Research Statement — Christopher Keyes

christopher.keyes@emory.edu
www.math.emory.edu/~ckeyes3

1. Overview

My research interests lie in **number theory**, but more specifically at the intersection of arithmetic geometry and arithmetic statistics. Briefly, **arithmetic geometry** involves studying or exploiting the geometry of equations defined over the rational numbers $\mathbb{Q}$ (or other fields of arithmetic interest such as number fields, global function fields, or finite fields). I interpret **arithmetic statistics** quite broadly to include quantitative questions involving objects or properties of interest to number theorists — prime numbers, rational points, or elliptic curves, to name a few.

To illustrate how these topics become entwined, consider a curve $C/\mathbb{Q}$ — a system of equations with rational coefficients of (complex) dimension one. Famously, **Faltings' theorem** states that if the geometric genus of C is at least 2, then the set of rational solutions $C(\mathbb{Q})$ is finite. In a sense, this is an arithmetic statistical result: finiteness is a quantitative statement about the rational points $C(\mathbb{Q})$.

This celebrated result leaves the door open for more questions. Enlarging our field of interest from $\mathbb{Q}$ to a number field $K/\mathbb{Q}$ of degree n, Faltings' theorem guarantees that the set of K-points $C(K)$ is finite, but it does not specify how many there are, nor does it tell us about the collection of all the **degree n points** of C. This brings us to the first of two broad questions around which my research interests are centered.

Question 1.1. *Given a curve $C/\mathbb{Q}$, what quantitative statements hold for the degree n points of C?*

Interpreted in this way, Faltings' theorem is a statement about the finiteness of degree 1 points on curves of genus at least 2. For $n > 1$, the smallest n for which C has infinitely many degree n points is known as the **arithmetic degree of irrationality** of $C/\mathbb{Q}$. Curves of arithmetic degree of irrationality 2 or 3 possess degree 2 or 3 maps to $\mathbb{P}^1$ or an ellptic curve E defined over $\mathbb{Q}$ [HS91, AH91], but this need not be the case more generally. [DF93].

To lead into the next broad question of interest, consider a curve C with a degree two map $C \to \mathbb{P}^1$, known as a **hyperelliptic curve**. Such a curve has infinitely many degree 2 points, arising geometrically from the pullback of rational points on $\mathbb{P}^1$ along the degree 2 map. Indeed, this provides a source of infinitely many points of any even degree, but how often does such a curve have any rational (or odd degree) points? Put much more generally:

Question 1.2. *Given some collection of curves over $\mathbb{Q}$, how many have a given property?*

In the case of the collection of hyperelliptics over $\mathbb{Q}$, suitably ordered by height, a landmark result of Bhargava–Gross–Wang [BGW17] is that a positive proportion have no odd degree points. In fact, the proportion of hyperelliptics with no $\mathbb{Q}$-points approaches 100% in the large genus limit [BGW17, Corollary 8].

Questions 1.1 and 1.2 guide my past, ongoing, and proposed research. In §2, I describe my work on the arithmetic of **superellptic curves**; this includes studying the proportion of superelliptics which are everywhere locally soluble [BK21b] and an asymptotic lower bound on the number of nonisomorphic field extensions of fixed degree arising from degree n points on a fixed such curve [Key22, BK21a]. In §3, I detail an ongoing project attempting to classify the **modular curves** $X_0(N)$ by their arithmetic degree of irrationality. In §4, I mention two other published works, on a generalization of **Mertens' product theorem** [APnKK22] and counting **arithmetical structures** on general graphs [KR21]. In §5 I describe the potential for undergraduate involvement in my research program.

2. Arithmetic of superelliptic curves

Let $m \geq 2$ be a positive integer and $m \mid d$. A **superelliptic curve** $C_f/\mathbb{Q}$ is given by an equation

$$C_f \colon y^m = f(x,z) = \sum_{i=0}^{d} c_i x^i z^{d-i}, \tag{2.1}$$

in the weighted projective space $\mathbb{P}\left(1, \frac{d}{m}, 1\right)$. These curves are notable for their degree m cover $C_f \to \mathbb{P}^1$ which has cyclic Galois group (after base change to a field containing the m-th roots of unity). When $m = 2$ in (2.1), C_f is known as a **hyperelliptic curve**.

2.1. Local solubility in families of curves.

In the spirit of Question 1.2, we ask how often a curve C_f given by (2.1) has a rational point. This turns out to be quite difficult, as the very first step in the approach of [BGW17] for the $m = 2$ case — determining how often C_f is everywhere locally soluble – was open for general m, so we begin there. A curve (or more generally a variety) $C/\mathbb{Q}$ is **everywhere locally soluble** if $C(\mathbb{Q}_p) \neq \emptyset$ for all primes p and $C(\mathbb{R}) \neq \emptyset$. Being everywhere locally soluble is necessary, but not sufficient, for C to have a $\mathbb{Q}$-point. If $C(\mathbb{Q}_p) = \emptyset$ (resp. $C(\mathbb{R}) = \emptyset$), we say C has a **local obstruction** at the place p (resp. the infinite place).

Poonen–Stoll [PS99b, PS99a] using the sieve of Ekedahl [Eke91] showed that the proportion of hyperelliptic curves over $\mathbb{Q}$ which are locally soluble is positive. Bhargava–Cremona–Fisher computed explicitly the proportion of locally soluble plane cubics ($\approx 75.96\%$) [BCF16] and genus one curves ($\approx 97.3\%$) [BCF21] by expressing the local densities as rational functions of p (in forthcoming work, they compute this quantity for hyperelliptic curves of genus $g > 1$).

In a joint work with Lea Beneish [BK21b], we study the proportion of superelliptic curves over $\mathbb{Q}$ which are everywhere locally soluble. To define this proportion, we take the limit

$$\rho_{m,d} = \lim_{B\to\infty} \frac{\#\{(c_i)_{i=0}^d \in \mathbb{Z}^d \mid C_f \text{ is everywhere locally soluble and } |c_i| \leq B \text{ for all } i\}}{\#\{(c_i)_{i=0}^d \in \mathbb{Z}^d \mid |c_i| \leq B \text{ for all } i\}},$$

where C_f is given by the equation (2.1).

Our first result shows that for fixed m, d, the proportion of locally soluble superelliptic curves $\rho_{m,d}$ is positive. The proof uses methods of Poonen–Stoll and Bright–Browning–Loughran [BBL16, PS99b, PS99a] and further shows that $\rho_{m,d}$ can be computed as a product of suitably defined local densities $\rho_{m,d}(p)$.

Theorem 2.1 (Beneish–K., [BK21b, Theorem A])**.** *For all $m \geq 2$ and d such that $m \mid d$ and $(m, d) \neq (2, 2)$, we have*

$$\rho_{m,d} = \rho_{m,d}(\infty) \prod_p \rho_{m,d}(p) > 0.$$

At a finite place p, the local density $\rho_{m,d}(p)$ is defined to be the normalized Haar measure of the subset of $(c_i)_{i=0}^d \in \mathbb{Z}_p^d$ for which C_f has a $\mathbb{Q}_p$-point. By studying the lifting of $\mathbb{F}_p$-points on the reduction $\overline{C_f}$ to $\mathbb{Q}_p$-points on C_f, we are in many cases able to give effective bounds for $\rho_{m,d}(p)$. These can be used to determine bounds for the limiting behavior as $d \to \infty$ for fixed m. In particular, for primes $m > 2$ as d ranges over multiples of m, we find that

$$\liminf_{d\to\infty} \rho_{m,d} \geq 0.83511 \quad \text{and} \quad \limsup_{d\to\infty} \rho_{m,d} \leq 1 - \frac{1}{2^9} \approx 0.99804 \quad \text{(see [BK21b, Theorem B])}.$$

We can think of the above as constraining the everywhere local solubility behavior in the **large genus limit**. In this limit (with m prime), between about 83% and 99.8% of superelliptic curves everywhere locally soluble.

In the special case of $(m, d) = (3, 6)$, a more careful investigation yields exact determinations of $\rho_{3,6}(p)$ for all primes p, which when combined with Theorem 2.1 leads to the following result.

Theorem 2.2 (Beneish–K., [BK21b, Theorem C])**.** *We have $\rho_{3,6} \approx 96.94\%$. More precisely, there exist rational functions $R_1(t)$ and $R_2(t)$ such that the local density $\rho_{3,6}(p)$ is given by*

$$\rho_{3,6}(p) = \begin{cases} R_1(p), & p \equiv 1 \pmod 3 \text{ and } p > 43 \\ R_2(p), & p \equiv 2 \pmod 3 \text{ and } p > 2. \end{cases}$$

The explicit formulae are given in [BK21b, (C.1)]. The asymptotic behavior of $R_i(t)$ are described by

$$1 - R_1(t) \sim \frac{2}{3} t^{-4} \quad \text{and} \quad 1 - R_2(t) \sim \frac{53}{144} t^{-7}.$$

The rational functions R_1 and R_2 have total degree 57 in p, much larger than in the genus 1 or plane cubic cases [BCF16, BCF21]. The computer algebra system Sage [Sag21] was used to compute R_1 and R_2 exactly, with the code available on Github at `https://github.com/c-keyes/Density-of-locally-soluble-SECs`. Considerable care and additional computations were needed to determine the local densities $\rho_{3,6}(2), \rho_{3,6}(3)$, and $\rho_{3,6}(p)$ for primes $p \equiv 1 \pmod 3$ such that $p \leq 43$.

2.2. Fields generated by points on curves.

Let $C/\mathbb{Q}$ be a curve (or more generally a variety) and fix an algebraic closure $\overline{\mathbb{Q}}$. We say a field $K/\mathbb{Q}$ is **generated by a point of** C if $K = \mathbb{Q}(P)$ for some $P \in C(\overline{\mathbb{Q}})$.

That is, K is the minimal field of definition for a degree n point of C. For $n \geq 1$ an integer and X a positive real number, we define the quantity $N_{n,C}(X)$ to be the number of such extensions with degree $[K : \mathbb{Q}] = n$ and bounded absolute discriminant $|\mathrm{Disc}(K)| \leq X$. We further take $N_{n,C}(X, G)$ to be the number of those extensions with $\mathrm{Gal}(\widetilde{K}/\mathbb{Q}) \simeq G$, where $\widetilde{K}$ denotes the Galois closure of K.

In their paper on **Diophantine Stability**, Mazur and Rubin [MR18] ask to what extent the set of fields generated by algebraic points determines the identity of the curve C. Motivated by this, and Question 1.1, we want to understand how $N_{n,C}(X)$ grows as $X \to \infty$, and how this asymptotic depends on both the geometry of C and the degree n. When C is an elliptic curve, Lemke Oliver and Thorne [LT19] show there are $\gg X^{c_n - \epsilon}$ number fields $K/\mathbb{Q}$ of degree $n \geq 2$ and discriminant at most X, such that the Mordell–Weil rank of $C(K)$ is greater than that of $C(\mathbb{Q})$, and C/K has specified root number. Here c_n is a positive constant and tends to $1/4$ from below as $n \to \infty$.

In successive papers [Key22] and [BK21a] (the second of which is joint with Lea Beneish), we consider the case where $C = C_f$ is a hyperelliptic, then superelliptic curve. For this application, it suffices to consider the affine equation,

$$C_f : y^m = f(x), \tag{2.2}$$

where $f(x) \in \mathbb{Q}[x]$ is m-th power free of degree d.

When $m = 2$ and C_f is a hyperelliptic curve, the main result of [Key22] is an asymptotic lower bound for $N_{n,C}(X, S_n)$ when n is large relative to d, generalizing that of Lemke Oliver and Thorne.

Theorem 2.3 (K., [Key22, Theorems 1.1, 1.2])**.** *Let C be a hyperelliptic curve with genus $g \geq 1$ given by an equation (2.2) with $m = 2$. If $n \geq d$ is divisible by $\gcd(d, 2)$ then*

$$N_{n,C}(X, S_n) \gg X^{c_n}$$

where c_n is a positive constant depending on g tending to $1/4$ as $n \to \infty$.

Moreover, the constant c_n can be determined explicitly. When $d = 2g + 1$ is odd we have

$$c_n = \frac{1}{4} - \frac{gn^2 - (g^2 - 2g - 3)n - 2g^2}{2n^2(n-1)}$$

unconditionally, and for sufficiently large n this improves to

$$c_n = \frac{1}{4} - \frac{gn + g^2 - 2g}{2n(n-1)},$$

with similar formulas in the case where d is even.

The cases of d odd and even are treated separately in the proof of Theorem 2.3. We recall that when d is even, the restriction to even degrees n is necessary, as a positive proportion of hyperelliptic curves over $\mathbb{Q}$ have no points defined over any extension of odd degree [BGW17].

The approach of the proof is to produce a family of polynomials whose roots give rise to points on C. We contrive this family to consist almost entirely of irreducible polynomials of the desired degree n with Galois group S_n, proving this with **Newton polygons**. We then count this family, adjusting for multiplicity of the fields generated, to give a lower bound for $N_{n,C}(X, S_n)$. Finally, to obtain the improved exponent c_n, we use the best known upper bounds on the number of fixed degree number fields with bounded discriminant due to Lemke Oliver and Thorne [LT22], valid for sufficiently large n.

In [BK21a] we extend results from [Key22] to superelliptic curves. Using similar methods, we succeed in giving lower bounds for $N_{n,C}(X)$ with explicit nonvanishing exponents depending on n, m, and d, albeit without the Galois group condition of Theorem 2.3.[1]

Theorem 2.4 (Beneish–K., [BK21a, Theorems 1, 2])**.** *Let C be a superelliptic curve with equation (2.2) and suppose n is a multiple of $\gcd(m, d)$ satisfying $n \geq \max(d, \mathrm{lcm}(m, d) - m - d + 1, 2m^2 - m)$. Then we have*

$$N_{n,C}(X) \gg X^{c_n}, \tag{2.3}$$

where c_n is an explicit constant depending on m, d, and n, with $c_n \to \frac{1}{m^2}$ as $n \to \infty$. The implied constant in (2.3) depends only on n and (the equation for) C.

Moreover, if C has a rational point, then for all sufficiently large degrees n with no divisibility requirements, (2.3) holds, c_n has an explicit description, and satisfies $c_n \to \frac{1}{m^2}$ as $n \to \infty$.

[1] Note that the publicly available preprint [BK21a] does not reflect the most up-to-date version as presented here.

Included in [BK21a, §7] is a discussion of whether and how often superelliptic curves given by (2.2) have points of degree n *not* divisible by $\gcd(m,d)$, and the possible geometric sources for such points. In particular, we highlight a new result that uses an argument of Gunther and Morrow [GM19, Proposition 2.6] which shows that 100% (with respect to an appropriate height) of genus g hyperelliptic curves with a Weierstrass rational point have finitely many points of degree $n < g$ that are not the pullback of a degree $n/2$ point on $\mathbb{P}^1$. Here we take the **height** of a degree d polynomial $f(x) = \sum_{i=0}^{d} c_i x^i$ to be $\mathrm{ht}(f) = \max\{|c_i|\}$.

Proposition 2.5 (Beneish–K., [BK21a, Proposition 7.2])**.** *Suppose m and d are positive even integers and k is an odd prime satisfying*

- $4 \mid m \mid d$*,*
- $m \leq k$*,*
- $N = 2k < \frac{d}{2} - 1$*.*

Then for a positive proportion of squarefree degree d polynomials $f(x)$ ordered by height, the superelliptic curve given by $C\colon y^m = f(x)$ has at most finitely points of degree N.

2.3. Ongoing and future work on superelliptic curves.

Theorems 2.1 and 2.2 suggest a number of follow up questions. Perhaps most tantalizing is that of global solubility.

Question 2.6. *What proportion of superelliptic curves C_f with $(m,d) = (3,6)$ have a $\mathbb{Q}$-point?*

Since we know how often such curves are everywhere locally soluble, this is equivalent to determining how often such curves fail the **Hasse principle**. In this vein, Browning [Bro17] computed such a proportion for a certain family of cubic surfaces, while Bhargava–Gross–Wang [BGW17] show that a positive proportion of hyperelliptic curves have no rational points.

As we have seen, [BGW17] goes a step further, showing that a positive proportion of hyperelliptic curves have no points over any odd degree extension of $\mathbb{Q}$. Equivalently, a positive proportion have **index** 2, where the index of a variety over $\mathbb{Q}$ is the greatest common divisor of the degrees $[\mathbb{Q}(P):\mathbb{Q}]$ of algebraic points P. A superelliptic curve C_f given by (2.1) has index at most m, since it always has points of degree m from pulling back rational points on $\mathbb{P}^1$ along the degree m map.

Question 2.7. *How often does a superelliptic curve C_f have index exactly m?*

We speculate that for a prime $m > 2$, something similar to the hyperelliptic case holds and propose to study this via the explicit **descent** methods for $\mathrm{Pic}^1(C_f)$ developed by Creutz [Cre13, Cre20]. A point on C_f of degree coprime to m gives rise to a degree 1 divisor class defined over $\mathbb{Q}$, i.e. an element of $\mathrm{Pic}^1(C_f)(\mathbb{Q})$. When $\mathrm{Pic}^1(C_f)(\mathbb{Q})$ is provably empty however, as is the case for e.g.

$$C_f\colon y^3 = 3(x^6 + x^4 + 4x^3 + 2x^2 + 4x + 3) \qquad\qquad \text{(see [Cre13, Example 7.3])},$$

we deduce that C_f has no points of degree prime to $m = 3$ and in particular no $\mathbb{Q}$-points. We hope that by sufficiently automating these descent computations and providing index data for many curves C_f, we will be able to make precise conjectures about how often curves of this form have higher degree points of nontrivial degree, and possibly prove them by modifying the descent machinery.

With a more complete grasp on if and/or when C_f has points at all of various degrees, we have several potential directions for inquiry. One is to attempt to generalize Proposition 2.5, to better understand how often C_f has only finitely many points of some fixed degree n *not* arising from the pullback along the superelliptic map of a degree n/m point on $\mathbb{P}^1$.

The hyperelliptic case gives reason to suggest that such generalizations can only hold for bounded degrees n. Consider that a point P arising from a degree $n/2$ point P' on $\mathbb{P}^1$ has a constrained Galois group, and in particular $G \neq S_n$. Theorem 2.3 implies that for n sufficiently large and even, a hyperelliptic curve C_f has infinitely many degree n points P with $G = S_n$, hence not arising from the pullback of such a P'.

We now return to the challenge of counting minimal fields of definition for these degree n points.

Question 2.8. *What can we say about $N_{n,C_f}(X,G)$ for various subgroups $G \subseteq S_n$?*

We expect that the lower bound of Theorem 2.4 holds for $G = S_n$, as in Theorem 2.3 in the hyperelliptic case. Little is known for proper subgroups $G \subsetneq S_n$, and it would be of interest to find such subgroups which occur either as often as S_n (in analogy to the prevalence of S_4- and D_4-quartic extensions $K/\mathbb{Q}$) or rarely in some appropriate sense. Another avenue is to explore the low degree n case, where we expect somewhat different behavior by the discussion above, and make more precise quantitative statements about $N_{n,C}(X, G)$ for various G. In the setting of Proposition 2.5 for example, we know that groups G not arising from a pullback can only occur finitely often, but it would be interesting to see if some cannot arise at all, or do not arise for most superelliptic curves C_f.

3. Higher degree points on modular curves

Let $n \geq 1$ be an integer. For which $N \geq 1$ do there exist infinitely many elliptic curves E/K with a K-rational torsion point (resp. cyclic N-isogeny), where $K/\mathbb{Q}$ is a number field of degree n? We can rephrase this into a question about degree n points on a **modular curve**.

Question 3.1. *For which N does the modular curve $X_1(N)$ (resp. $X_0(N)$) have infinitely many degree n points?*

The rational points on a modular curve X correspond to elliptic curves defined over $\mathbb{Q}$ with additional torsion data, along with finitely many rational cusps. For example, if $X_1(N)$ has a non-cuspidal rational point, then there exists an elliptic curve $E/\mathbb{Q}$ and a rational point $P \in E(\mathbb{Q})$ of exact order N. Faltings' theorem then implies that whenever $X_1(N)$ has genus at least 2, there are at most finitely many such pairs $(E/\mathbb{Q}, P)$. The proof of Mazur's famous theorem (which actually preceded Faltings' theorem) classifying the possible torsion subgroups over an elliptic curve over $\mathbb{Q}$ goes a step further, actually determining which modular curves have non-cuspidal rational points.

For higher n, this question is intimately related to the geometry of the modular curve. For instance, a curve X of genus $g > 1$ with a rational point has infinitely many quadratic points if and only if it has a degree 2 map to $\mathbb{P}^1_{\mathbb{Q}}$, or a degree 2 map to an elliptic curve $E/\mathbb{Q}$ of positive rank [HS91]. We call these situations **hyperelliptic** and **bielliptic**, respectively. The modular curves $X_0(N)$ which are hyperelliptic and bielliptic were classified by [Ogg74] and [Bar99], respectively.

We say a curve $X/\mathbb{Q}$ is **n-gonal** over $\mathbb{Q}$ if n is the minimal degree of a map $X \to \mathbb{P}^1$ defined over $\mathbb{Q}$. Similarly, X is **n-elliptic** over $\mathbb{Q}$ if n is the minimal degree of $X \to E$ defined over $\mathbb{Q}$, where E is a curve of genus 1. We can similarly define the gonality and ellipticity over $\overline{\mathbb{Q}}$, by allowing these maps to be defined over the algebraic closure instead.

We focus our attention on $X_0(N)$. Much effort has been made to determine the gonalities and ellipticities of these modular curves, with an eye to classifying they have infinitely many degree n points.

- Ogg, Hasegawa–Shimura, Jeon–Park, and most recently Najman–Orlić have determined the list of n-gonal $X_0(N)$ for $2 \leq n \leq 5$ [Ogg74, HS99, JP05, NO22].
- Bars and Jeon determined the 2- and 3-elliptic $X_0(N)$, respectively [Bar99, Jeo21].
- Several authors have also made progress in studying these properties for quotients of $X_0(N)$, namely by their Atkin–Lehner involutions [Has95, Has97, FH99, Jeo18, BGR19, BGK20].

Work of Harris–Silverman and Abramovich–Harris shows that for $n = 2$ or 3, having infinitely many degree n points is equivalent to having a degree n map to $\mathbb{P}^1_{\mathbb{Q}}$ or an elliptic curve $E/\mathbb{Q}$ of positive rank [HS91, AH91]. Thus the answer to Question 3.1 is known for $X_0(N)$ in the $n \leq 3$ cases.

We were working to complete the classification of levels N for which $X_0(N)$ is 4-gonal when Najman–Orlić posted their preprint [NO22] in July 2022. Building on these new results, we propose to study several remaining questions regarding $X_0(N)$ and its degree n points for $n \geq 4$.

Question 3.2. *Which $X_0(N)$ are n-elliptic over $\mathbb{Q}$ (or $\overline{\mathbb{Q}}$) for $n \geq 4$?*

When $n = 4$, there are some obvious candidates for 4-elliptic modular curves $X_0(N)$, namely those with a bielliptic Atkin–Lehner quotient $X_0(N)/w$, which were recently classified [Jeo18, BGR19, BKS22]. It would be interesting to determine whether there exist any 4-elliptic maps *not* factoring through such a quotient. The $n = 5$ case may prove somewhat easier since 5 is prime. For either choice of n, geometric tools, computation of the Jacobian deomposition, and finite field methods would likely all play a role.

Question 3.3. *Do there exist $X_0(N)$ with infinitely many degree n points, but possessing no maps of degree n to $\mathbb{P}^1_{\mathbb{Q}}$ or to an elliptic curve $E/\mathbb{Q}$?*

Consider $n = 4$ case, where work of Abramovich–Harris [AH91] implies that such curves must have genus 7. $X_0(97)$ is the only modular curve $X_0(N)$ of genus 7 to have gonality strictly greater than 4; it is geometrically 5-gonal, and 6-gonal over $\mathbb{Q}$ [NO22]. Moreover, there exist no elliptic curves $E/\mathbb{Q}$ of conductor 97, so we find no maps $X_0(97) \to E$ defined over $\mathbb{Q}$. Thus $X_0(97)$ cannot have infinitely many degree 4 points coming from pullback of rational points on a curve of genus at most 1.

However, a genus 7 curve $X/\mathbb{Q}$ may possess infinitely many degree 4 points if the image of $\mathrm{Sym}^4(X)$ in its Jacobian $J = \mathrm{Pic}^0(X)$ contains a (translate of a) positive-dimensional abelian variety. In the case of $X_0(97)$, its Jacobian has an isogeny decomposition

$$J_0(97) \sim A_1 \times A_2$$

where A_1, A_2 are simple of dimensions 3 and 4 respectively. These dimensions are large enough to rule out a positive-dimensional abelian translate in $\mathrm{Sym}^4(X_0(97))$ [DF93, Corollary 3.6], so Question 3.3 is answered for $n = 4$: the only modular curves with infinitely many degree 4 points are those which are 4-gonal or 4-elliptic.

The question remains open for $n > 4$. Debarre–Fahlaoui [DF93] gave constructions of curves with infinitely many points of degree $n \geq 4$, but lacking a degree n map to a curve of genus at most 1. Recent classification results of Kadets–Vogt [KV22] for $n \leq 5$ show that these and similar constructions are the only other way to obtain infinitely many degree n points on X (modulo some small genus cases). By investigating how modular curves $X_0(N)$ fit into this classification and when they come from Debarre–Fahlaoui constructions, we hope to answer Question 3.3 for values of $n \geq 5$.

4. Other topics

4.1. Generalization of Mertens' product theorem to Chebotarev sets.

Recall the celebrated **Mertens product theorem** [Mer74],

$$\prod_{p \leq x} \left(1 - \frac{1}{p}\right) \sim \frac{e^{-\gamma}}{\log x}, \qquad \text{as } x \to \infty, \tag{4.1}$$

where γ is the Euler constant. See, for example, [MV06, Theorem 2.7] for a modern discussion and proof. This theorem has been generalized appropriately to products over primes in number fields [Ros99], points on varieties over finite fields [Leb07], and primes in congruence classes [Wil74], [LZ07]. Some of these results are summarized in Table 1.

In a joint publication with Santiago Arango-Piñeros and Daniel Keliher [APnKK22], we generalize this further to **Chebotarev sets** of primes in a number field and give a power saving error term. More precisely, fix a Galois extension of number fields E/F with group G. If $C \subset G$ is a conjugacy class, let $\mathcal{C}(x)$ denote the set of unramified primes P in $\mathcal{O}_F$ with Artin symbol $\mathrm{Frob}_P = \left(\frac{E/F}{P}\right) = C$ and bounded norm $NP \leq x$.

Theorem 4.1 (Arango-Piñeros–Keliher–K., [APnKK22, Theorem A])**.** *With the notation as above, we have*

$$\prod_{P \in \mathcal{C}(x)} \left(1 - \frac{1}{NP}\right) = \left(\frac{e^{-\gamma(E/F,C)}}{\log x}\right)^{|C|/|G|} + O\left(\frac{1}{(\log x)^{|C|/|G|+1}}\right) \tag{4.2}$$

when $x \to \infty$, and the implied constant depends on the extension E/F and C. Furthermore, the constant $e^{-\gamma(E/F,C)}$ is given by

$$e^{-\gamma(E/F,C)} = e^{-\gamma_F} \prod_{P \in \Sigma_F} \left(1 - \frac{1}{NP}\right)^{\alpha(E/F,C;P)} \tag{4.3}$$

where $\gamma_F = \gamma + \log \varkappa_F$, with $\varkappa_F$ denoting the residue of the Dedekind zeta function $\zeta_F(s)$ at $s = 1$, and

$$\alpha(E/F,C;P) = \begin{cases} -1, & P \mid \Delta, \\ \frac{|G|}{|C|} - 1, & \mathrm{Frob}_P = C, \\ -1, & \mathrm{Frob}_P \neq C. \end{cases} \tag{4.4}$$

The proof of Theorem 4.1 follows a similar argument as that of Williams [Wil74], by using **character orthogonality** and studying several Euler products, including the **Artin L-function** of E/F. Some care must be taken when G has representations of dimension greater than one. The description of the constant (4.3) is analogous to the description given in [LZ07].

We conclude the paper with several applications: explicit descriptions of the constants when E/F is quadratic, abelian, or an S_3 sextic, and an application to **primes represented by quadratic forms**. We highlight the latter application here. If Q is a binary quadratic form, let $\mathcal{Q}$ be the set of primes represented by Q, and $\mathcal{Q}(x)$ the subset of such primes $p \leq x$.

Corollary 4.2 (Arango-Piñeros–Keliher–K., [APnKK22, Corollary 4.2])**.** *Let Q be a primitive, irreducible, positive definite, integral binary quadratic form with discriminant D. Let E be the ring class field of the order of D. Then*

$$\prod_{P\in\mathcal{Q}(x)}\left(1-\frac{1}{p}\right)=\left(\frac{e^{-\gamma(E/\mathbb{Q},C)}}{\log x}\right)^{\frac{|C|}{2h(D)}}\prod_{\substack{p\mid\Delta_E\\ p\in\mathcal{Q}}}\left(1-\frac{1}{p}\right)+O\left(\frac{1}{(\log x)^{1+\frac{|C|}{2h(D)}}}\right),$$

where $C\subset \mathrm{Gal}(E/\mathbb{Q})$ is the conjugacy class corresponding to $\mathcal{Q}$ via class field theory and $h(D)$ is the class number of forms of discriminant D.

Table 1: Prime Number theorems vs. Mertens-type theorems

Trivial extension $E=F=\mathbb{Q}$	Prime Number theorem $\sum_{p\leq x} 1\sim\frac{x}{\log x}$	Mertens' theorem $\prod_{p\leq x}\left(1-\frac{1}{p}\right)\sim\frac{e^{-\gamma}}{\log x}$
Cyclotomic extension $E=\mathbb{Q}(\zeta_b), F=\mathbb{Q}$	Dirichlet's theorem $\sum_{\substack{p\leq x\\ p\equiv a \mod b}} 1\sim\frac{1}{\varphi(b)}\frac{x}{\log x}$	Williams' theorem [Wil74] $\prod_{\substack{p\leq x\\ p\equiv a \mod b}}\left(1-\frac{1}{p}\right)\sim\left(\frac{e^{-\gamma(a,b)}}{\log x}\right)^{1/\varphi(b)}$
Number field $E=F\supseteq\mathbb{Q}$	Laundau's theorem $\sum_{NP\leq x} 1\sim\frac{x}{\log x}$	Rosen's theorem [Ros99] $\prod_{NP\leq x}(1-\frac{1}{NP})\sim\frac{e^{-\gamma_E}}{\log x}$
Galois extension $E\supseteq F\supseteq\mathbb{Q}$	Chebotarev's theorem $\sum_{\substack{NP\leq x\\ \mathrm{Frob}_P=C}} 1\sim\frac{\lvert C\rvert}{\lvert G\rvert}\frac{x}{\log x}$	(4.2) $\prod_{\substack{NP\leq x\\ \mathrm{Frob}_P=C}}(1-\frac{1}{NP})\sim\left(\frac{e^{-\gamma(E/F,C)}}{\log x}\right)^{\lvert C\rvert/\lvert G\rvert}$

4.2. Arithmetical structures on graphs.

Let G be a connected undirected graph with n vertices labeled $v_1,\ldots,v_n$, containing no loops but possibly multiedges and use $E(G)$ to refer to the edge set of G, δ_{ij} to denote the number of edges between v_i and v_j, and $\deg v$ for the degree of the vertex v. An **arithmetical structure on** G is a pair $(\boldsymbol{r},\boldsymbol{d})\in\mathbb{N}^n\times\mathbb{N}^n$, such that $\gcd(\boldsymbol{r})=\gcd(r_1,...,r_n)=1$, satisfying the system

$$r_i d_i=\sum_{j\neq i} r_j\delta_{ij} \text{ for all } 1\leq i\leq n. \tag{4.5}$$

Note that specifying $\boldsymbol{r}$ such that $r_i \mid \sum_{j\neq i} r_j\delta_{ij}$ is sufficient to recover $\boldsymbol{d}$. Thus we may simply refer to $\boldsymbol{r}$ as an arithmetical structure on G. We use $A(G)$ to denote the set of arithmetical structures on a graph G.

While arithmetical structures are interesting as purely combinatorial objects, they were introduced to study special fibers of relative proper minimal models of curves by Lorenzini, who also proved $\#A(G)$ is finite [Lor89]. More is known about $\#A(G)$ in only a few special cases:

- if G is a path or cycle $\#A(G)$ may be computed exactly [BCC+18],
- if G is a bident graph, there are known bounds for $\#A(G)$ [ABDL+20], and
- if G is a path with one doubled edge, there are conjectured asymptotics for $\#A(G)$ [GW20].

In joint work with Tomer Reiter [KR21], we give the first known general upper bound for $\#A(G)$ in terms of only n and $\#E(G)$ [KR21].

Theorem 4.3 (Keyes–Reiter, [KR21])**.** *Let G be a connected, undirected graph on n vertices, with no loops but possible multiedges. Then the following is an upper bound for the number of arithmetical structures on G.*

$$\#A(G) \leq \frac{n!}{2} \cdot \#E\left(G\right)^{2^{n-2}-1} \cdot \#E\left(G\right)^{2^{n-1} \cdot \frac{1.538 \log(2)}{(n-1)\log(2)+\log(\log(\#E(G)))}}.$$

The proof of Theorem 4.3 involves inductively applying a construction [KR21, Construction 2.1] which associates to an arithmetical structure $(\boldsymbol{r}, \boldsymbol{d})$ on G another arithmetical structure $(\boldsymbol{r}', \boldsymbol{d}')$ on G', where G' has $n-1$ vertices. This construction generalizes the **smoothing** process used in [BCC$^+$18], [ABDL$^+$20], and [GW20]. In certain special cases, it is the inverse of Lorenzini's **blowup** construction [Lor89, 1.8], and extends observations made by Corrales and Valencia about the arithmetical structures on the **clique-star** transform [CV18].

We refine our observations further for the case where G is a graph with n vertices and m edges between every pair of distinct vertices, denoted $G = mK_n$. Our earlier construction may then be leveraged to compute $\#A_{\mathrm{dec}}(mK_n)$ exactly for $n = 3, 4, 5$ and relatively small values of m, where here $A_{\mathrm{dec}}(mK_n)$ denotes arithmetical structures with decreasing r_i values. Comparing these true values with a refined version of Theorem 4.3 [KR21, Corollary 3.4] shows an gap in effectivity increasing with both m and n.

One may find a better asymptotic upper bound for $\#A_{\mathrm{dec}}(mK_n)$ by making the connection to **Egyptian fractions**, or integer solutions to the equation

$$\frac{1}{x_1} + \cdots + \frac{1}{x_n} = \frac{1}{m}. \tag{4.6}$$

In [KR21, Theorem 4.1], we demonstrate that $A(mK_n)$ is in one-to-one correspondence with the set of solutions to (4.6) with $x_i \leq x_{i+1}$, extending an observation of Harris and Louwsma when $m = 1$ [HL20]. The upper bounds of Elsholtz and Planitzer [EP20] and Browning and Elsholtz [BE11] on nondecreasing solutions to (4.6) then give asymptotically superior upper bounds for $\#A(mK_n)$, but lack the explicit constants of our results.

5. Student involvement

My belief that including undergraduate students in the research process provides a tremendous value to both the researcher and the student comes from first hand experience. As an undergrad, I participated in a summer research program with a faculty member after my third year at Tufts University, which helped to focus my mathematical interests and opened the door to a future engaging in mathematical research.

Later, as a PhD student at Emory University, I started a **Directed Reading Program** (DRP), which pairs undergrads with a graduate student to learn an advanced topic together. While not performing original research, the students develop their ability to learn independently and practice how to present a topic to a mathematical audience. Through this one-on-one mentorship with a graduate student, my hope is that the DRP provides students considering an academic career with insight into the day-to-day life and challenges of a PhD student. Meanwhile, the graduate student mentors gain not only mentorship experience but also strengthen their own grasp of the topic.

Moving forward, I hope to develop undergraduate researchers by leading advanced reading courses, serving on thesis committees, and involving students in my research program directly. Given the breadth of my research interests, I expect to find several suitably accessible problems of interest to an undergraduate student curious about number theory and arithmetic geometry. See three potential examples below.

- Compute (to a high degree of precision) the constant $e^{-\gamma(E/F,C)}$ in (4.2) and/or make the second order term explicit. This would require the student to learn some analytic number theory techniques and likely involve some computation.
- For a specific family of graphs G_n, improve Theorem 4.3 by giving a sharper upper bound on $\#A(G_n)$. Better yet, determine (or bound) the asymptotic growth rate of $\#A(G_n)$ (see [BCC$^+$18], [ABDL$^+$20], [GW20] for the cases of paths and cycles, bidents, and paths with a doubled edge).
- Use a computer algebra software (such as `Magma` or `Sage`) to enumerate (or disprove the existence of) g^1_d's on $X_0(N)$ or other modular curves, thereby verifying and/or extending the results of [NO22].

References

[ABDL$^+$20] Kassie Archer, Abigail C. Bishop, Alexander Diaz-Lopez, Luis D. García Puente, Darren Glass, and Joel Louwsma. Arithmetical structures on bidents. *Discrete Math.*, 343(7):111850, 23, 2020.

[AH91] Dan Abramovich and Joe Harris. Abelian varieties and curves in $W_d(C)$. *Compositio Math.*, 78(2):227–238, 1991.

[APnKK22] Santiago Arango-Piñeros, Daniel Keliher, and Christopher Keyes. Mertens' theorem for Chebotarev sets. *Int. J. Number Theory*, 18(8):1823–1842, 2022.

[Bar99] Francesc Bars. Bielliptic modular curves. *Journal of Number Theory*, 76(1):154–165, 1999.

[BBL16] M. J. Bright, T. D. Browning, and D. Loughran. Failures of weak approximation in families. *Compos. Math.*, 152(7):1435–1475, 2016.

[BCC$^+$18] Benjamin Braun, Hugo Corrales, Scott Corry, Luis David García Puente, Darren Glass, Nathan Kaplan, Jeremy L. Martin, Gregg Musiker, and Carlos E. Valencia. Counting arithmetical structures on paths and cycles. *Discrete Math.*, 341(10):2949–2963, 2018.

[BCF16] Manjul Bhargava, John Cremona, and Tom Fisher. The proportion of plane cubic curves over $\mathbb{Q}$ that everywhere locally have a point. *Int. J. Number Theory*, 12(4):1077–1092, 2016.

[BCF21] Manjul Bhargava, John Cremona, and Tom Fisher. The proportion of genus one curves over $\mathbb{Q}$ defined by a binary quartic that everywhere locally have a point. *Int. J. Number Theory*, 17(4):903–923, 2021.

[BE11] T. D. Browning and C. Elsholtz. The number of representations of rationals as a sum of unit fractions. *Illinois J. Math.*, 55(2):685–696 (2012), 2011.

[BGK20] Francesc Bars, Josep González, and Mohamed Kamel. Bielliptic quotient modular curves with N square-free. *J. Number Theory*, 216:380–402, 2020.

[BGR19] Francesc Bars and Josep González Rovira. Bielliptic modular curves $X_0^*(N)$ with square-free levels. *Math. Comp.*, 88(320):2939–2957, 2019.

[BGW17] Manjul Bhargava, Benedict H. Gross, and Xiaoheng Wang. A positive proportion of locally soluble hyperelliptic curves over $\mathbb{Q}$ have no point over any odd degree extension. *J. Amer. Math. Soc.*, 30(2):451–493, 2017. With an appendix by Tim Dokchitser and Vladimir Dokchitser.

[BK21a] Lea Beneish and Christopher Keyes. Fields generated by points on superelliptic curves. https://arxiv.org/abs/2103.16672, 2021.

[BK21b] Lea Beneish and Christopher Keyes. On the proportion of locally soluble superelliptic curves. https://arxiv.org/abs/2111.04697, 2021.

[BKS22] Francesc Bars, Mohamed Kamel, and Andreas Schweizer. Bielliptic quotient modular curves of $X_0(N)$. https://arxiv.org/abs/2208.10957, 2022.

[Bro17] T. D. Browning. Many cubic surfaces contain rational points. *Mathematika*, 63(3):818–839, 2017.

[Cre13] Brendan Creutz. Explicit descent in the Picard group of a cyclic cover of the projective line. In *ANTS X—Proceedings of the Tenth Algorithmic Number Theory Symposium*, volume 1 of *Open Book Ser.*, pages 295–315. Math. Sci. Publ., Berkeley, CA, 2013.

[Cre20] Brendan Creutz. Generalized Jacobians and explicit descents. *Math. Comp.*, 89(323):1365–1394, 2020.

[CV18] Hugo Corrales and Carlos E. Valencia. Arithmetical structures on graphs. *Linear Algebra Appl.*, 536:120–151, 2018.

[DF93] Olivier Debarre and Rachid Fahlaoui. Abelian varieties in $W_d^r(C)$ and points of bounded degree on algebraic curves. *Compositio Math.*, 88(3):235–249, 1993.

[Eke91] Torsten Ekedahl. An infinite version of the Chinese remainder theorem. *Comment. Math. Univ. St. Paul.*, 40(1):53–59, 1991.

[EP20] Christian Elsholtz and Stefan Planitzer. The number of solutions of the Erdős-Straus equation and sums of k unit fractions. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics*, 150(3):1401–1427, 2020.

[FH99] Masahiro Furumoto and Yuji Hasegawa. Hyperelliptic quotients of modular curves $X_0(N)$. *Tokyo J. Math.*, 22(1):105–125, 1999.

[GM19] Joseph Gunther and Jackson S. Morrow. Irrational points on hyperelliptic curves. https://arxiv.org/abs/1709.02041, 2019.

[GW20] Darren Glass and Joshua Wagner. Arithmetical structures on paths with a doubled edge. *Integers*, 20:Paper No. A68, 18, 2020.

[Has95] Yuji Hasegawa. Table of quotient curves of modular curves $X_0(N)$ with genus 2. *Proceedings of the Japan Academy, Series A, Mathematical Sciences*, 71(10):235 – 239, 1995.

[Has97] Yuji Hasegawa. Hyperelliptic modular curves $X_0^*(N)$. *Acta Arith.*, 81(4):369–385, 1997.

[HL20] Zachary Harris and Joel Louwsma. On arithmetical structures on complete graphs. *Involve*, 13(2):345–355, 2020.

[HS91] Joe Harris and Joe Silverman. Bielliptic curves and symmetric products. *Proc. Amer. Math. Soc.*, 112(2):347–356, 1991.

[HS99] Yuji Hasegawa and Mahoro Shimura. Trigonal modular curves. *Acta Arith.*, 88(2):129–140, 1999.

[Jeo18] Daeyeol Jeon. Bielliptic modular curves $X_0^+(N)$. *J. Number Theory*, 185:319–338, 2018.

[Jeo21] Daeyeol Jeon. Modular curves with infinitely many cubic points. *J. Number Theory*, 219:344–355, 2021.

[JP05] Daeyeol Jeon and Euisung Park. Tetragonal modular curves. *Acta Arith.*, 120(3):307–312, 2005.

[Key22] Christopher Keyes. Growth of points on hyperelliptic curves over number fields. *J. Théor. Nombres Bordeaux*, 34(1):271–294, 2022.

[KR21] Christopher Keyes and Tomer Reiter. Bounding the number of arithmetical structures on graphs. *Discrete Mathematics*, 344(9):112494, 2021.

[KV22] Borys Kadets and Isabel Vogt. Subspace configurations and low degree points on curves. https://arxiv.org/abs/2208.01067, 2022.

[Leb07] Philippe Lebacque. Generalised Mertens and Brauer-Siegel theorems. *Acta Arithmetica*, 130(4):333–350, 2007.

[Lor89] Dino J. Lorenzini. Arithmetical graphs. *Math. Ann.*, 285(3):481–501, 1989.

[LT19] Robert J. Lemke Oliver and Frank Thorne. Rank growth of elliptic curves in nonabelian extensions. *Int. Math. Res. Not. IMRN*, Dec. 2019.

[LT22] Robert J. Lemke Oliver and Frank Thorne. Upper bounds on number fields of given degree and bounded discriminant. *Duke Mathematical Journal*, pages 1 – 11, 2022.

[LZ07] Alessandro Languasco and Alessandro Zaccagnini. A note on Mertens' formula for arithmetic progressions. *Journal of Number Theory*, 127:37–46, 2007.

[Mer74] Franz Mertens. Ein beitrag zur analytischen zahlentheorie. *J. reine angew. Math.*, 78:46–62, 1874.

[MR18] Barry Mazur and Karl Rubin. Diophantine stability. *Amer. J. Math.*, 140(3):571–616, 2018. With an appendix by Michael Larsen.

[MV06] Hugh Montgomery and Robert Vaughan. *Multiplicative Number Theory.* Number 97 in Cambridge studies in advanced mathematics. Cambridge University Press, 2006.

[NO22] Filip Najman and Petar Orlić. Gonality of the modular curve $X_0(N)$. https://arxiv.org/abs/2207.11650, 2022.

[Ogg74] Andrew P. Ogg. Hyperelliptic modular curves. *Bull. Soc. Math. France*, 102:449–462, 1974.

[PS99a] Bjorn Poonen and Michael Stoll. The Cassels-Tate pairing on polarized abelian varieties. *Ann. of Math. (2)*, 150(3):1109–1149, 1999.

[PS99b] Bjorn Poonen and Michael Stoll. A local-global principle for densities. In *Topics in number theory (University Park, PA, 1997)*, volume 467 of *Math. Appl.*, pages 241–244. Kluwer Acad. Publ., Dordrecht, 1999.

[Ros99] Michael Rosen. A generalization of Mertens' theorem. *Journal of the Ramanujan Mathematical Society*, 14:1–20, 1999.

[Sag21] Sage Developers. *SageMath, the Sage Mathematics Software System (Version 8.2)*, 2021. https://www.sagemath.org.

[Wil74] Kenneth S. Williams. Mertens' theorem for arithmetic progressions. *J. Number Theory*, 6:353–359, 1974.