{k-1}}{(k-1)!}dt \cdot (\log R)^k,$$

where $v_d := \frac{\log d}{\log R}$.

[10]This also follows from the relatively easy proof of Theorem 1.1 of IK [26].

4.6. **Selberg's combinatorial approach, II.** A completely analogous calculation, but now applying the reciprocity law with

$$L(d) := \frac{\lambda(d)\omega^*(d)}{\phi(d)} \quad \text{and} \quad Y(r) := \frac{y^*(r)\omega^*(r)}{\phi_\omega(r)},$$

yields that

$$S_2 := \sideset{}{'}\sum_{\substack{d_1,d_2\\ D:=[d_1,d_2]}} \lambda(d_1)\lambda(d_2)\frac{\omega^*(D)}{\phi(D)} = \sum_r \frac{y^*(r)^2\omega^*(r)}{\phi_\omega(r)}. \tag{4.10}$$

We need to determine $y^*(r)$ in terms of the $y(r)$, which we achieve by applying the reciprocity law twice:

$$\begin{aligned} y^*(r) &= \mu(r)\frac{\phi_\omega(r)}{\omega^*(r)} \sum_{d:\ r|d} \frac{\omega^*(d)}{\phi(d)}\mu(d)\frac{d}{\omega(d)} \sum_{n:\ d|n} \frac{y(n)\omega(n)}{\phi_\omega(n)} \\ &= \frac{r}{\phi(r)} \sum_{n:\ r|n} \frac{y(n)}{\phi_\omega(n/r)} \sum_{d:\ d/r|n/r} \mu(d/r)\frac{\omega^*(d/r)d/r}{\phi(d/r)}\omega(n/d) \\ &= r \sideset{}{'}\sum_{n:\ r|n} \frac{y(n)}{\phi(n)} = \frac{r}{\phi(r)} \sideset{}{'}\sum_{m:\ (m,r)=1} \frac{y(mr)}{\phi(m)} \\ &\sim \int_{\frac{\log r}{\log R}}^1 F(t)dt \cdot \log R, \end{aligned}$$

where the last estimate was obtained by applying (4.9) with $k = 1$, and taking care with the Euler product.

We now can insert this into (4.10), and apply (4.9) with k replaced by $k-1$, noting that $\kappa(g^*) = 1/C(a)$, to obtain

$$C(a)S_2 = C(a)\sum_r \frac{y^*(r)^2\omega^*(r)}{\phi_\omega(r)} \sim (\log R)^{k+1} \cdot \int_0^1 \left(\int_t^1 F(u)du\right)^2 \frac{t^{k-2}}{(k-2)!}dt.$$

4.7. **Finding a positive difference; the proof of Theorem 3.1.** From these estimate, we deduce that $C(a)$ times (4.4) is asymptotic to $x(\log 3x)(\log R)^k$ times

$$k\,\frac{\log R}{\log 3x} \cdot \int_0^1 \left(\int_t^1 F(u)du\right)^2 \frac{t^{k-2}}{(k-2)!}dt - \int_0^1 F(t)^2\frac{t^{k-1}}{(k-1)!}dt. \tag{4.11}$$

Define

$$\rho_k(F) := k\int_0^1 \left(\int_t^1 F(u)du\right)^2 \frac{t^{k-2}}{(k-2)!}dt \Big/ \int_0^1 F(t)^2\frac{t^{k-1}}{(k-1)!}dt. \tag{4.12}$$

Assume that the Elliott-Halberstam conjecture holds with exponent $\frac{1}{2}+\eta$, so that we may take $R = \sqrt{Q}$. Hence we deduce that if

$$\frac{1}{2}\left(\frac{1}{2}+\eta\right)\ \rho_k(F) > 1$$

for some F that satisfies the above hypotheses, then (4.11) implies that (4.4), and so (4.1), is > 0

We now need to select a suitable function $F(t)$ to proceed. A good choice is $F(t) = \frac{(1-t)^\ell}{\ell!}$. Using the beta integral identity

$$\int_0^1 \frac{v^k}{k!}\frac{(1-v)^\ell}{\ell!}dv = \frac{1}{(k+\ell+1)!},$$

we obtain

$$\int_0^1 F(t)^2 \frac{t^{k-1}}{(k-1)!}dt = \int_0^1 \frac{(1-t)^{2\ell}}{\ell!^2}\frac{t^{k-1}}{(k-1)!}dt = \frac{1}{(k+2\ell)!}\binom{2\ell}{\ell},$$

and

$$\int_0^1 \left(\int_t^1 F(u)du\right)^2 \frac{t^{k-2}}{(k-2)!}dt = \int_0^1 \left(\frac{(1-t)^{\ell+1}}{\ell+1}\right)^2 \frac{t^{k-2}}{(k-2)!}dt = \frac{1}{(k+2\ell+1)!}\binom{2\ell+2}{\ell+1}.$$

Therefore (4.12) is > 0 if (3.4) holds, and so we deduce Theorem 3.1.

In particular if the Elliott-Halberstam conjecture holds with exponent $\frac{1}{2}+\eta$, then we select ℓ to be a sufficiently large integer for which $1+2\eta > \left(1+\frac{1}{2\ell+1}\right)^2$. Selecting $k=(2\ell+1)^2$ we deduce that for every admissible k-tuple, there are infinitely many n for which the k-tuple, evaluated at n, contains two primes.

5. Zhang's modifications of GPY

At the end of the previous section we saw that if the Elliott-Halberstam conjecture holds with any exponent $> \frac{1}{2}$, then for every admissible k-tuple, there are infinitely many n for which the k-tuple contains two primes. However the Elliott-Halberstam conjecture remains unproven.

In (3.5) we stated Zhang's result, which breaks the $\sqrt{x}$-barrier in such results, but at the cost of restricting the moduli to being y-smooth, and restricting the arithmetic progressions $a \pmod q$ to having the same value of a as we vary over q . Can the Goldston-Pintz-Yıldırım argument be modified to handle these restrictions?

5.1. **Averaging over arithmetic progressions.**

In the GPY argument we need estimates for the number of primes in the arithmetic progressions $m + a_i \pmod D$ where $m \in \Omega_i(D)$. When using the Bombieri-Vinogradov Theorem, it does not matter that $m + a_i$ varies as we vary over D; but it does matter when employing Zhang's Theorem 3.2.

Zhang realized that one can exploit the structure of the set $O_i(D) = \Omega_i(D) + a_i$, since it is constructed from the $O_i(p)$ with $p|D$ using the Chinese Remainder Theorem, to get around this issue:

Let $\nu(D)$ denote the number of prime factors of (squarefree) D, so that $\tau(D) = 2^{\nu(D)}$. Any squarefree D can be written as $[d_1, d_2]$ for $3^{\nu(D)}$ pairs d_1, d_2, which means that we need an appropriate upper bound on

$$\leqslant \sideset{}{'}\sum_{D\leqslant Q} 3^{\nu(D)} \sum_{b\in O_i(D)} \left|\Theta(X;D,b) - \frac{X}{\phi(D)}\right|$$

where $Q = R^2$ and $X = x$ or $2x$, for each i.

Let L be the lcm of all of the D in our sum. Then the set, $O_i(L)$, reduced mod D, gives $|O_i(L)|/|O_i(D)|$ copies of $O_i(D)$ and so

$$\frac{1}{|O_i(D)|} \sum_{b\in O_i(D)} \left|\Theta(X;D,b) - \frac{X}{\phi(D)}\right| = \frac{1}{|O_i(L)|} \sum_{b\in O_i(L)} \left|\Theta(X;D,b) - \frac{X}{\phi(D)}\right|.$$

Hence we need to divide and multiply by $|O_i(D)|$ in each term of the above sum. Since $|O_i(D)| = \omega^*(D) \leqslant (k-1)^{\nu(D)}$, the above is therefore

$$\begin{aligned} &\leqslant \sideset{}{'}\sum_{D\leqslant Q} \tau(D)^A \cdot \frac{1}{|O_i(D)|} \sum_{b\in O_i(D)} \left|\Theta(X;D,b) - \frac{X}{\phi(D)}\right| \\ &= \frac{1}{|O_i(L)|} \sum_{b\in O_i(L)} \sideset{}{'}\sum_{D\leqslant Q} \tau(D)^A \cdot \left|\Theta(X;D,b) - \frac{X}{\phi(D)}\right| \\ &\leqslant \max_{a\in\mathbb{Z}} \sideset{}{'}\sum_{\substack{D\leqslant Q\\ (D,a)=1}} \tau(D)^A \cdot \left|\Theta(X;D,a) - \frac{X}{\phi(D)}\right| \end{aligned}$$

where $2^A = 3(k-1)$.

It now needs a standard technical argument to bound this using Theorem 3.2: By Cauchy's Theorem, the square of this is

$$\leqslant \sum_{D\leqslant Q} \frac{\tau(D)^{2A}}{D} \cdot \sideset{}{'}\sum_{D\leqslant Q} D \left|\Theta(X;D,b) - \frac{X}{\phi(D)}\right|^2.$$

The first sum is bounded by $(c\log Q)^{9k^2}$, and we have $D|\Theta(X;D,b)| \leqslant (X+D)\log X$, trivially, and so

$$D\left|\Theta(X;D,b) - \frac{X}{\phi(D)}\right| \ll X\log X.$$

The result now follows by applying Theorem 3.2.

5.2. **Restricting the support to smooth integers.** Zhang simply took the same coefficients $y(r)$ as above, but now restricted to y-smooth integers; and called this restricted class of coefficients, $z(r)$. Evidently the sum in (4.7) with $z(r)$ in place of $y(r)$, is bounded above by the sum in (4.7). The sum in (4.10) with $z(r)$ in place of $y(r)$, is a little more tricky, since we need a lower bound. Zhang proceeds by showing that if L is sufficiently large and δ sufficiently small, then the two sums differ by only a negligible amount. In particular we will prove Zhang's Theorem when

$$162\eta + 90\delta < 1.$$

Zhang's argument here holds when $L = 863, k = L^2$ and $\eta = 1/(L-1)$.

It should be noted that Motohashi and Pintz [32] had already given an argument to accomplish the goals of this section, in the hope that someone might prove an estimate like (3.5)!

6. Goldston-Pintz-Yildirim in higher dimensional analysis

In the set-up in the argument of Goldston, Pintz and Yıldırım, we saw that we study the divisors d of the product of the values of the k-tuple; that is

$$d|\mathcal{P}(n) = (n+a_1)\dots(n+a_k).$$

with $d \leqslant R$.

The latest breakthrough stems from the idea of instead studying the k-tuples of divisors $d_1, d_2, \dots, d_k$ of each individual element of the k-tuple; that is

$$d_1|n+a_1,\ d_2|n+a_2, \dots, d_k|n+a_k.$$

Now, instead of $d \leqslant R$, we take $d_1 d_2 \dots d_k \leqslant R$.

6.1. **The set up.** One can proceed much as in the previous section, though technically it is easier to restrict our attention to when n is an appropriate congruence class mod m where m is the product of the primes for which $\omega(p) < k$. (because, if $\omega(p) = k$ then p can only divide one $n + a_i$ at a time). Hence we study

$$S_0 := \sum_{r\in\Omega(m)} \sum_{\substack{n\sim x \\ n\equiv r \pmod m}} \left(\sum_{j=1}^k \theta(n+a_j) - h\log 3x \right) \left(\sum_{d_i|n+a_i \text{ for each } i} \lambda(d_1,\dots,d_k) \right)^2$$

which upon expanding, as $(d_i, m)|(n+a_i, m) = 1$, equals

$$\sum_{\substack{d_1,\dots,d_k\geqslant 1 \\ e_1,\dots,e_k \geqslant 1 \\ (d_ie_i,m)=1 \text{ for each } i}} \lambda(d_1,\dots,d_k)\lambda(e_1,\dots,e_k) \sum_{r\in\Omega(m)} \sum_{\substack{n\sim x \\ n\equiv r \pmod m \\ [d_i,e_i]|n+a_i \text{for each } i}} \left(\sum_{j=1}^k \theta(n+a_j) - h\log 3x \right).$$

Next notice that $[d_i, e_i]$ is coprime with $[d_j, e_j]$ whenever $i \neq j$, since their gcd divides $(n + a_j) - (n + a_i)$, which divides m, and so equals 1 as $(d_ie_i, m) = 1$. Hence, in our internal sum, the values of n belong to an arithmetic progression with modulus $m\prod_i [d_i, e_i]$. Also notice that if $n + a_j$ is prime then $d_j = e_j = 1$.

Therefore, ignoring error terms,

$$S_0 = \sum_{1\leqslant \ell\leqslant k} \frac{\omega(m)}{\phi(m)} S_{2,\ell} \cdot x - h\frac{\omega(m)}{m} S_1 \cdot x\log 3x$$

where

$$S_1 := \sum_{\substack{d_1,\dots,d_k\geqslant 1 \\ e_1,\dots,e_k \geqslant 1 \\ (d_i,e_j)=1 \text{ for } i\neq j}} \frac{\lambda(d_1,\dots,d_k)\lambda(e_1,\dots,e_k)}{\prod_i [d_i,e_i]}$$

and

$$S_{2,\ell} := \sum_{\substack{d_1,\dots,d_k\geqslant 1\\ e_1,\dots,e_k\geqslant 1\\ (d_i,e_j)=1 \text{ for } i\neq j\\ d_\ell=e_\ell=1}} \frac{\lambda(d_1,\dots,d_k)\lambda(e_1,\dots,e_k)}{\prod_i \phi([d_i,e_i])}.$$

6.2. **The combinatorics.** The reciprocity law generalizes quite beautifully to higher dimension: Suppose that $L(d)$ and $Y(r)$ are two sequences of complex numbers, indexed by $d, r \in \mathbb{Z}^k_{\geqslant 1}$, and non-zero only when each d_i (or r_i) is squarefree. Then

$$L(d_1,\dots,d_k) = \prod_{i=1}^k \mu(d_i) \sum_{\substack{r_1,\dots,r_k\geqslant 1\\ d_i|r_i \text{ for all } i}} Y(r_1,\dots,r_k)$$

if and only if

$$Y(r_1,\dots,r_k) = \prod_{i=1}^k \mu(r_i) \sum_{\substack{d_1,\dots,d_k\geqslant 1\\ r_i|d_i \text{ for all } i}} L(d_1,\dots,d_k).$$

We use this much as above, in the first instance with

$$L(d_1,\dots,d_k) = \frac{\lambda(d_1,\dots,d_k)}{d_1,\dots,d_k} \text{ and } Y(r_1,\dots,r_k) = \frac{y(r_1,\dots,r_k)}{\phi_k(r_1\dots r_k)}$$

where

$$y(r_1,\dots,r_k) = F\left(\frac{\log r_1}{\log R},\dots,\frac{\log r_k}{\log R}\right)$$

with $F \in \mathbb{C}[t_1,\dots,t_k]$, such that that there is a uniform bound on all of the first order partial derivatives, and F is only supported on

$$T_k := \{(t_1,\dots,t_k) : \text{ Each } t_j \geqslant 1, \text{ and } t_1+\dots+t_k \leqslant 1\}.$$

Proceeding much as before we obtain

$$S_1 \sim \sum_{r_1,\dots,r_k\geqslant 1} \frac{y(r_1,\dots,r_k)^2}{\phi_k(r_1\dots r_k)}. \tag{6.1}$$ S1value

6.3. **Sums of multiplicative functions.** By (SD 4.8) we have

$$\sum_{\substack{1\leqslant n\leqslant N\\ (n,m)=1}} \frac{\mu^2(n)}{\phi_k(n)} = \prod_{p|m} \frac{p-1}{p} \prod_{p\nmid m} \frac{(p-1)\phi_{k-1}(p)}{p\ \phi_k(p)} \cdot (\log N + O(1)) \tag{6.2}$$ sumeval

We apply this k times; firstly with m replaced by $mr_1\dots r_{k-1}$ and n by r_k, then with m replaced by $mr_1\dots r_{k-2}$, etc By the end we obtain

$$C_m(a) \sum_{\substack{1\leqslant r_1\leqslant R_1,\\ \dots,\\ 1\leqslant r_k\leqslant R_k}} \frac{\mu^2(r_1\dots r_k m)}{\phi_k(r_1,\dots,r_k)} = \prod_i (\log R_i + O(1)), \tag{6.3}$$ step0

where

$$C_m(a) := \prod_{p|m}\left(1-\frac{1}{p}\right)^{-k}\prod_{p\nmid m}\left(1-\frac{k}{p}\right)\left(1-\frac{1}{p}\right)^{-k}.$$

From this, and partial summation, we deduce from (S1value 6.1), that

$$C_m(a)S_1 \sim (\log R)^k \cdot \int_{t_1,\dots,t_k\in T_k} F(t_1,\dots,t_k)^2 dt_k\dots dt_1. \tag{6.4}$$

ResultS1

Had we stopped our calculation one step earlier we would have found

$$C_m(a)\sum_{\substack{1\leqslant r_1\leqslant R_1,\\ \dots,\\ 1\leqslant r_{k-1}\leqslant R_{k-1}}} \frac{\mu^2(r_1\dots r_{k-1}m)}{\phi_k(r_1,\dots,r_{k-1})} = \frac{m}{\phi(m)}\cdot\prod_i(\log R_i + O(1)), \tag{6.5}$$

step1

6.4. **The combinatorics, II.** We will deal only with the case $\ell = k$, the other cases being analogous. Now we use the higher dimensional reciprocity law with

$$L(d_1,\dots,d_{k-1}) = \frac{\lambda(d_1,\dots,d_{k-1},1)}{\phi(d_1\dots d_{k-1})} \quad\text{and}\quad Y_k(r_1,\dots,r_{k-1}) = \frac{y_k(r_1,\dots,r_{k-1})}{\phi_k(r_1\dots r_{k-1})}$$

where $d_k = r_k = 1$, so that, with the exactly analogous calculations as before,

$$S_{2,k} \sim \sum_{r_1,\dots,r_{k-1}\geqslant 1} \frac{y_k(r_1,\dots,r_{k-1})^2}{\phi_k(r_1\dots r_{k-1})}.$$

Using the reciprocity law twice to determine the $y_k(r)$ in terms of the $y(n)$, we obtain that

$$y_k(r_1,\dots,r_{k-1}) \sim \frac{\phi(m)}{m}\cdot\int_{t\geqslant 0} F(\rho_1,\dots,\rho_{k-1},t)dt\cdot\log R$$

where each $r_i = N^{\rho_i}$. Therefore, using (step1 6.5), we obtain

$$C_m(a)S_{2,k} \sim \int_{0\leqslant t_1,\dots,t_{k-1}\leqslant 1}\left(\int_{t_k\geqslant 0} F(t_1,\dots,t_{k-1},t_k)dt_k\right)^2 dt_{k-1}\dots dt_1\cdot\frac{\phi(m)}{m}(\log R)^{k+1}. \tag{6.6}$$

ResultS2

6.5. **Finding a positive difference.** By the Bombieri-Vinogradov Theorem we can take $R = x^{1/4-o(1)}$, so that, by (ResultS1 6.4) and (ResultS2 6.6), $C_m(a)S_0$ equals $\frac{\omega(m)}{m}x(\log 3x)(\log R)^k$ times

$$\frac{1}{4}\sum_{\ell=1}^{k}\int_{\substack{0\leqslant t_i\leqslant 1 \text{ for}\\ 1\leqslant i\leqslant k,\ i\neq\ell}}\left(\int_{t_\ell\geqslant 0}F(t_1,\dots,t_k)dt_\ell\right)^2\prod_{\substack{1\leqslant j\leqslant k\\ i\neq\ell}}dt_j - h\int_{t_1,\dots,t_k\in T_k}F(t_1,\dots,t_k)^2dt_k\dots dt_1 + o(1).$$

One can show that the optimal choice for F must be symmetric. Hence $S_0 > 0$ follows if there exists a symmetric F (with the restrictions above) for which the ratio

$$\rho(F) := \frac{k\int_{t_1,\dots,t_{k-1}\geqslant 0}\left(\int_{t_k\geqslant 0}F(t_1,\dots,t_k)dt_k\right)^2 dt_{k-1}\dots dt_1}{\int_{t_1,\dots,t_k\geqslant 0}F(t_1,\dots,t_k)^2dt_k\dots dt_1}.$$

satisfies $\rho(F) > 4h$.

TechProp

Proposition 6.1. *Fix $h \geqslant 1$. Suppose that there exists $F \in \mathbb{C}(x_1, \dots, x_k)$ which is measurable, supported on T_k, for which there is a uniform bound on the first order partial derivatives and such that $\rho(F) > 4h$. Then, for every admissible k-tuple of linear forms, there are infinitely many integers n such that there are $> h$ primes amongst the k linear forms when evaluated at n. If the Elliott-Halberstam conjecture holds then we only need that $\rho(F) > 2h$.*

6.6. **A special case.** If $F(t_1, \dots, t_k) = f(t_1 + \dots + t_k)$ then since

$$\int_{\substack{t_1,\dots,t_k \geqslant 0 \\ t_1+\dots+t_k=t}} dt_{k-1}\dots dt_1 = \frac{t^{k-1}}{(k-1)!},$$

we deduce that

$$\rho(F) = \rho_k(f)$$

as defined in (4.12); that is, reduce to the original GPY argument.

We need to make some choices for F that do not lead back to the original GPY argument, in the hope that we can do better; evidently we should avoid selecting F to be a function of one variable. Since F is symmetric it makes sense to define the symmetric sums as $P_j = \sum_{i=1}^k t_i^j$; in the GPY argument F was a function of P_1. A first guess might be to work now with functions of P_1 and P_2, so as to consider functions F that do not appear in the GPY argument.

6.7. **Maynard's Fs, and gaps between primes.** For $k = 5$ let

$$F(t_1, \dots, t_5) = 70P_1P_2 - 49P_1^2 - 75P_2 + 83P_1 - 34.$$

A calculation yields that

$$\rho(F) = \frac{1417255}{708216} > 2.$$

Therefore, by Proposition 6.1, if we assume the Elliott-Halberstam conjecture with $h = 1$ then for every admissible 5-tuple of linear forms, there are infinitely many integers n such that there are at least two primes amongst the five linear forms when evaluated at n. In particular, from the admissible forms $\{x, x+2, x+6, x+8, x+12\}$ we deduce that there are infinitely many pairs of distinct primes that differ by no more than 12. Also from the admissible forms $\{x+1, 2x+1, 4x+1, 8x+1, 16x+1\}$ we deduce that there are infinitely many pairs of distinct primes, p, q for which $(p-1)/(q-1) = 2^j$ for $j = 0, 1, 2, 3$ or 4.

Unconditionally, Maynard shows that there exists a polynomial of the form

$$\sum_{\substack{a,b \geqslant 0 \\ a+2b \leqslant 11}} c_{a,b}(1-P_1)^a P_2^b$$

with $k = 105$, for which

$$\rho(F) = 4.0020697\dots$$

How does Maynard prove this? With F of the above form, one sees that both the numerator and denominator of $\rho(F)$ are quadratic forms in the variables $c_{a,b}$. There are 42 such coefficients, and we let a be the vector of c-values. Therefore there exist easily calculable matrices M_1 and M_2 for which the numerator of F is $a^T M_2 a$, and the denominator is $a^T M_1 a$. By the theory of Lagrangian multipliers, Maynard shows that

$$M_1^{-1} M_2 a = \rho(F) a$$

so that $\rho(f)$ can be taken to be the largest eigenvalue of $M_1^{-1} M_2$, and a the corresponding eigenvector. These calculations are easily completed using a computer algebra package and yield the result above.

By Proposition 6.1 with $h = 1$, we deduce that for every admissible 105-tuple of linear forms, there are infinitely many integers n such that there are at least two primes amongst the 105 linear forms when evaluated at n.

6.8. F **as a product of one dimensional functions.** We make the choice that

$$F(t_1, \dots t_k) = \begin{cases} g(kt_1) \dots g(kt_k) & \text{if } t_1 + \dots + t_k \leqslant 1 \\ 0 & \text{otherwise,} \end{cases}$$

where g is some integrable function supported only on $[0, T]$. Let $\gamma := \int_{t \geqslant 0} g(t)^2 dt$, so that the denominator of $\rho(F)$ is

$$I_k = \int_{t \in T_k} f(t_1, \dots t_k)^2 dt_k \dots dt_1 \leqslant \int_{t_1, \dots, t_k \geqslant 0} (g(kt_1) \dots g(kt_k))^2 dt_k \dots dt_1 = k^{-k} \gamma^k.$$

We rewrite the numerator of $\rho(F)$ as $L_k - M_k$ where

$$L_k := k \int_{t_1, \dots, t_{k-1} \geqslant 0} \left(\int_{t_k \geqslant 0} g(kt_1) \dots g(kt_k) dt_k \right)^2 dt_{k-1} \dots dt_1 = k^{-k} \gamma^{k-1} \left(\int_{t \geqslant 0} g(t) dt \right)^2.$$

As $g(t)$ is only supported in $[0, T]$ we have, by Cauchying and letting $u_j = kt_j$,

$$\begin{aligned} M_k : &= \int_{t_1, \dots, t_{k-1} \geqslant 0} \left(\int_{t_k \geqslant 1 - t_1 - \dots - t_{k-1}} g(kt_1) \dots g(kt_k) dt_k \right)^2 dt_{k-1} \dots dt_1 \\ &\leqslant k^{-k} T \int_{\substack{u_1, \dots, u_k \geqslant 0 \\ u_1 + \dots + u_k \geqslant k}} g(u_1)^2 \dots g(u_k)^2 du_1 \dots du_k. \end{aligned}$$

Now assume that $\mu := \int_t t g(t)^2 dt \leqslant (1 - \eta) \int_t g(t)^2 dt = (1 - \eta)\gamma$ for some given $\eta > 0$; that is, that the "weight" of g^2 is centered around values of $t \leqslant 1 - \eta$. We have

$$1 \leqslant \eta^{-2} \left(\frac{1}{k}(u_1 + \dots + u_k) - \mu/\gamma \right)^2$$

whenever $u_1 + \dots + u_k \geqslant k$. Therefore,

$$\begin{aligned} M_k &\leqslant \eta^{-2} k^{-k} T \int_{u_1, \dots, u_k \geqslant 0} g(u_1)^2 \dots g(u_k)^2 \left(\frac{1}{k}(u_1 + \dots + u_k) - \mu/\gamma \right)^2 du_1 \dots du_k \\ &= \eta^{-2} k^{-k-1} T \int_{u_1, \dots, u_k \geqslant 0} g(u_1)^2 \dots g(u_k)^2 (u_1^2 - \mu^2/\gamma^2) du_1 \dots du_k \\ &= \eta^{-2} k^{-k-1} \gamma^{k-1} T \left(\int_{u \geqslant 0} u^2 g(u)^2 du - \mu^2/\gamma \right) \leqslant \eta^{-2} k^{-k-1} \gamma^{k-1} T \int_{u \geqslant 0} u^2 g(u)^2 du, \end{aligned}$$

by symmetry. We deduce that

$$\rho(F) \geqslant \frac{\left(\int_{t\geqslant 0} g(t)dt\right)^2 - \frac{\eta^{-2}T}{k}\int_{u\geqslant 0} u^2 g(u)^2 du}{\int_{t\geqslant 0} g(t)^2 dt}. \tag{6.7}$$ 1stLowerBo

Notice that we can multiply g through by a scalar and not effect the value in (6.7).

6.9. **The optimal choice.** We wish to find the value of g that maximizes the right-hand side of (6.7). This can be viewed as an optimization problem:

Maximize $\int_{t\geqslant 0} g(t)dt$, subject to the constraints $\int_{t\geqslant 0} g(t)^2dt = \gamma$ and $\int_{t\geqslant 0} tg(t)^2 dt = \mu$.

One can approach this using the calculus of variations or even by discretizing g and employing the technique of Lagrangian multipliers. The latter gives rise to (a discrete form of)

$$\int_{t\geqslant 0} g(t)dt - \alpha\left(\int_{t\geqslant 0} g(t)^2 dt - \gamma\right) - \beta\left(\int_{t\geqslant 0} tg(t)^2 dt - \mu\right),$$

for unknowns α and β. Differentiating with respect to $g(v)$ for each $v \in [0,T]$, we obtain

$$1 - 2\alpha g(v) - 2\beta v g(v) = 0;$$

that is, after re-scaling,

$$g(t) = \frac{1}{1+At} \quad \text{for} \quad 0 \leqslant t \leqslant T,$$

for some real $A > 0$. We select T so that $1 + AT = e^A$, and let $A > 1$. We then calculate the integrals in (6.7):

$$\begin{aligned} \gamma = \int_t g(t)^2 dt &= \frac{1}{A}(1 - e^{-A}), \\ \int_t tg(t)^2 dt &= \frac{1}{A^2}\left(A - 1 + e^{-A}\right), \\ \int_t t^2 g(t)^2 dt &= \frac{1}{A^3}\left(e^A - 2A - e^{-A}\right), \\ \text{and} \qquad \int_t g(t)dt &= 1, \\ \text{so that} \qquad \eta &= \frac{1-(A-1)e^{-A}}{A(1-e^{-A})} > 0, \end{aligned}$$

which is necessary. (6.7) then becomes

$$\rho(F) \geqslant \frac{A}{(1-e^{-A})} - \frac{e^{2A}}{Ak}\left(1 - 2Ae^{-A} - e^{-2A}\right)\frac{(1-e^{-A})^2}{(1-(A-1)e^{-A})^2} \geqslant A - \frac{e^{2A}}{Ak} \tag{6.8}$$ 2ndLowerBo

Taking $A = \frac{1}{2}\log k + \frac{1}{2}\log\log k$, we deduce that

$$\rho(F) \geqslant \frac{1}{2}\log k + \frac{1}{2}\log\log k - 2.$$

Hence, for every $m \geqslant 1$ we find that $\rho(F) > 4m$ provided $e^{8m+4} < k\log k$.

This implies the following result:

Theorem 6.2. *For any given integer $m \geqslant 2$, let k be the smallest integer with $k \log k > e^{8m+4}$. For* any *admissible k-tuple of linear forms $L_1, \dots, L_k$ there exists infinitely many integers n such that at least m of the $L_j(n)$, $1 \leqslant j \leqslant k$ are prime.*

For any $m \geqslant 1$, we let k be the smallest integer with $k \log k > e^{8m+4}$, so that $k > 10000$; in this range it is known that $\pi(k) \leqslant \frac{k}{\log k - 4}$. Next we let $x = 2k \log k > 10^5$ and, for this range it is known that $\pi(x) \geqslant \frac{x}{\log x}(1 + \frac{1}{\log x})$. Hence

$$\pi(2k \log k) - \pi(k) \geqslant \frac{2k \log k}{\log(2k \log k)} \left(1 + \frac{1}{\log(2k \log k)}\right) - \frac{k}{\log k - 4}$$

and this is $> k$ for $k \geqslant 311$ by an easy calculation. We therefore apply the theorem with the k smallest primes $> k$, which form an admissible set $\subset [1, 2k \log k]$, to obtain:

Corollary 6.3. *For any given integer $m \geqslant 2$, let $B_m = e^{8m+5}$. There are infinitely many integers x for which there are at least m distinct primes within the interval $[x, x + B_m]$.*

By a slight modification of this construction, Maynard obtains $B_m \ll m^3 e^{4m}$ in [29].

Part 2. Primes in arithmetic progressions; breaking the $\sqrt{x}$-barrier

Our goal, in the rest of the article, is to sketch the ideas behind the proof of Yitang's extraordinary result, given in (3.5), that primes are well-distributed on average in the arithmetic progressions $a \pmod q$ with q a little bigger than $\sqrt{x}$. We will see how this question fits into a more general framework, as developed by Bombieri, Friedlander and Iwaniec [3], so that Zhang's results should also allow us to deduce analogous results for interesting arithmetic sequences other than the primes.

To begin with we will need to discuss a key technique of analytic number theory, the idea of creating important sequences through convolutions:

7. Convolutions in number theory

The convolution of two functions f and g, written $f * g$, is defined by

$$(f * g)(n) := \sum_{ab=n} f(a)g(b),$$

for every integer $n \geqslant 1$, where the sum is over all pairs of positive integers a, b whose product is n. Hence if $\tau(n)$ counts the number of divisors of n then

$$\tau = 1 * 1,$$

where 1 is the function with $1(n) = 1$ for every $n \geqslant 1$. We already saw, in (2.2), that if $L(n) = \log n$ then $\mu * L = \Lambda$, where $\Lambda(n) = \log p$ if n is a power of prime p, and $\Lambda(n) = 0$ otherwise. In the GPY argument we used that $(1 * \mu)(n) = 0$ if $n > 1$.

There is no better way to understand why convolutions are useful than to present a famous argument of Dirichlet, estimating the average of $\tau(n)$. Now , if n is squarefree and has k prime factors then $\tau(n) = 2^k$, so we see that $\tau(n)$ varies greatly depending on the arithmetic structure of n, but the average is more stable:

$$\begin{aligned}\frac{1}{x}\sum_{n\leqslant x}\tau(n) &= \frac{1}{x}\sum_{n\leqslant x}\sum_{d|n}1 = \frac{1}{x}\sum_{d|n}\sum_{\substack{n\leqslant x\\ d|n}}1 = \frac{1}{x}\sum_{d\leqslant x}\left[\frac{x}{d}\right]\\ &= \frac{1}{x}\sum_{d\leqslant x}\left(\frac{x}{d}+O(1)\right) = \sum_{d\leqslant x}\frac{1}{d} + O\left(\frac{1}{x}\sum_{d\leqslant x}1\right).\end{aligned}$$

One can approximate $\sum_{d\leqslant x}\frac{1}{d}$ by $\int_1^x dt/t = \log x$. Indeed the difference tends to a limit, the Euler-Mascheroni constant $\gamma := \lim_{N\to\infty}\frac{1}{1}+\frac{1}{2}+\ldots+\frac{1}{N}-\log N$. Hence we have proved that the integers up to x have $\log x + O(1)$ divisors, on average, which is quite remarkable for such a wildly fluctuating function.

Dirichlet studied this argument and noticed that when we approximate $[x/d]$ by $x/d + O(1)$ for large d, say for those d in $(x/2, x]$, then this is not really a very good approximation, and gives a large cumulative error term, $O(x)$. However we know that $[x/d] = 1$ for each of these d, and so we can estimate this sum by $x/2 + O(1)$, which is much more precise. In general we write $n = dm$, where d and m are integers. When d is small then we should fix d, and count the number of such m, with $m \leqslant x/d$ (as we did above); but when m is small, then we should fix m, and count the number of d with $d \leqslant x/m$. In this way our sums are all over long intervals, which allows us to get an accurate approximation of their value:

$$\begin{aligned}\frac{1}{x}\sum_{n\leqslant x}\tau(n) &= \frac{1}{x}\sum_{n\leqslant x}\sum_{dm=n}1 = \frac{1}{x}\sum_{d\leqslant\sqrt{x}}\sum_{\substack{n\leqslant x\\ d|n}}1 + \frac{1}{x}\sum_{m<\sqrt{x}}\sum_{\substack{n\leqslant x\\ m|n}}1 - \frac{1}{x}\sum_{d\leqslant\sqrt{x}}\sum_{m<\sqrt{x}}1\\ &= \frac{1}{x}\sum_{d\leqslant\sqrt{x}}\left(\frac{x}{d}+O(1)\right) + \frac{1}{x}\sum_{m<\sqrt{x}}\left(\frac{x}{m}+O(1)\right) - 1 + O\left(\frac{1}{\sqrt{x}}\right)\\ &= \log x + 2\gamma - 1 + O\left(\frac{1}{\sqrt{x}}\right),\end{aligned}$$

since $\sum_{n\leqslant N} 1/n = \log N + \gamma + O(1/N)$, an extraordinary improvement upon the earlier error term.

7.1. **Vaughan's identity.** We will need a more convoluted identity than (2.2) [VMidentity] to prove our estimates for primes in arithmetic progressions. There are several possible suitable identities, the simplest of which is due to Vaughan [40] [vaughan]:

$$\textit{Vaughan's identity:}\qquad \Lambda_{\geqslant V} = \mu_{<U} * L - \mu_{<U} * \Lambda_{<V} * 1 + \mu_{\geqslant U} * \Lambda_{\geqslant V} * 1 \qquad (7.1)$$ Vaughident

where $g_{>W}(n) = g(n)$ if $n > W$ and $g(n) = 0$ otherwise; and $g = g_{\leqslant W} + g_{>W}$. To verify this identity, we manipulate the algebra of convolutions:

$$\begin{aligned}\Lambda_{\geqslant V} &= \Lambda - \Lambda_{<V} = (\mu * L) - \Lambda_{<V} * (1 * \mu) \\ &= \mu_{<U} * L + \mu_{\geqslant U} * L - \mu_{<U} * \Lambda_{<V} * 1 - \mu_{\geqslant U} * \Lambda_{<V} * 1 \\ &= \mu_{<U} * L - \mu_{<U} * \Lambda_{<V} * 1 + \mu_{\geqslant U} * (\Lambda * 1 - \Lambda_{<V} * 1),\end{aligned}$$

GeneralBV

8. Distribution in arithmetic progressions

8.1. **General sequences in arithmetic progressions.** One can ask whether *any* given sequence $(\beta(n))_{n\geqslant 1} \in \mathbb{C}$ is well-distributed in arithmetic progressions modulo q. We begin by formulating an appropriate analogy to (3.2), which should imply non-trivial estimates in the range $q \leqslant (\log x)^A$ for any fixed $A > 0$: We say that β satisfies a *Siegel-Walfisz condition* if, for any fixed $A > 0$, and whenever $(a, q) = 1$, we have

$$\left| \sum_{\substack{n\leqslant x \\ n\equiv a \pmod q}} \beta(n) - \frac{1}{\phi(q)} \sum_{\substack{n\leqslant x\\ (n,q)=1}} \beta(n) \right| \ll_A \frac{\|\beta\| x^{\frac12}}{(\log x)^A} ,$$

with $\|\beta\| = \|\beta\|_2$ where, as usual,

$$\|\beta\|_2 := \left(\sum_{n\leqslant x} |\beta(n)|^2 \right)^{\frac12} .$$

Using Cauchy's inequality one can show that this assumption is "non-trivial" only for $q < (\log x)^{2A}$; that is, when x is very large compared to q.

Using the large sieve, Bombieri, Friedlander and Iwaniec [3] were able to prove two results that are very surprising, given the weakness of the hypotheses. In the first they showed that if β satisfies a Siegel-Walfisz condition,[11] then it is well-distributed for *almost all* arithmetic progressions $a \pmod q$, for *almost all* $q \leqslant x/(\log x)^B$:

Theorem 8.1. *Suppose that the sequence of complex numbers $\beta(n), n \leqslant x$ satisfies a Siegel-Walfisz condition. For any $A > 0$ there exists $B = B(A) > 0$ such that*

$$\sum_{q\leqslant Q} \sum_{a:\ (a,q)=1} \left| \sum_{n\equiv a \pmod q} \beta(n) - \frac{1}{\phi(q)} \sum_{(n,q)=1} \beta(n) \right|^2 \ll \|\beta\|^2 \frac{x}{(\log x)^A}$$

where $Q = x/(\log x)^B$.

The analogous result for $\Lambda(n)$ is known as the *Barban-Davenport-Halberstam theorem* and in that special case one can even obtain an asymptotic.

Before proceeding, let us assume, for the rest of this article, that we are given two sequences of complex numbers as follows:

[11] Their condition appears to be weaker than that assumed here, but it can be shown to be equivalent.

- $\alpha(m),\ M < m \leqslant 2M$ and $\beta(n),\ N < n \leqslant 2N$, with $x^{1/3} < N \leqslant M \leqslant x^{2/3}$ and $MN \leqslant x$.
- $\beta(n)$ satisfies the Siegel-Walfisz condition.
- $\alpha(m) \ll \tau(m)^A(\log x)^B$ and $\beta(n) \ll \tau(n)^A(\log x)^B$ (there inequalities are satisfied by $\mu, 1, \Lambda, L$ and any convolutions of these sequences).

In their second result, Bombieri, Friedlander and Iwaniec, showed that rather general convolutions are well-distributed[12] for *all* arithmetic progressions $a \pmod q$, for *almost all* $q \leqslant x^{1/2}/(\log x)^B$.

Theorem 8.2. *Suppose that $\alpha(m)$ and $\beta(n)$ are as above. For any $A > 0$ there exists $B = B(A) > 0$ such that*

$$\sum_{q\leqslant Q} \max_{a:\ (a,q)=1} \left| \sum_{n\equiv a \pmod q} (\alpha*\beta)(n) - \frac{1}{\phi(q)} \sum_{(n,q)=1} (\alpha*\beta)(n) \right| \ll \|\alpha\|\|\beta\| \frac{x^{1/2}}{(\log x)^A}$$

where $Q = x^{1/2}/(\log x)^B$.

This allowed them to give a proof of the Bombieri-Vinogradov theorem for primes, using Vaughan's identity (7.1), that seems to be less dependent on very specific properties of the primes. The subject, though, has long been stuck with the bound $x^{1/2}$ on the moduli.[13]

Bombieri, Friedlander and Iwaniec [3] made the following conjecture, and noted that in many applications, it suffices to work with a fixed (as is true in the application here).

Conjecture 8.3. *Suppose that $\alpha(m)$ and $\beta(n)$ are as above. For any $A, \epsilon > 0$, and every integer a, we have*

$$\sum_{\substack{q\leqslant Q\\ (q,a)=1}} \left| \sum_{n\equiv a \pmod q} (\alpha*\beta)(n) - \frac{1}{\phi(q)} \sum_{(n,q)=1} (\alpha*\beta)(n) \right| \ll \|\alpha\|\|\beta\| \frac{x^{1/2}}{(\log x)^A}$$

where $Q = x^{1-\epsilon}$.

The extraordinary work of Zhang breaks through the $\sqrt{x}$ barrier in some generality, working with moduli slightly larger than $x^{1/2}$, though his moduli are y-smooth, with $y = x^\delta$. The key result is as follows:

[12] This possibility has its roots in a paper of Motohashi [31].

[13] There had been some partial progress with moduli $> x^{1/2}$, as in [4], but no upper bounds which "win" by an arbitrary power of $\log x$ (which is what is essential to applications).

Vdyadicrange

Theorem 8.4. *Suppose that $\alpha(m)$ and $\beta(n)$ are as above. There exist constants $\eta, \delta > 0$ such, for any $A > 0$, for any integer a,*

$$\sum_{\substack{q \leqslant Q \\ P(q) \leqslant x^\delta \\ (q,a)=1 \\ q \text{ squarefree}}} \left| \sum_{\substack{n \leqslant x \\ n \equiv a \pmod q}} (\alpha * \beta)(n) - \frac{1}{\phi(q)} \sum_{\substack{n \leqslant x \\ (n,q)=1}} (\alpha * \beta)(n) \right| \ll_A \|\alpha\| \|\beta\| \frac{x^{1/2}}{(\log x)^A}$$

where $Q = x^{1/2+\eta}$.

We then deduce the same result but now for α and β supported in $x^{1/3} < m, n \leqslant x^{2/3}$ with $mn \leqslant x$, by dissecting this range up into into dyadic ranges (that is, $M < m \leqslant 2M$ and $N < n \leqslant 2N$) and smaller ranges, as well as possible, and then carefully accounting for the (m, n) pairs missed.

8.2. **The deduction of the main theorem for primes.** We will bound each term that arises from Vaughan's identity, (7.1), with $U = V = x^{1/3}$, rewritten as

$$\Lambda = \Lambda_{<x^{1/3}} + \mu_{<x^{1/3}} * L - (\mu * \Lambda)_{<x^{1/3}} * 1_{\geqslant x^{2/3}} - \mu_{<x^{1/3}} * \Lambda_{<x^{1/3}} * 1_{<x^{2/3}} + \mu_{\geqslant x^{1/3}} * \Lambda_{\geqslant x^{1/3}} * 1.$$

The first term is acceptably small, simply by taking absolute values. For the second term we write $(\mu_{<x^{1/3}} * L)(n) = \sum_{um=n,\ u<x^{1/3}} \mu(u) \log m$, to obtain the difference

$$\sum_{\substack{u<x^{1/3} \\ (u,q)=1}} \mu(u) \left(\sum_{\substack{x/u<m\leqslant 2x/u \\ m\equiv a/u \pmod q}} \log m - \frac{1}{\phi(q)} \sum_{\substack{x/u<m\leqslant 2x/u \\ (m,q)=1}} \log m \right)$$

Writing $M = x/u$, the inner sum is the difference between the sum of $\log m$ in $(M, 2M]$ over an arithmetic progression $b \pmod q$ with $(b, q) = 1$, minus the average of such sums. Now if $n_- = [M/q]$ and $n_+ = [2M/q]$, then, since $\log q[m/q] < \log m < \log q([m/q] + 1)$, such a sum is $> \sum_{n_- \leqslant n \leqslant n_+ - 1} \log qn$ and is $< \sum_{n_-+1 \leqslant n \leqslant n_+ + 1} \log qn$. The difference between these bounds in $\ll \log M$, and hence this is our bound on the term in parentheses. Summing over u yields a bound that is acceptably small.

We deal with the third term, by the same argument as just above, since we obtain an inner sum of 1, over the values of m in an interval of an arithmetic progression; and then we obtain a bound that is acceptably small.

We are left to work with two sums of convolutions:

$$\sum_{\substack{mn \asymp x \\ mn \equiv a \pmod q}} (\mu_{<x^{1/3}} * \Lambda_{<x^{1/3}})(m) 1_{<x^{2/3}}(n) \quad \text{and} \quad \sum_{\substack{mn \asymp x \\ mn \equiv a \pmod q}} (\Lambda_{\geqslant x^{1/3}} * 1)(m) \mu_{\geqslant x^{1/3}}(n),$$

where $x^{1/3} \ll m, n \ll x^{2/3}$, and each convolution takes the form $\alpha(m)\beta(n)$ with $\alpha(m)$ and $\beta(n)$ as above. The result then follows from Zhang's result as discussed at the end of the last subsection.

8.3. **Further reductions.** We reduce Theorem BVdyadicrange 8.4 further. The first observation is that we can restrict our moduli to those with $< C \log\log x$ prime factors, for some large $C > 0$, since the moduli with more prime factors are rare and thus contribute little to the sum. Since the moduli are y-smooth, they can be factored as qr where $N/(yx^\epsilon) < r \leqslant N/x^\epsilon$. Since the modulus does not have a lot of prime factors, one can deduce that the smallest prime factor of q, denoted $p(q)$, is $\geqslant D_0 := x^{\epsilon/\log\log x}$. Hence we may also now assume

- $r \in (R, 2R]$ with $P(r) \leqslant y$ with $y := x^\delta$.
- $q \in (Q, 2Q]$ with $D_0 < p(q) \leqslant P(q) \leqslant y$.
- $N/(yx^\epsilon) < R \leqslant N/x^\epsilon$ and $x^{1/2}/(\log x)^B < QR \leqslant x^{1/2+\eta}$

In polymath8 [34], some gains are made by working instead with the full set of moduli that have this kind of convenient factorizations, rather than restrict attention just to those moduli which are y-smooth.

We begin by noting that

$$\sum_{n\equiv a \pmod{qr}} \gamma(n) - \frac{1}{\phi(qr)} \sum_{(n,qr)=1} \gamma(n) =$$

$$\sum_{n\equiv a \pmod{qr}} \gamma(n) - \frac{1}{\phi(q)} \sum_{\substack{(n,q)=1\\ n\equiv a \pmod r}} \gamma(n) + \frac{1}{\phi(q)}\left(\sum_{\substack{(n,q)=1\\ n\equiv a \pmod r}} \gamma(n) - \frac{1}{\phi(r)} \sum_{\substack{(n,q)=1\\ (n,r)=1}} \gamma(n)\right)$$

with $\gamma = \alpha * \beta$. We sum the absolute value of these terms, over the moduli $d \in [D, 2D]$, factored into qr as above. Since $\beta(n)$ satisfies the Siegel-Walfisz criterion, we can deduce that $\beta(n)1_{(n,q)=1}$ also satisfies it, and therefore Theorem BF12 8.2 is applicable for $\alpha(m) * \beta(n)1_{(n,q)=1}$; this allows us to bound the sum of the second terms here, suitably. Hence it remains to prove

$$\sum_{\substack{q\in[Q,2Q]\\ D_0<p(q)\leqslant P(q)\leqslant y}} \sum_{\substack{r\in[R,2R],\\ P(r)\leqslant y\\ qr \text{ squarefree}}} \left| \sum_{\substack{n\equiv a \pmod r\\ n\equiv b \pmod q}} (\alpha*\beta)(n) - \sum_{\substack{n\equiv a \pmod r\\ n\equiv b' \pmod q}} (\alpha*\beta)(n) \right| \ll_A \|\alpha\|\|\beta\| \frac{x^{1/2}}{(\log x)^A}, \tag{8.1}$$

straw-2

for any integers a, b, b' with $p(abb') > y$.

9. Removing the weights, and an unweighted arithmetic progression

prelim-red

9.1. **Removing the weights.** The sums in (straw-2 8.1) are complicated, and the innermost sum is over an unknown function $\alpha * \beta$. In this section we use Cauchy's inequality to "unfold" the sum, so as to remove the weight from the innermost sum:

In the left-hand side of (8.1) [straw-2] we replace the absolute value in the (q,r) term by a complex number $c_{q,r}$ of absolute value 1, to obtain, after a little re-arranging:

$$\sum_r \sum_m \alpha(m) \left(\sum_q \sum_{n:\ mn \equiv a \pmod r} c_{q,r} \beta(n) (1_{mn\equiv b \pmod q} - 1_{mn \equiv b' \pmod q}) \right).$$

By the Cauchy-Schwarz inequality the square of this is

$$\leqslant \sum_r \sum_m |\alpha(m)|^2 \leqslant R\|\alpha\|^2$$

times

$$\sum_r \sum_m \left| \sum_q \sum_{n:\ mn \equiv a \pmod r} c_{q,r} \beta(n) (1_{mn\equiv b \pmod q} - 1_{mn \equiv b' \pmod q}) \right|^2 . \tag{9.1}$$

sq

When we expand the square, we obtain the sum of four terms of the form

$$\begin{aligned} &\pm \sum_r \sum_m \sum_{q_1,q_2} \sum_{\substack{n_1,n_2 \\ mn_1 \equiv mn_2 \equiv a \pmod r}} c_{q_1,r}\overline{c_{q_2,r}} \beta(n_1)\overline{\beta(n_2)} 1_{mn_1\equiv b_1 \pmod{q_1}} 1_{mn_2 \equiv b_2 \pmod{q_2}} \\ = &\pm \sum_r \sum_{q_1,q_2} \sum_{\substack{n_1,n_2 \\ n_1 \equiv n_2 \pmod r}} c_{q_1,r}\overline{c_{q_2,r}} \beta(n_1)\overline{\beta(n_2)} \cdot \sum_m 1_{\substack{m \equiv b_1/n_1 \pmod{q_1} \\ m\equiv b_2/n_2 \pmod{q_2} \\ m \equiv a/n_1 \pmod r}} \end{aligned} \tag{9.2}$$

TheUnfolde

where we get "+" when $b_1 = b_2 = b$ or b', and "−" otherwise, since $(mn, qr) = 1$.

We have achieved our goal of having an unweighted innermost sum. Indeed, if it is non-zero,[14] then it is just the number of integers in an interval of an arithmetic progression with common difference $r[q_1, q_2]$.

9.2. **The main terms.** The number of integers in an interval of length M, from an arithmetic progression with common difference $r[q_1, q_2]$ is

$$\frac{M}{r[q_1,q_2]} + O(1).$$

We study now the sum of the "main terms", the $M/r[q_1, q_2]$. Firstly, for the terms with $(q_1, q_2) = 1$ the main terms sum to

$$\pm \sum_r \sum_{\substack{q_1,q_2 \\ (q_1,q_2)=1}} \sum_{\substack{n_1,n_2 \\ n_1 \equiv n_2 \pmod r}} c_{q_1,r}\overline{c_{q_2,r}} \beta(n_1)\overline{\beta(n_2)} \cdot \frac{M}{rq_1q_2},$$

which is independent of the values of b_1, b_2 and hence cancel, when we sum over the four terms (and the two '+', and two '−', signs). For the terms with $(q_1, q_2) \neq 1$ we have $(q_1, q_2) \geqslant D_0$ (since the prime factors of the q_i are all $\geqslant D_0$), and it is not difficult to show that these are $\ll x(\log x)^{O(1)}/RD_0$, which is acceptably small.

[14] This sum cannot possibly contain any integers, and so is 0, if the congruences are incompatible. Since $(r, q_1q_2) = 1$ they are compatible unless $b_1/n_1 \equiv b_2/n_2 \pmod{(q_1, q_2)}$. Note that this criterion is irrelevant if $(q_1, q_2) = 1$.

TheUnfolded

9.3. **The error terms and the advent of exponential sums.** The "$O(1)$"s in (9.2) can add up to a total that is far too large. One can show that in most of the terms of the sum, the common difference of the arithmetic progression is larger than the length of the interval, so the correct count is either 0 or 1: It is hardly surprising that an error term of "$O(1)$" is too insensitive to help us.

To proceed, instead of approximating, we will give a *precise* formula for the number of integers in an arithmetic progression in an interval, using a sum of exponentials. By the Chinese Remainder Theorem, we can rewrite our triple of congruence conditions

$$m \equiv b_1/n_1 \pmod{q_1},\ m \equiv b_2/n_2 \pmod{q_2},\ m \equiv a/n_1 \pmod{r}$$

as one,

$$m \equiv m_0(n_1, n_2) \pmod{q}$$

where $q = rg\ell_1\ell_2$, when there is a solution, which happens if and only if $b_1/n_1 \equiv b_2/n_2 \pmod{g}$, where $g = (q_1, q_2)$ and we now define $\ell_1 = q_1/g,\ \ell_2 = q_2/g$.

To identify whether m is in a given interval I, we use Fourier analysis. The *discrete Fourier transform* is defined by

$$\hat{f}(h) := \sum_{b \pmod{q}} f(b) e_q(hb),$$

for any function f of period q. If f is any such function and $I(.)$ is the characteristic function for the interval $(M, 2M]$, then

$$\sum_{m \in I} f(m) = \frac{1}{q} \sum_{h \pmod{q}} \hat{I}(h) \hat{f}(-h), \tag{9.3}$$

Plancherel

is an example of Plancherel's formula. This has a "main term" at $h = 0$ (which is the same as the main term we found above, in that special case). The coefficients $\hat{I}(h)$ are easily evaluated and bounded:

$$\hat{I}(h) = \sum_{m=M+1}^{2M} e_q(hm) = e_q(2hM) \cdot \frac{e_q(hM) - 1}{e_q(h) - 1}.$$

The numerator has absolute value $\leqslant 2$ and, using the Taylor expansion, the denominator has absolute value $\asymp |h|/q$. Hence

$$|\hat{I}(h)| \ll \min\{M, q/|h|\},$$

We apply (9.3) with $f = \sum_i c_i 1_{m \equiv a_i \pmod{q}}$, take absolute values, and use our bounds for $|\hat{I}(h)|$, to obtain

$$\left| \sum_i c_i \left(\sum_{\substack{m \asymp M \\ m \equiv a_i \pmod{q}}} 1 - \frac{M}{q} \right) \right| \ll \sum_{\substack{0 \leqslant j \leqslant J \\ H_j := 2^j q/M}} \frac{1}{H_j} \sum_{1 \leqslant |h| \leqslant H_j} \left| \sum_i c_i e_q(a_i h) \right|. \tag{9.4}$$

ExponExpan

The error terms in (9.2) are bounded by

$$\sum_{r\asymp R}\sum_{g\leqslant G}\sum_{\substack{\ell_1,\ell_2\asymp Q/g\\(\ell_1,\ell_2)=1}}\left|\sum_{\substack{n_1,n_2\asymp N\\ n_1\equiv n_2 \pmod r\\ b_1/n_1\equiv b_2/n_2 \pmod g}}\beta(n_1)\overline{\beta(n_2)}\cdot\left(\sum_{\substack{m\asymp M\\ m\equiv m_0(n_1,n_2) \pmod{rg\ell_1\ell_2}}}1-\frac{M}{rg\ell_1\ell_2}\right)\right|$$

which, by (9.4), is

$$\ll\sum_{r\asymp R}\sum_{g\leqslant G}\sum_{\substack{\ell_1,\ell_2\asymp Q/g\\(\ell_1,\ell_2)=1}}\sum_{\substack{0\leqslant j\leqslant J\\ H_j:=2^jG/g}}\frac{1}{H_j}\sum_{1\leqslant|h|\leqslant H_j}\left|\sum_{\substack{n_1,n_2\asymp N\\ n_1\equiv n_2 \pmod r\\ n_2\equiv(b_2/b_1)n_1 \pmod g}}\beta(n_1)\overline{\beta(n_2)}e_{rg\ell_1\ell_2}(m_0(n_1,n_2)h)\right|.$$

We write $n_1=n,\ n_2=n+kr$, replace the n_2 variable with k, and define $m_k(n)=m_0(n_1,n_2)$. To simplify matters shall proceed with r,g,k and j fixed, and then sum over these at the end, so we are reduced to studying

$$\sum_{\substack{\ell_1,\ell_2\asymp L\\(\ell_1,\ell_2)=1}}\frac{1}{H}\sum_{1\leqslant|h|\leqslant H}\left|\sum_{\substack{n\asymp N\\(b_2-b_1)n\equiv b_1kr \pmod g}}\beta(n)\overline{\beta(n+kr)}e_{rg\ell_1\ell_2}(m_k(n)h)\right| \tag{9.5}$$

where $L=Q/g$.

10. Linnik's dispersion method

The proof of Zhang's Theorem, and indeed of all the results in the literature of this type, use Linnik's dispersion method. The idea is to express the fact that n belongs to an arithmetic progression using Fourier analysis; summing up over n gives us a main term plus a sum of exponential sums, and then the challenge is to bound each of these exponential sums.

Often the sums come with weights, and judicious use of Cauchying allows one to work with an unweighted, but more complicated exponential sum. We will discuss bounds on exponential sums later in this section. These exponential sums are often *Kloosterman sums*, which one needs to bound. Individual Kloosterman sums can often by suitably bounded by Weil's or Deligne's Theorem. However, sometimes one needs to get good bounds on averages of Kloosterman sums, a question that was brilliantly attacked by Deshouillers and Iwaniec [7], using the (difficult) spectral theory of automorphic forms. Indeed all previous work, breaking the $\sqrt{x}$ barrier, such as [12], [3]) uses these types of estimates. One of the remarkable aspects of Zhang's work is that he avoids these penible techniques, and the restrictions that come with them.

Zhang was able to use only existing bounds on Kloosterman sums to prove his Theorem, though he does use the sophisticated estimate of Birch and Bombieri from the appendix of [14]. Polymath8 indicates how even this deeper result can be avoided, so that the proof can be given using only “standard” estimates, which is what we do here.

10.1. **Removing the weights again.** To remove the β weights from (9.5), we begin by replacing the absolute value in (9.5) by the appropriate complex number c_{h,ℓ_1,ℓ_2} of absolute value 1, and re-organize to obtain

$$\sum_{\substack{n\asymp N\\ (b_2-b_1)n\equiv b_1kr \pmod g}} \beta(n)\overline{\beta(n+kr)} \sum_{\substack{\ell_1,\ell_2\asymp L\\ (\ell_1,\ell_2)=1}} \frac{1}{H}\sum_{1\leqslant |h|\leqslant H} c_{h,\ell_1,\ell_2} e_{rg\ell_1\ell_2}(m_k(n)h). \tag{10.1}$$

PreCauchy

We now Cauchy on the outer sum, which allows us to peel off the β's in the term

$$\sum_n |\beta(n)\beta(n+kr)|^2 \leqslant \sum_n |\beta(n)|^4 = \|\beta\|_4^4,$$

times the more interesting term

$$\sum_n \left| \sum_{\substack{\ell_1,\ell_2\asymp L\\ (\ell_1,\ell_2)=1}} \frac{1}{H}\sum_{1\leqslant |h|\leqslant H} c_{h,\ell_1,\ell_2} e_{rg\ell_1\ell_2}(m_k(n)h) \right|^2 .$$

We simply expand this sum, and take absolute values for each fixed $h, j, \ell_1, \ell_2, m_1, m_2$, to obtain

$$\leqslant \frac{1}{H^2} \sum_{1\leqslant |h|,|j|\leqslant H_i} \sum_{\substack{\ell_1,\ell_2,m_1,m_2\asymp L\\ (\ell_1,\ell_2)=(m_1,m_2)=1}} \left| \sum_{\substack{n\asymp N\\ (b_2-b_1)n\equiv b_1kr \pmod g}} e_{rg\ell_1\ell_2}(m_k(n)h) e_{rgm_1m_2}(-m_k(n)j) \right| .$$

Finally we have pure exponential sums, albeit horribly complicated.

10.2. **Exponential sums with complicated moduli.** If $(r,s) = 1$ then there are integers a, b for which

$$ar + bs = 1.$$

Note that although there are infinitely many possibilities for the pair of integers a, b, the values of $a \pmod s$ and $b \mod r$ are uniquely defined. If we divide the previous equation by rs, and multiply by m, and then take $e(.)$ of both sides, we obtain

$$e_{rs}(m) = e_s(am)\cdot e_r(bm).$$

This allows us to write the exponential, in our last sum, explicitly. After some analysis, we find that our exponential sum take the form

$$\sum_{\substack{n\asymp N\\ n\equiv a \pmod q}} e_{d_1}\left(\frac{C_1}{n}\right) e_{d_2}\left(\frac{C_2}{n+kr}\right), \tag{10.2}$$

ExpSum

for some constants C_1, C_2 (where $d_1 = rg[\ell_1,\ell_2]$, $d_2 = [m_1,m_2]$ and q divides g) which depend on many variables but are independent of n. With a change of variable $n \to qn + a$ we transform this to another sum of the same shape but instead over all n in an interval.

10.3. **Exponential sums: From the incomplete to the complete.** We now have the sum of the exponential of a function of n, over the integers in an interval. There are typically many integers in this sum, so this is unlike what we encountered earlier (when we were summing 1). The terms of the sum are periodic of period dividing $[d_1, d_2]$ and it is not difficult to sum the terms over a complete period. Hence we can restrict our attention to "incomplete sums" where the sum does not include a complete period.

We can now employ (9.3) once more. The coefficients $\hat{I}(h)$ are well understood, but the $\hat{f}(h)$ now take the form

$$\sum_{n \pmod q} e_{d_1}\left(\frac{C_1}{n} + hn\right) e_{d_2}\left(\frac{C_2}{n+\Delta} + hn\right),$$

a "complete" exponential sum.

The trick here is that we can factor the exponential into its prime factor exponentials and then, by the Chinese Remainder Theorem, this sum *equals* the product over the primes p dividing q, of the same sum but now over $n \pmod p$ with the appropriate $e_p(*)$. Hence we have reduced this question to asking for good bounds on exponential sums of the form

$$\sum_{n \pmod p} e_p\left(\frac{a}{n} + \frac{b}{n+\Delta} + cn\right).$$

Here we omit values of n for which a denominator is 0. As long as this does not degenerate (for example, it would degenerate if $p|a, b, c$) then Weil's Theorem implies that this is $\leqslant \kappa p^{1/2}$, for some constant $\kappa > 0$. Therefore the complete sum over $n \pmod q$ is $\leqslant \kappa^{\nu(q)} q^{1/2}$. This in turn allows us to bound our incomplete sum (10.2), and to bound the term at the end of the previous section.

The calculations to put this into practice are onerous, and we shall omit these details here. At the end one finds that the bounds deduced are acceptably small if

$$x^{1/2} \geqslant N > x^{(2+\epsilon)/5}$$

where $\epsilon > 12\eta + 7\delta$. However this is not quite good enough, since we need to be able to take N as small as $x^{1/3}$.

We can try a modification of this proof, the most successful being where, before we Cauchy equation (10.1) we also fix the ℓ_1 variable. This variant allows us to extend our range to all

$$N > x^{\frac{1}{3}+\epsilon}$$

where $\epsilon > \frac{14}{3}\eta + \frac{7}{2}\delta$. We are very close to the exponent $\frac{1}{3}$, but it seems that we are destined to just fail.

11. Complete exponential sums: Combining information the Graham-Ringrose way

The "square-root cancellation" for incomplete exponential sums of the form $|\sum_n e_q(f(n))|$ for various moduli q, with the sum over n in an interval of length $N < q$ is not quite good enough to obtain our results.

Graham and Ringrose [17] proved that we can improve the (analogous) incomplete character sum bounds when q is smooth. Here we follow Polymath8 [34], who showed how to modify the Graham-Ringrose argument to incomplete exponential sums. This will allow us to reduce the size of N in the above argument and prove our result.

11.1. **Formulating the improved incomplete exponential sum result.**

For convenience we will write the entry of the exponential sum as $f(n)$, which should be thought of as taking the form $a/n + b/(n+\Delta) + cn$, though the argument is rather more general. We assume that $N < q$, so that the Weil bound gives

$$\left|\sum_n e_q(f(n))\right| \ll \tau(q)^A q^{1/2}. \tag{11.1}$$

weil1

for some constant A which depends only on the degree of f.

In what follows we will assume that q is factored as $q = q_1 q_2$, and we will deduce that

$$\left|\sum_n e_q(f(n))\right| \ll \left(q_1^{1/2} + q_2^{1/4}\right) \tau(q)^A (\log q) N^{1/2}. \tag{11.2}$$

vdc-1

If q is y-smooth then we let q_1 be the largest divisor of q that is $\leqslant (qy)^{1/3}$ so that it must be $> (q/y^2)^{1/3}$, and so $q_2 \leqslant (qy)^{2/3}$. Hence the last bound implies

$$\left|\sum_n e_q(f(n))\right| \ll \tau(q)^A (qy)^{1/6} (\log q) N^{1/2}.$$

It is this bound that we insert into the machinery of the previous section, and it allows use to extend our range to all

$$N > x^{\frac{3}{10}+\epsilon}$$

where ϵ is bounded below by a (positive) linear combination of η and δ. In order that we can stretch the range down to *all* $N > x^{\frac{1}{3}}$, this method requires that

$$162\eta + 90\delta < 1.$$

11.2. **Proof of (11.2).**

We may assume

$$q_1 \leqslant N \leqslant q_2$$

else if $N < q_1$ we have the trivial bound $\leqslant N < (q_1 N)^{1/2}$, and if $N > q_2$ then (11.1) implies the result since $q^{1/2} = (q_1 q_2)^{1/2} < (q_1 N)^{1/2}$.

The main idea will be to reduce our incomplete exponential sum mod q, to a sum of incomplete exponential sums mod q_2. Now

$$e_q(f(n+kq_1)) = e_{q_1}(f(n)/q_2)\ e_{q_2}(f(n+kq_1)/q_1)$$

so that, by a simple change of variable, we have

$$\sum_n e_q(f(n)) = \sum_n e_q(f(n+kq_1))) = \sum_n\ e_{q_1}(f(n)/q_2)\ e_{q_2}(f(n+kq_1)/q_1).$$

Now, if we sum this over all $k, 1 \leqslant k \leqslant K := \lfloor N/q_1 \rfloor$, then we have

$$K\sum_n e_q(f(n)) = \sum_n e_{q_1}(f(n)/q_2)\ \sum_{k=1}^{K} e_{q_2}(f(n+kq_1)/q_1),$$

and so

$$\begin{aligned}\left|K\sum_n e_q(f(n))\right|^2 &\leqslant \left(\sum_n \left|\sum_{k=1}^{K} e_{q_2}(f(n+kq_1)/q_1)\right|\right)^2 \\ &\ll N\sum_n \left|\sum_{k=1}^{K} e_{q_2}(f(n+kq_1)/q_1)\right|^2 \\ &= N\sum_{1\leqslant k,k'\leqslant K}\ \sum_n e_{q_2}(g_{k,k'}(n)),\end{aligned}$$

where $g_{k,k'}(n) := (f(n+kq_1) - f(n+k'q_1))/q_1 \pmod{q_2}$ if $n+kq_1,\ n+k'q_1 \in I$, and $g_{k,k'}(n) := 0$ otherwise. If $k = k'$ then $g_{k,k}(n) = 0$, and so these terms contribute $\leqslant KN^2$.

We now apply the bound of (11.1) taking $f = g_{k,k}$ for $k \neq k'$. Calculating the sum yields (11.2).

11.3. **Better results.** In [34] the authors obtain better results using somewhat deeper techniques.

By replacing the set of y-smooth integers by the much larger class of integers with divisors in a pre-specified interval (and such that those divisors have divisors in a different pre-specified interval, etc., since one can iterate the proof in the previous section) they improve the restriction to

$$84\eta + 48\delta < 1.$$

Following Zhang they also gained bounds on certain higher order convolutions (of the shape $\alpha * 1 * 1 * 1$), though here needing deeper exponential sum estimates, and were then able to improve the restriction to (slightly better than)

$$43\eta + 27\delta < 1.$$

11.4. **Final remark.** It is worth noting that one can obtain the same quality of results only assuming a bound $\ll p^{2/3-\epsilon}$ for the relevant exponential sums in finite fields.

References

[1] E. Bombieri, *On the large sieve*, Mathematatika **12** (1965), 201–225.

[2] E. Bombieri and H. Davenport *Small difference between prime numbers*, Proc. Roy. Soc. Ser. A **293** (1966), 1-18.

[3] E. Bombieri, J. Friedlander and H. Iwaniec*Primes in arithmetic progressions to large moduli*, Acta Math. **156** (1986), no. 3-4, 203–251.

[4] E. Bombieri, J. Friedlander and H. Iwaniec, *Primes in arithmetic progressions to large moduli. II*, Math. Ann. **277** (1987), no. 3, 361–393.

[5] E. Bombieri, J. Friedlander and H. Iwaniec, *Primes in arithmetic progressions to large moduli. III*, J. Amer. Math. Soc. **2** (1989), no. 2, 215–224.

[6] P. Deligne, *La conjecture de Weil. II*, Publications Mathématiques de l'IHÉS **52** (1980), 137–252.

[7] J.-M. Deshouillers and H. Iwaniec, *Kloosterman Sums and Fourier Coefficients of Cusp Forms*, Inventiones mathematicae **70** (1982/83), 219-219.

[8] P. D. T. A. Elliott and H. Halberstam, *A conjecture in prime number theory*, Symp. Math. **4** (1968), 59–72.

[9] E. Fouvry, *A new form of the error term in the linear sieve*, Acta Arith., **37** (1980), 307–320.

[10] E. Fouvry, *Autour du théorème de Bombieri-Vinogradov*, Acta Math. **152** (1984), no. 3-4, 219–244.

[11] E. Fouvry and H. Iwaniec, *On a theorem of Bombieri-Vinogradov type*, Mathematika **27** (1980), no. 2, 135–152 (1981).

[12] E. Fouvry and H. Iwaniec, *Primes in arithmetic progressions*, Acta Arith. **42** (1983), no. 2, 197–218.

[13] J. Friedlander and A. Granville, *Limitations to the equi-distribution of primes. I*, Ann. of Math. **129** (1989), 363-382.

[14] J. Friedlander and H. Iwaniec, *Incomplete Kloosterman sums and a divisor problem*, With an appendix by Bryan J. Birch and Enrico Bombieri. Ann. of Math. (2) **121** (1985), no. 2, 319–350.

[15] D. Goldston, J. Pintz and C. Yıldırım, *Primes in tuples. I*, Ann. of Math. **170** (2009), no. 2, 819–862.

[16] D. Goldston, S. Graham, J. Pintz and C. Yıldırım, *Small gaps between primes or almost primes*, Trans. Amer. Math. Soc. **361** (2009), no. 10, 5285–5330.

[17] S. W. Graham and C. J. Ringrose, *Lower bounds for least quadratic nonresidues*, Analytic number theory (Allerton Park, IL, 1989), 269–309, Progr. Math., 85, Birkhäuser Boston, Boston, MA, 1990.

[18] A. Granville and K. Soundararajan, *Multiplicative number theory; the pretentious approach*, to appear.

[19] B.J. Green and T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Annals of Mathematics **167** (2008), 481-547.

[20] B.J. Green, T. Tao and T. Ziegler*An inverse theorem for the Gowers $U^{s+1}[N]$-norm*, Annals of Mathematics **176** (2012), 1231-1372.

[21] G. H. Hardy and J. E. Littlewood, *Some problems of "Partitio Numerorum", III: On the expression of a number as a sum of primes*, Acta Math. **44** (1923), 1–70.

[22] D. R. Heath-Brown, *Prime numbers in short intervals and a generalized Vaughan identity*, Canad. J. Math. 34 (1982), no. 6, 1365–1377.

[23] H. A. Helfgott, *Major arcs for Goldbach's theorem*, to appear.

[24] D. Hensley and I. Richards, *On the incompatibility of two conjectures concerning primes*, Analytic number theory (Proc. Sympos. Pure Math., Vol. XXIV, St. Louis Univ., St. Louis, Mo., 1972), pp. 123–127. Amer. Math. Soc., Providence, R.I., 1973.

[25] D. Hensley and I. Richards, *Primes in intervals*, Acta Arith. **25** (1973/74), 375–391.

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

[27] H. D. Kloosterman, *On the representation of numbers in the form $ax^2 + by^2 + cz^2 + dt^2$*, Acta Mathematica **49** (1926), pp. 407–464.

[28] H. Maier, *Small differences between prime numbers*, Michigan Math. J. **35** (1988), 323344.

[29] J. Maynard, *Small gaps between primes*, preprint.

[30] L. J. Mordell, *On a sum analogous to a Gauss's sum*, Quart. J. Math. Oxford Ser. **3** (1932), 161–167.
[31] Y. Motohashi, *An induction principle for the generalization of Bombieri's Prime Number Theorem*, Proc. Japan.. Acad. **52** (1976) 273-275.
[32] Y. Motohashi and J. Pintz, *A smoothed GPY sieve*, Bull. Lond. Math. Soc. **40** (2008), no. 2, 298–310.
[33] J. Pintz, *Polignac Numbers, Conjectures of Erdős on Gaps between Primes, Arithmetic Progressions in Primes, and the Bounded Gap Conjecture*, preprint.
[34] D.H.J. Polymath, *A new bound for small gaps between primes*, preprint.
[35] A. Schinzel, *Remarks on the paper "Sur certaines hypothéses concernant les nombres premiers"*, Acta Arith. **7** (1961/1962) 1–8.
[36] A. Selberg, *On elementary methods in prime number-theory and their limitations*, in Proc. 11th Scand. Math. Cong. Trondheim (1949), Collected Works, Vol. I, 388397, Springer-Verlag, Berlin-Göttingen-Heidelberg, 1989.
[37] P. Shiu, *A Brun-Titchmarsh theorem for multiplicative functions*, J. Reine Angew. Math. 313(1980), 161–170.
[38] K. Soundararajan, *Small gaps between prime numbers: the work of Goldston-Pintz-Yildirim*, Bull. Amer. Math. Soc. (N.S.) **44** (2007), no. 1, 1–18.
[39] T. Tao, *private communication.*
[40] R. C. Vaughan, *Sommes trigonométriques sur les nombres premiers*, C. R. Acad. Sci. Paris Sér. A **285** (1977), 981–983.
[41] A. I. Vinogradov, *The density hypothesis for Dirichlet L-series*, Izv. Akad. Nauk SSSR Ser. Mat. **29** (1965), 903–934.
[42] A. Weil, *Numbers of solutions of equations in finite fields*, Bulletin of the American Mathematical Society **55** (1949), 497–508.
[43] Y. Zhang, *Bounded gaps between primes*, to appear, Annals of Mathematics.

Département de mathématiques et de statistiques, Université de Montréal, Montréal QC H3C 3J7, Canada.

E-mail address: andrew@dms.umontreal.ca