On Primes in Quadratic Progressions

Joint Work with Stephan Baier

Liangyi Zhao

Nanyang Technological University

It is attributed to Dirichlet that any linear polynomial with integer coefficients represents infinitely many primes provided the coefficients are co-prime. G. H. Hardy and J. E. Littlewood gave the following conjecture in 1922 based on their circle method.

Conjecture 1. *Suppose $a, b, c \in \mathbb{Z}$ with $a > 0$, $\gcd(a, b, c) = 1$, $a + b$ and c are not both even, and $D = b^2 - 4ac$ is not a square. Let $P_f(x)$ be the number of primes $p \leq x$ of the form $p = f(n) = an^2 + bn + c$ with $n \in \mathbb{Z}$. Then*

$$P_f(x) \sim \gcd(2, a+b) \frac{\mathfrak{S}(D)}{\sqrt{a}} \frac{\sqrt{x}}{\log x} \prod_{\substack{p|a, p|b \\ p>2}} \frac{p}{p-1}, \tag{1}$$

where

$$\mathfrak{S}(D) = \prod_{\substack{p \nmid a \\ p>2}} \left(1 - \frac{\left(\frac{D}{p}\right)}{p-1}\right). \tag{2}$$

Here and after, $\left(\frac{D}{p}\right)$ denotes the Legendre symbol.

$$\left(\frac{D}{p}\right) = \begin{cases} 1, & \text{if } D \equiv \square \not\equiv 0 \pmod{p} \\ -1, & \text{if } D \not\equiv \square \pmod{p} \\ 0, & \text{if } p|D. \end{cases}$$

The conjecture has resisted attack to the extent that its simplest case for n^2+1 is not even resolved. Indeed, no polynomial of degree two or higher is known to represent infinitely many primes.

For polynomials of higher degrees, Hypothesis H of A. Schinzel and W. Sierpiński gives that if f is an irreducible polynomial with integer coefficients and there is no prime p such that $f(n) \equiv 0 \pmod{p}$ for every $n \in \mathbb{N}$, then f represents infinitely many primes. P. T. Bateman and R. A. Horn gave a more explicit version, with an asymptotic formula, of Hypothesis H.

A. Granville and R. A. Mollin studied $P_f(x)$, where f belongs to some family of quadratic polynomials.

1. Majorant predicted by (1), unconditionally uniform in f, and uniform in x under the Riemann hypothesis for the Dirichlet L-function $L(s,(D/\cdot))$.

2. For large R and N with $R^{\varepsilon} < N < \sqrt{R}$,

$$\#\left\{n \leq N : n^2 + n + A \in \mathbb{P}\right\} \asymp L\left(1, \left(\frac{1-4A}{\cdot}\right)\right)^{-1} \frac{N}{\log N}$$

   holds for a positive proportion of integers A with $R < A < 2R$.

3. An asymptotic formula for $P_f(x)$ holds for x in some range under the assumption of the existence of a Siegel zero for the relevant Dirichlet L-function.

Let E be an elliptic curve over $\mathbb{Q}$ with good reduction at a prime p. A theorem of H. Hasse states that

$$\#E_p = p + 1 - \lambda_E(p), \text{ with } |\lambda_E(p)| < 2\sqrt{p}.$$

The Lang-Trotter conjecture predicts an asymptotic formula for the number of primes $p \leq x$ such that $\lambda_E(p)$ equals a fixed integer r. If E has "complex multiplication" and $r \neq 0$, then the primes p satisfying $\lambda_E(p) = r$ lie in quadratic progression.

For example, the endomorphism ring of the elliptic curve $E : y^2 = x^3 - x$ is $\mathbb{Z}[i]$. It turns out that $p = n^2 + 1$ for some integer n if and only if $\lambda_E(p) = \pm 2$.

In cases like this, (1) would follow from Lang-Trotter conjecture.

The von Mangoldt function $\Lambda(n)$, the usual weight with which primes are counted, is defined as follows.

$$\Lambda(n) = \begin{cases} \log p, & \text{if } n = p^l \text{ for some } p \in \mathbb{P} \text{ and } l \in \mathbb{N}, \\ 0, & \text{otherwise.} \end{cases}$$

For $n^2 + k$ for with some fixed $k \in \mathbb{N}$ together with the weight of Λ, the conjecture (1) takes the following simpler form.

$$\sum_{n \le x} \Lambda(n^2 + k) \sim \mathfrak{S}(-4k)x.$$

Theorem 1. *[Baier, Z, 2007] Suppose that $z \ge 3$. Given $B > 0$, we have, for $z^{1/2+\varepsilon} \le K \le z/2$,*

$$\sum_{1 \le k \le K} \left| \sum_{z < n^2 + k \le 2z} \Lambda(n^2 + k) - \mathfrak{S}(-4k) \sum_{z < n^2 + k \le 2z} 1 \right|^2 \ll \frac{Kz}{(\log z)^B}.$$

From Theorem 1, the following corollary can be deduced immediately.

Corollary 1. *[Baier, Z, 2007] Given $A, B > 0$ and $\mathfrak{S}(k)$ as defined above, we have, for $z^{1/2+\varepsilon} \leq K \leq z/2$, that*

$$\sum_{z<n^2+k\leq 2z} \Lambda(n^2+k) = \mathfrak{S}(-4k) \sum_{z<n^2+k\leq 2z} 1 + O\left(\frac{\sqrt{z}}{(\log z)^B}\right)$$

holds for all natural numbers k not exceeding K with at most $O\left(K(\log z)^{-A}\right)$ exceptions.

It can be easily shown that $\mathfrak{S}(-4k)$ converges and

$$\mathfrak{S}(-4k) \gg \frac{1}{\log k} \gg \frac{1}{\log K} \gg \frac{1}{\log z}.$$

The above inequality shows that the main terms above are indeed dominating for the k's under consideration if $B > 1$ and that we truly have an "almost all" result.

The following sharpened version of Theorem 1 for short segments of quadratic progressions on average was proved.

Theorem 2. *[Baier, Z. 2007] If* $z \geq 3$, $z^{2/3+\varepsilon} \leq \Delta \leq z^{1-\varepsilon}$, $z^{1/2+\varepsilon} \leq K \leq z/2$ *and* $B > 0$, *then*

$$\int_z^{2z} \sum_{1 \leq k \leq K} |S_\Lambda - \mathfrak{S}(-4k) S_1|^2 \, \mathrm{d}t \ll \frac{\Delta^2 K}{(\log z)^B},$$

where

$$S_\Lambda = \sum_{t < n^2 + k \leq t + \Delta} \Lambda(n^2 + k), \text{ and } S_1 = \sum_{t < n^2 + k \leq t + \Delta} 1.$$

Under GRH for Dirichlet L-functions, the Δ-range in Theorem 2 can be extended to $z^{1/2+\varepsilon} \leq \Delta \leq z^{1-\varepsilon}$. Theorem 2 gives that

$$\sum_{t < n^2 + k \leq t + \Delta} \Lambda(n^2 + k) \sim \mathfrak{S}(-4k) \sum_{t < n^2 + k \leq t + \Delta} 1$$

holds for almost all k and t in the indicated ranges.

These results improve some earlier results of the authors where we used the circle method together with some lemmas in harmonic analysis due to P. X. Gallagher and H. Mikawa and the large sieve for real characters of D. R. Heath-Brown.

In the earlier work, k was restricted to be square-free and K can only be in the much smaller range of $z(\log z)^{-A} \leq K \leq z/2$. Our approach in the proof of Theorem 2 is a variant of the dispersion method of J. V. Linnik.

Expanding the modulus square in Theorem 2, three terms arise. Asymptotic formulas are established on average for all three and when combined the main terms cancel giving the desired result.

Note that n^2+1 is a prime if and only if $n+i$ is a Gaussian prime. Hence the problem is equivalent to counting Gaussian primes on the line $\Im z = 1$. Therefore, the problem can be approximated by counting Gaussian primes in narrow strips or sectors which can be studied using Hecke L-functions.

In this direction, C. Ankeny and P. Kubilius showed independently that under the Riemann hypothesis for Hecke L-functions for $\mathbb{Q}[i]$ there exist infinitely many Gaussian primes of the form $\pi = m+ni$ with $n < c\log|\pi|$, where c is some positive constant. From this, one infers the infinitude of primes of the form $p = m^2+n^2$ with $n < c\log p$.

Using sieve methods for $\mathbb{Z}[i]$, G. Harman and P. Lewis showed unconditionally that there exist infinitely many primes of the above form with $n \leq p^{0.119}$.

It can be seen that n^2+1 represents an infinitude of primes if and only if there are infinitely many primes p such that the fractional part of $\sqrt{p}$ is very small, namely $< 1/\sqrt{p}$.

A. Balog, G. Harman and S. Baier dealt with the following related question. Given $0 \leq \lambda \leq 1$ and a real number θ, for what positive numbers τ can one prove that there exist infinitely many primes p for which the inequality

$$\left\{p^\lambda - \theta\right\} < p^{-\tau}$$

is satisfied?

This problem in turn is related to estimating the number of primes of the form $[n^c]$, where $c > 1$ is fixed and n runs over the positive integers. Primes of this form are referred to as Pyateckiĭ-Šapiro primes.

It was established by C. Hooley that if D is not a perfect square then the greatest prime factor of $n^2 - D$ exceeds n^θ infinitely often if $\theta < \theta_0 = 1.1001\cdots$.

J.-M. Deshouillers and H. Iwaniec improved this to the effect that $n^2 + 1$ has infinitely often a prime factor greater than $n^{\theta_0 - \varepsilon}$, where $\theta_0 = 1.202\cdots$ satisfies $2 - \theta_0 - 2\log(2 - \theta_0) = \frac{5}{4}$. The improvement comes from utilizing mean-value estimates of Kloosterman sums of J.-M. Deshouillers and H. Iwaniec.

The result of Deshouillers and Iwaniec can also be generalized to $n^2 - D$ by Hooley's arguments.

Moreover, H. Iwaniec showed that there are infinitely many integers n such that n^2+1 is the product of at most two primes.

This result improved a previous one of P. Kuhn that n^2+1 is the product of at most three primes for infinitely many integers n and can be extended to any irreducible polynomial an^2+bn+c with $a>0$ and c odd.

J. B. Friedlander and H. Iwaniec, using results on half-dimensional sieve of H. Iwaniec, obtained lower bounds for the number of integers with no small prime divisors represented by a quadratic polynomial.

J. B. Friedlander and H. Iwaniec also proved the celebrated result that there exist infinitely many primes of the form $m^2 + n^4$ (with an asymptotic formula).

The set of integers of the form $m^2 + n^4$ contains the set of integers of the form $m^2 + 1$ but is still very sparse. The number of such integers not exceeding x is $O(x^{3/4})$. It is generally very difficult to detect primes in sparse sets.

We approximate the problem of representation of primes by $m^2 + 1$ in the following way. For a natural number n let $s(n)$ be the square-free kernel of n; i.e. $s(n) = n/m^2$, where m^2 is the largest square dividing n. We note that $s(n) = 1$ if and only if n is a perfect square. We consider primes of the form $n + 1$, where $s(n)$ is small.

Theorem 3. *[Baier, Z. 2006] Let $\varepsilon > 0$. Then there exist infinitely many primes p such that $s(p-1) \leq p^{5/9+\varepsilon}$.*

The set of natural numbers n with $s(n) \leq n^{5/9+\varepsilon}$ is also very sparse. More precisely, the number of $n \leq x$ with $s(n) \leq n^{5/9+\varepsilon}$ is $O(x^{7/9+\varepsilon/2})$.

We actually prove that the number of primes not exceeding x fitting the description in Theorem 3 is at least $cx^{7/9-\varepsilon}$ for some $c > 0$ and fixed. Theorem 3 can be reformulated as follows.

Theorem 3′. *Let $\varepsilon > 0$. Then there exist infinitely many primes of the form $p = am^2 + 1$ such that $a \leq p^{5/9+\varepsilon}$.*

Theorem 3 can be deduced from a Bombieri-Vinogradov type theorem for square moduli.

Theorem 4. *[Baier, Z. 2006] For any $\varepsilon > 0$ and fixed $A > 0$, we have*

$$(3) \qquad \sum_{q \le x^{2/9-\varepsilon}} q \max_{\substack{a \\ \gcd(a,q)=1}} \left| \psi(x; q^2, a) - \frac{x}{\varphi(q^2)} \right| \ll \frac{x}{(\log x)^A},$$

where

$$\psi(x; q, a) = \sum_{\substack{n \le x \\ n \equiv a \bmod q}} \Lambda(n)$$

and $\varphi(q)$ is the number of units in $\mathbb{Z}/q\mathbb{Z}$.

Theorem 4 improves some results of H. Mikawa and T. P. Peneva and P. D. T. A. Elliott. The key ingredient in the proof of Theorem 4 is the large sieve for square moduli which was studied both independently and jointly by the speaker and S. Baier.

The classical Bombieri-Vinogradov theorem gives

$$\sum_{q \le \sqrt{x}/(\log x)^{A+5}} \max_{\substack{a \\ \gcd(a,q)=1}} \left| \psi(x; q, a) - \frac{x}{\varphi(q)} \right| \ll \frac{x}{(\log x)^A}.$$

Hence the analogous statement for square moduli should have $q \le x^{1/4}(\log x)^{-A}$ in the sum over q in (3). Hence Theorem 4 is not the complete analogue of the classical theorem.

This is due to the fact that we established only results weaker than the expected analogue of the classical large sieve in the large sieve for square moduli.

The latter imperfection is caused by the fact that only a result weaker than the expected was established concerning the spacing of Farey fractions with square denominators.

If any of the above-mentioned expectations can be established (spacing of special Farey fractions, large sieve for square moduli or (3) with the extended range for q with $q \leq x^{1/4-\varepsilon}$), it would follow that there exist infinitely many primes p such that $s(p-1) \leq p^{1/2+\varepsilon}$.

We can get the same result under the assumption of the GRH for Dirichlet L-functions.

We note that the set of n such that $s(n) \leq n^{1/2+\varepsilon}$ is "almost" as sparse as the set of numbers m^2+n^4 considered by Friedlander and Iwaniec. Indeed, the number of $n \leq x$ such that $s(n) \leq n^{1/2+\varepsilon}$ is $O(x^{3/4+\varepsilon/2})$.

It is conceivable that an Elliott-Halberstam type hypothesis holds for primes in arithmetic progressions to square moduli, *i.e.*, that (3) holds with the exponent $1/2 - \varepsilon$ in place of $2/9 - \varepsilon$.

This would imply that there exist infinitely many primes p such that $s(p-1) \leq p^{\varepsilon}$.

A result of this kind comes very close to the conjecture that there exist infinitely many primes of the form $n^2 + 1$ since the number of $n \leq x$ such that $s(n) \leq n^{\varepsilon}$ is $O(x^{1/2+\varepsilon/2})$.