Research Statement: Frank Thorne

My research has been in classical analytic number theory. Specifically, my work has focused on modern techniques used to study the distribution of the primes, and on applications of these techniques beyond their traditional settings.

Let $\pi(n)$ denote the number of primes $\leq n$. The classical prime number theorem states that $\pi(n) \sim \frac{n}{\log n}$, and thus the probability that a randomly chosen integer n is prime is roughly $1/\log n$. It was suggested by Harald Cramér that the primes can be well modeled as independent random variables, and although contradictions have been found to random models (see Section 2 as well as the survey article of Soundararajan [16]), these models still give predictions which are widely believed to be accurate. For example, the Hardy-Littlewood prime k-tuple conjecture incorporates an arithmetical correction factor into Cramér's model, and predicts, for example, that the number of twin primes $TP(x)$ less than x is

$$TP(x) \sim 2\frac{x}{\log^2 x} \prod_{p \neq 2} \left(1 - \frac{1}{p}\right)^{-2} \left(1 - \frac{2}{p}\right). \tag{1}$$

This conjecture is widely believed and very well supported by numerical evidence.

Proofs of these conjectures seem to be hopelessly beyond reach. However, substantial progress has been made on the related problem of studying the distribution of the function

$$g(n) := \frac{p_{n+1} - p_n}{\log n}. \tag{2}$$

Certainly (1) implies that $g(n)$ should be arbitrarily close to 0 infinitely often, and this was recently proved in an Annals paper [3] by Goldston, Pintz, and Yıldırım, by the construction of an appropriate lower bound sieve. In Section 1, I will describe their work further, and I will discuss my own related work concerning small gaps between almost primes. In addition, I will discuss how my results yield applications to distribution questions concerning divisibility of class numbers, ranks of elliptic curves, and Fourier coefficients of modular forms.

I will also consider a related question on the distribution of primes in somewhat longer intervals. If $A > 2$, the probabilistic model predicts that the asymptotic

$$\frac{\pi(n + \log^A n) - \pi(n)}{\log^{A-1} n} \sim 1 \tag{3}$$

should hold as $n \to \infty$. Surprisingly, Maier [11] was able to use an ingenious matrix construction to disprove (3) for each A. Subsequent work of Balog, Friedlander, Granville, Soundararajan, Wooley, and others showed that Maier's result is not an isolated phenomenon, but rather that "unexpected irregularities" occur quite naturally in the distribution of the primes as well as more general arithmetic sequences. In Section 2, I will discuss these ideas further, and I will describe my work extending Maier's matrix method to number fields and function fields, where I was able to prove many of the analogous results.

1. Small Gaps Between Primes and Almost Primes

As mentioned in the introduction, Goldston, Pintz, and Yıldırım proved in a recent paper [3] that

$$\liminf_{n \to \infty} \frac{p_{n+1} - p_n}{\log n} = 0. \tag{4}$$

In follow-up work, Goldston, Graham, Pintz, and Yıldırım gave an alternate proof [4] of (4) based on the Selberg sieve. Furthermore, they showed in [4, 5] that their method was highly adaptable, and could also bound gaps between almost prime numbers. In particular, they proved that

$$\liminf_{n \to \infty} \big(q_{n+1} - q_n\big) \leq 6,$$

where q_n denotes the nth E_2 number (i.e., the nth square-free number with exactly two prime factors). They further obtained bounds between nonconsecutive E_2 numbers q_n and $q_{n+\nu}$ for any ν, or between E_2 numbers whose prime factors are both congruent to 1 modulo 4.

The reason for this adaptability is the following. The proofs in [3, 4, 5] proceed by considering a sum of the shape

$$S = \sum_{n=N}^{2N} \left(\sum_{h \in \mathcal{H}} \chi(n+h) - 1 \right) \left(\sum_{d \mid \prod_h (n+h)} \lambda_d \right)^2, \tag{5}$$

where $\chi(n)$ is the characteristic function of the primes or of a related sequence, $\mathcal{H}$ is a finite set of integers, and the λ_d are (real-valued) Selberg sieve coefficients, which have the property that the squared term is very small if $\prod_h (n+h)$ is divisible by many small primes. If $S > 0$ asymptotically for large N and a fixed choice of $\mathcal{H}$, then this argument proves the existence of bounded gaps between the integers counted by $\chi(n)$.

Therefore, one can prove the existence of bounded gaps in many sequences for which the sums originating from (5) can be estimated. In [19] I was able to prove the following rather general theorem:

Theorem 1. *Suppose $\mathcal{P}$ is a set of primes satisfying 'Condition BV' (to be described), ν is a positive integer, $r \geq 2$ is a positive integer, and q_n denotes the nth E_r number whose prime factors are all in $\mathcal{P}$. Then*

$$\liminf_{n \to \infty} (q_{n+\nu} - q_n) < C(r, \nu, \mathcal{P}),$$

for an explicit constant $C(r, \nu, \mathcal{P})$.

The precise statement of my result in [19] is somewhat more technical and involves a generalization to linear forms (originating in [5]), which I have not described here.

'Condition BV' refers to a condition similar to the Bombieri-Vinogradov theorem, which states that the primes of $\mathcal{P}$ must be well-distributed in arithmetic progressions on average. Work of Murty and Murty [12] implies that my theorem applies in particular to *Chebotarev* sets, i.e., sets of primes defined by the property that for some Galois extension $K/\mathbb{Q}$, their Frobenius elements lie in a fixed union of conjugacy classes of $\mathrm{Gal}(K/\mathbb{Q})$.

My work was largely motivated by results of Ono [13] and Soundararajan [15], which allowed me to prove the existence of bounded gaps between integers of more general arithmetic interest. As an example, fix an elliptic curve $E/\mathbb{Q}$ and its family of D-quadratic twists

$$E(D): \quad Dy^2 = x^3 + ax^2 + bx + c.$$

The $\mathbb{Q}$-rational points on each curve $E(D)$ form a finitely generated abelian group, and much work has been done to study the distribution of the ranks of $E(D)$ as D varies. One such result was obtained by Ono [13], who constructed an infinite family of D for which the rank of $E(D)$ is 0. These D are essentially products of primes in a certain Chebotarev set S_E, and using this construction I obtained the following theorem:

Theorem 2. *Let $E/\mathbb{Q}$ be an elliptic curve without a $\mathbb{Q}$-rational torsion point of order 2. Then there are a constant C_E and infinitely many pairs of square-free integers m and n with $|m - n| < C_E$, for which* $\mathrm{rk}(E(m)) = \mathrm{rk}(E(n)) = 0$.

I was also able to address a related question concerning divisibility of class groups of imaginary quadratic fields. In [15], Soundararajan proves that if d is a positive square-free integer $\equiv 1 \pmod 8$ whose prime factors are all congruent to $\pm 1 \pmod 8$, then the class group $\mathrm{Cl}(\mathbb{Q}(\sqrt{-d}))$ contains an element of order 4. Using this characterization I obtained the following result:

Corollary 3. *There are infinitely many pairs of E_2 numbers, say m and n, such that the class groups* $\mathrm{Cl}(\mathbb{Q}(\sqrt{-m}))$ *and* $\mathrm{Cl}(\mathbb{Q}(\sqrt{-n}))$ *each contain elements of order 4, with $|m - n| \leq 64$.*

This result is perhaps most interesting because the constant 64 is much smaller than might be expected. This constant corresponds to a certain 6-tuple of linear forms, and a careful analysis of the relation between

these linear forms and the E_2 numbers being counted shows that these numbers "appear" to have density $1/2$.

My results have further applications to the distribution of almost-prime ideals and to the distribution of nonzero Fourier coefficients of modular forms; see [19] for further details.

2. Maier Matrices and Unexpected Irregularities

As mentioned in the introduction, Maier [11] proved in 1985 that "unexpected" irregularities exist in the distribution of primes in short intervals. In particular he proved that for any $A > 2$ there exists a constant $\delta_A > 0$ such that

$$\limsup_{n\to\infty} \frac{\pi(n+\log^A n)-\pi(n)}{\log^{A-1} n} \geq 1+\delta_A, \quad \liminf_{n\to\infty} \frac{\pi(n+\log^A n)-\pi(n)}{\log^{A-1} n} \leq 1-\delta_A. \tag{6}$$

The proof is by the "Maier matrix" method, which I will describe very briefly here. (See the nice article by Granville [6] for a more thorough survey.) Let Q be a certain product of small primes, let x_1 and x_2 be integers with x_2 substantially larger than x_1, and let y be an integer with $y < Q$. Consider the following matrix of integers:

$$\begin{bmatrix} Qx_1+1 & Qx_1+2 & \dots & Qx_1+y \\ Q(x_1+1)+1 & Q(x_1+1)+2 & \dots & Q(x_1+1)+y \\ \vdots & \vdots & \vdots & \vdots \\ Qx_2+1 & Qx_2+2 & \dots & Qx_2+y \end{bmatrix}$$

The columns form arithmetic progressions modulo Q, and a theorem of Gallagher implies that if Q meets appropriate conditions on the associated Dirichlet L-functions, then each column will contain roughly the expected number of primes. Therefore, the number of primes in the matrix is determined by the number of integers $i \in [1, y]$ which are coprime to Q, and this is not necessarily asymptotic to $y\phi(Q)/Q$.

In related work, Shiu [14] used a variant of Maier's construction to prove that if a and m are integers with $(a, m) = 1$, then there exist arbitrarily long strings of consecutive primes which are all $\equiv a \pmod m$. For example, if $a = 1$, one constructs a matrix similar to the one above, where primes $\not\equiv 1 \pmod m$ are excluded from the product Q. Most primes in the matrix will then be $\equiv 1 \pmod m$, and Shiu's result easily follows.

In light of the well-known analogy between number fields and function fields, it is natural to ask whether similar results should hold in the polynomial ring $\mathbb{F}_q[t]$, for a fixed finite field $\mathbb{F}_q$. In [20], I adapted the Maier matrix method to $\mathbb{F}_q[t]$ and proved that the following similar results indeed hold:

Theorem 4. *For any fixed $A > 0$, there exists a constant $\delta_A > 0$ (depending also on q) such that*

$$\limsup_{k\to\infty} \sup_{\deg f = k} \frac{\pi(f, \lceil A\log k\rceil)}{q^{\lceil A\log k\rceil+1}/k} \geq 1+\delta_A, \quad \liminf_{k\to\infty} \inf_{\deg f = k} \frac{\pi(f, \lceil A\log k\rceil)}{q^{\lceil A\log k\rceil+1}/k} \leq 1-\delta_A.$$

Here $\pi(f, i)$ denotes the number of irreducible monic polynomials p with $\deg(f-p) \leq i$.

Theorem 5. *Suppose that k is a positive integer, and a and m are monic polynomials with $(a, m) = 1$. Then there exists a string of consecutive primes*

$$p_{r+1} \equiv p_{r+2} \equiv \cdots \equiv p_{r+k} \equiv a \pmod m.$$

Furthermore, for sufficiently large k, these primes may be chosen so that their common degree D satisfies

$$\frac{1}{\phi(m)}\left(\frac{\log D}{(\log\log D)^2}\right)^{1/\phi(m)} \ll k. \tag{7}$$

The implied constant depends only on q, and "consecutive" is to be understood with respect to lexicographic order.

Maier's matrix method may be similarly generalized to certain number fields. Let K be an imaginary quadratic field of class number 1, so that the ring of integers $\mathcal{O}_K$ naturally forms a lattice in $\mathbb{C}$. Generalizing Shiu's theorem, I have proved the existence of "bubbles" of congruent primes:

Theorem 6. *Suppose K is as above, k is a positive integer, and a and q are elements of $\mathcal{O}_K$ with $q \neq 2$ and $(a, q) = 1$. Then there exists a "bubble"*

$$B(r, x_0) = \{x \in \mathbb{C} : |x - x_0| < r\}$$

with at least k primes, all of which are congruent to ua modulo q for units $u \in \mathcal{O}_K$. Furthermore, for sufficiently large k, x_0 will satisfy

$$\frac{\omega_K}{\phi_K(q)} \left(\frac{\log\log|x_0| \log\log\log\log|x_0|}{(\log\log\log|x_0|)^2} \right)^{\omega_K/\phi_K(q)} \ll k. \tag{8}$$

Here ω_K denotes the number of units in $\mathcal{O}_K$, and $\phi_K(q) := |(\mathcal{O}_K/(q))^\times|$.

I am especially interested in a series of papers (see [6, 16] for overviews) which builds Maier's example (6) of "unexpected irregularities" into a general theory. In a 1989 paper [2], Friedlander and Granville proved that similar irregularities occur in estimating the number of primes $\leq x$ in arithmetic progressions to large moduli $\sim x \log^{-A} x$. Their result in particular contradicted the strong form of the Elliott-Halberstam conjecture. In related work, Balog and Wooley [1] showed that irregularities in short intervals of the form (6) are not restricted to the primes, and in particular occur in the distribution of sums of two squares.

In a 2007 Annals paper [8], Granville and Soundararajan developed these examples into a general "uncertainty" principle which governs the oscillation in the mean values of multiplicative functions. Suppose that $\mathcal{A}$ is an arbitrary set of integers having 'arithmetic structure'. Loosely speaking, this means that the frequency with which integers d divide elements of $\mathcal{A}$ is often substantially less than $1/d$. Granville and Soundararajan then proved that $\mathcal{A}$ *must* exhibit irregularities of distribution in short intervals, arithmetic progressions to large moduli, or both. As special cases they recovered all of the motivating examples of the previous paragraph.

Building on my work in [20], I have proved in [22] that Granville and Soundararajan's mechanism translates nicely to the polynomial ring $\mathbb{F}_q[t]$, and that analogues of essentially all of the same results hold. Moreover, in the $\mathbb{F}_q[t]$ case I have often been able to be more precise about where irregularities occur. For example, I was able to prove a stronger version of Theorem 4 where, among other things, I established that irregular short intervals can be found in every sufficiently large degree.

3. Directions for Further Work

I have several specific ideas for further work. One in particular was suggested by Granville, who asked whether my results on bounded gaps can be recast in a general axiomatic setting, along the lines of [8]. I have shown that any simple modification of the method will only give results for E_r or P_r numbers; i.e., if the method proves the existence of bounded gaps between numbers in a sequence $\mathcal{A}$, then it proves the same between the P_r numbers in $\mathcal{A}$ for some r. However, it is possible that a more substantial modification (and in particular, a choice of sieve weights substantially different from Selberg's) may yield interesting results.

But more broadly speaking, I look forward to studying more of the exciting developments that have recently occurred in analytic number theory. I will describe just a few that have interested me here. Green and Tao [10] used ergodic theory to examine the dichotomy between structure and randomness in the primes, and proved the amazing result that the primes contain arbitrarily long arithmetic progressions. Granville and Soundararajan, as well as some of their collaborators, have developed the idea of pretentiousness [7] in analytic number theory, and have used this to prove (among other results) a sharpened form of the Pólya-Vinogradov inequality for character sums for primitive characters of odd order. In perhaps more far-reaching work, Taylor [18] was able to prove the Sato-Tate conjecture for a large class of elliptic curves. Although this result is analytic, Taylor's proof essentially is not, and instead proceeds using sophisticated techniques from arithmetic geometry to show the automorphy of the symmetric power L-functions associated to E.

My intention is to learn many of these techniques as thoroughly as possible, with an eye towards applications which may be outside their traditional settings. For example, I am very interested in distribution questions concerning ranks of elliptic curves, Fourier coefficients of modular forms, reductions of elliptic curves and more general varieties over finite fields, and the parity of the partition function. Many such questions have been successfully addressed by analytic methods, and many more problems remain open. Questions such as these will continue to motivate both my research in analytic number theory as well as my study of subjects such as modular and automorphic forms, combinatorics, and arithmetic geometry.

References

[1] A. Balog and T. Wooley, *Sums of two squares in short intervals*, Canad. J. Math. **52** (2000), 673-694.
[2] J. B. Friedlander and A. Granville, *Limitations to the equi-distribution of primes I*, Ann. of Math. **129** (1989), 363-382.
[3] D. A. Goldston, J. Pintz, and C.Y. Yıldırım, *Primes in tuples I*, Ann. of Math., to appear.
[4] D. A. Goldston, S. W. Graham, J. Pintz, and C. Y. Yıldırım, *Small gaps between primes or almost primes*, preprint.
[5] D. A. Goldston, S. W. Graham, J. Pintz, and C. Y. Yıldırım, *Small gaps between products of two primes*, preprint.
[6] A. Granville, *Unexpected irregularities in the distribution of prime numbers*, Proceedings of the International Congress of Mathematicians (Zürich, 1994), 388-399, Birkhäuser, Basel, 1995.
[7] A. Granville, *Pretentiousness in the distribution of prime numbers*, notes from a lecture at Illinois Number Theory Fest, Urbana, IL, 2007.
[8] A. Granville and K. Soundararajan, *An uncertainty principle for arithmetic sequences*, Ann. of Math. **165** (2007), no. 2, 593-635.
[9] A. Granville and K. Soundararajan, *Large character sums: pretentious characters and the Pólya-Vinogradov theorem*, J. Amer. Math. Soc. **20** (2007), 357-384.
[10] B. Green and T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. of Math., to appear.
[11] H. Maier, *Primes in short intervals*, Michigan Math. J. **32** (1985), 221-225.
[12] M. R. Murty and V. K. Murty, *A variant of the Bombieri-Vinogradov theorem*, Canadian Math. Soc. Conf. Proc., Vol. 7 (1987), 243-272.
[13] K. Ono, *Nonvanishing of quadratic twists of modular L-functions and applications to elliptic curves*, J. reine angew. math., **533** (2001), 81-97.
[14] D. K. L. Shiu, *Strings of congruent primes*, J. London Math. Soc. **61** (2000), 359-373.
[15] K. Soundararajan, *Divisibility of class numbers of imaginary quadratic fields*, J. London Math. Soc., **61** (2000), 681-690.
[16] K. Soundararajan, *The distribution of prime numbers*, Equidistribution in number theory, an introduction, 59-83, NATO Sci. Ser. II Math. Phys. Chem. **237**, Springer, Dordrecht, 2007.
[17] K. Soundararajan, *Small gaps betwen prime numbers: the work of Goldston-Pintz-Yıldırım*, Bull. Amer. Math. Soc. **44** (2007), 1-18.
[18] R. Taylor, *Automorphy for some l-adic lifts of automorphic mod l representations. II*, preprint.
[19] F. Thorne, *Bounded gaps between products of primes with applications to ideal class groups and elliptic curves*, Int. Math. Res. Not., recommended for publication.
[20] F. Thorne, *Irregularities in the distribution of primes in function fields*, J. Number Theory, accepted for publication.
[21] F. Thorne, *Bubbles of congruent primes*, submitted.
[22] F. Thorne, *An uncertainty principle for function fields*, preprint.